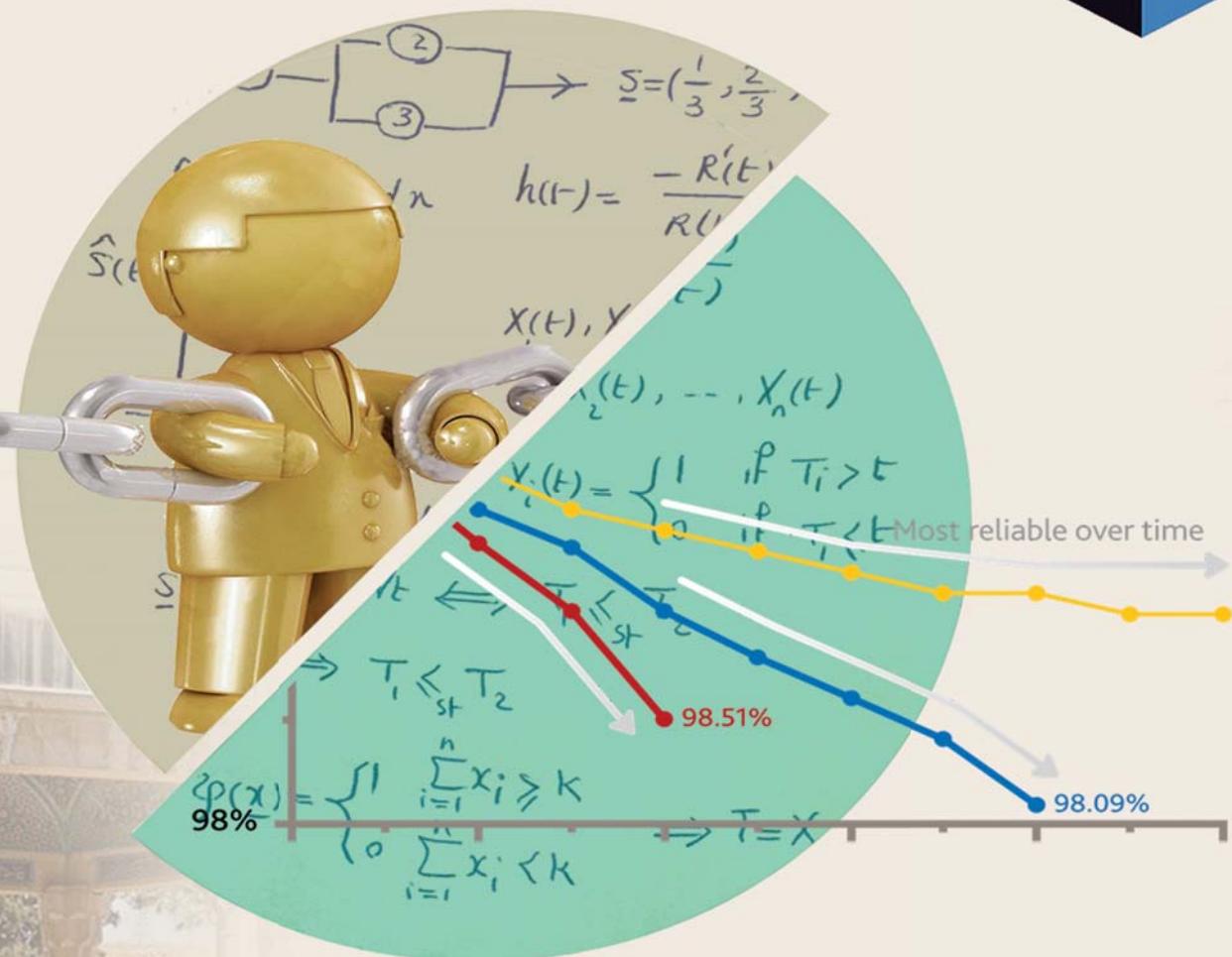


# Proceeding of the 4th Seminar on Reliability Theory

and its Applications

DEPARTMENT OF STATISTICS, SHIRAZ UNIVERSITY

25-26 April 2018



In the Name of God



**Proceeding of  
the 4th Seminar on  
Reliability Theory and its Applications**

**Department of Statistics,  
Shiraz University,  
Shiraz, Iran**

**25-26 April, 2018**

## Preface

Following the series of workshops on “Reliability Theory and its Applications” in Ferdowsi University of Mashhad and three seminars in University of Isfahan (2015), University of Tehran (2016) and Ferdowsi University of Mashhad (2017) we are pleased to organize the 4th Seminar on “**Reliability Theory and its Applications**” during 25-26 April, 2018 at the Department of Statistics, Shiraz University. On behalf of the organizing and scientific committees, we would like to extend a very warm welcome to all participants, hoping that their stay in Shiraz will be happy and fruitful. Hope that this seminar provides an environment of useful discussions and would also exchange scientific ideas through opinions. We wish to express our gratitude to the numerous individuals that have contributed to the success of this seminar, in which around 90 colleagues, researchers, and postgraduate students from universities and organizations have participated.

Finally, we would like to extend our sincere gratitude to the Research Council of Shiraz University, the administration of College of Sciences, the Ordered and Spatial Data Center of Excellence, the Islamic World Science Citation Center, the Fars Science and Technology Park, the Iranian Statistical Society, the Scientific Committee, the Organizing Committee, the referees, and the students and staff of the Department of Statistics at Shiraz University for their kind cooperation.

Somayeh Zarezadeh (Chair)

April, 2018

## Topics

The aim of the seminar is to provide a forum for presentation and discussion of scientific works covering theories and methods in the field of reliability and its application in a wide range of areas:

- Accelerated life testing
- Bayesian methods in reliability
- Case studies in reliability analysis
- Computational algorithms in reliability
- Data mining in reliability
- Degradation models
- Lifetime data analysis
- Lifetime distributions theory
- Maintenance modeling and analysis
- Networks reliability
- Optimization methods in reliability
- Reliability of coherent systems
- Safety and risk assessment
- Software reliability
- Stochastic aging
- Stochastic dependence in reliability
- Stochastic orderings in reliability
- Stochastic processes in reliability
- Stress-strength modeling
- Survival analysis

## Scientific Committee

1. Ahmadi, J., Ferdowsi University of Mashhad (Iran)
2. Asadi, M., University of Isfahan (Iran)

3. Asgharzadeh, A., Mazandaran University (Iran)
4. Behboodian, J., Shiraz University (Iran)
5. Gaudoin, O., Université Grenoble Alpes (France)
6. Haghghi, F., University of Tehran (Iran)
7. Hamadani, A. Z., Isfahan University of Technology (Iran)
8. Khaledi, B., Razi University (Iran)
9. Khanjari, M., Birjand University (Iran)
10. Kharrati-Kopaei, M., Shiraz University (Iran)
11. Khodadadi, A., Shahid Beheshti University (Iran)
12. Khorram, E., Amirkabir University of Technology (Iran)
13. Nematollahi, A. R., Shiraz University (Iran)
14. Ross, S., University of Southern California (USA)
15. Soofi, E. S., University of Wisconsin-Milwaukee (USA)
16. Towhidi, M., Shiraz University (Iran) (Chair)
17. Zarezadeh, S., Shiraz University (Iran)

## **Organizing Committee**

1. Borhani-Haghghi, A., Shiraz University
2. Jabbari Nooghabi, H., Ferdowsi University of Mashhad
3. Kharati-Kopaei, M., Shiraz University
4. Mirzaei, B., Shiraz University
5. Nematollahi, A. R., Shiraz University
6. Zarezadeh, S., Shiraz University (Chair)

# Contents

<b>A test for exponentiality based on gamma-divergence</b>	
Ahrari, V., Habibirad, A., Baratpour, S. ....	1
<b>Stochastic comparisons of mixture of coherent systems with dependent component lifetimes</b>	
Amini-Seresht, E. ....	12
<b>Regression analysis of clustered interval-censored failure time data with the proportional hazards model</b>	
Amoorahimi, E., Kelkinnama, M. ....	16
<b>Three-state systems with different types of components</b>	
Ashrafi,S. ....	25
<b>Planning of step-stress accelerated degradation test with tampered failure rate model under an inverse Gaussian process</b>	
Azizi, F., Mosayebi Omshi, E., Haghghi, F., Shemehsavar, S. ....	33
<b>Cumulative residual inaccuracy for minimum and maximum of order statistics</b>	
Daneshi, S., Nezakati, A., Tahmasebi, S. ....	41
<b><math>P(Y &lt; X)</math> using generalized order statistics and concomitant</b>	
Domma, F., Eftekharian, A., Razmkhah, M. ....	48
<b>An imputation approach to Cox proportional hazards model in the presence of interval-censored data</b>	
Fani,S. ....	56
<b>Statistical inferences for stress-strength in the Marshal-Olkin models based on progressively type-II censored samples</b>	
Ghanbari, S., Rezaei, A., Salehi, M. ....	64

<b>Comparison of equivalency and optimality between constant-stress and step-stress tests under type-I censoring for Rayleigh distri- bution</b>	
Hakamipour, N. ....	75
<b>Cumulative residual entropy in sequential order statistics and some characterizations</b>	
Hashempour, M., Doostparast, M. ....	84
<b>Some results on proportional mean past lifetime frailty model</b>	
Hooti, F., Ahmadi, J. ....	90
<b>On mean time to failure in age replacement</b>	
Izadi, M., Khaledi, B. ....	97
<b>Reliability estimation for the two-parameter exponential distribution based on records</b>	
Jafari, A. A., Bafecri Fadafan, S. ....	99
<b>Multi-objective redundancy allocation problem with entropy constraint using two approaches</b>	
Kamandi, R., Zarezadeh, S., Mirdehghan, M. ....	104
<b>Statistical inference for a repairable system with Weibull distribution subject to shocks</b>	
Kamranfar, H., Etminan, J., Chahkandi, M. ....	113
<b>On the coherent systems with two different types of dependent com- ponents</b>	
Kelkinnama, M. ....	121
<b>Testing exponentiality based on the residual Lin-Wong divergence</b>	
Khalili, M., Habibirad, A., Yousefzadeh, F. ....	128
<b>Some results on comparison among coherent systems</b>	
Khanjari Sadegh, M., Mirjalili, A. ....	137
<b>On the point reconstruction of order statistics from geometric distri- bution</b>	
Khatib Astaneh, B. ....	144

Bayesian inference for $R = P(X < Y)$ under adaptive type-II hybrid progressive censored samples in generalized inverted exponential distribution	
Kohansal, A. ....	151
An additive-multiplicative mean past life regression model	
Mansourvar, Z. ....	160
Parametric inference from system lifetime data under a proportional hazard rate model based on the minimum Jensen-Gini estimator	
Mehrali, Y. ....	167
Survival analysis of dependent competing risks with masked failure causes	
Mireh, S., Khodadadi, A., Haghghi, F. ....	176
Bi-objective reliability-redundancy allocation problem with cold standby strategy	
Mohammadi, Z., Zarezadeh, S., Dehbozorgi, M. ....	184
Bayesian designing of step-stress accelerated degradation test under an inverse Gaussian process with tampered failure rate model	
Mosayebi, E., Azizi, F., Shemehsavar, S., Haghghi, F. ....	192
Connections between diversion concepts and reliability measures	
Nasr Esfahani, S., Asadi, M. ....	201
Reliability analysis of $k$ -out-of- $n:F$ system with soft and hard failures	
Nezakati, E., Razmkhah, M. ....	209
On a goodness of fit test for normality based on Lin-Wong divergence with type-I censored data	
Pakgohar, A., Habibirad, A., Yousefzadeh, F. ....	217
Bayes estimation of mean lifetime in a proportional hazard rate model under progressive censoring	
Razmkhah, M., Khatib, B. ....	226
Application of skip-free markov chains for degradation	
Sajadi, F.A. ....	233

<b>On the reliability modeling of weighted <math>k</math>-out-of-<math>n</math> systems with randomly chosen components</b>	
Salehi, M., Shishebor, Z., Asadi, M. ....	238
<b>Bayesian and non-bayesian reliability analysis of the exponential family based on dual generalized order statistics</b>	
Salim, N., Dordizadeh, H. ....	246
<b>Optimal step stress test plan based on bivariate gamma degradation model</b>	
Shemehsavar, S., Mosaeby Omshi, E. ....	254
<b>A new class of bivariate failure time distributions in shock and competing risk models</b>	
Shoaee, Sh. ....	261
<b>On the behavior of the system failure rate</b>	
Tavangar, M. ....	270
<b>On the estimation of stress strength reliability parameter of new Burr distribution</b>	
Yari, G.H., Tondpour, Z. ....	278
<b>An RSS-based estimator of mean residual life</b>	
Zamanzade, E., Parvardeh, A., Asadi, M. ....	285



Forth Seminar on  
**Reliability Theory and its Applications**

April 25-26, 2018



## A test for exponentiality based on gamma-divergence

Ahrari, V.<sup>1</sup> Habibirad, A.<sup>2</sup> Baratpour, S.<sup>3</sup>

<sup>1,2,3</sup> Department of Statistics, Ferdowsi University of Mashhad

### Abstract

In this paper, first we use the already defined gamma-divergence and then introduce a goodness of fit test for exponentiality. This divergence measure is very robust with respect to outliers. In order to construct test statistics, two estimators are used as estimators of gamma-divergence. In the first one we consider the gamma-divergence of the equilibrium distribution function, which is well defined on the empirical distribution function (EDF) and is proposed as an EDF-based goodness of fit test statistic. The second one is an estimator in manner of Vasicek entropy estimator. Critical values of test statistics are computed by Monte Carlo simulations. Then, the power values of the proposed tests are compared with some known competing tests under various alternatives. Simulation results indicate that in comparison with the other tests statistics, our mentioned test statistics almost in most of the cases have higher power. Finally, an example containing outliers illustrate the importance and use of the proposed tests.

**Keywords:** Entropy estimator, Equilibrium distribution, Ggamma-divergence, Goodness of fit test, Outliers, Testing exponentiality.

## 1 Introduction

Suppose that  $X$  is a non-negative continuous random variable with cumulative distribution function (cdf)  $F(x)$  and probability density function (pdf)  $f(x)$ . Consider the following hypotheses

$$H_0 : f(x) = f_0(x, \theta) \quad v.s. \quad H_1 : f(x) \neq f_0(x, \theta),$$

---

<sup>1</sup>v.ahrari84@mail.um.ac.ir

where  $f_0(x, \theta) = \frac{1}{\theta}e^{-\frac{x}{\theta}}$ ,  $x > 0$ ,  $\theta > 0$ , and  $\theta$  is unknown. Testing for exponentiality has drawn the attention of many investigators, such that there are several goodness of fit tests in the literature for the exponential distribution. Many authors including Lilliefors (1969), Van-Soest (1969), Finkelstein and Schafer (1971), Stephens (1974), Harris (1976), Ebrahimi et al. (1992), Ahmad and Alwasel (1999), Alwasel (2001), Abbasnejad (2011), Baratpour and Habibirad (2012) and Park et al. (2014) presented different tests for exponentiality.

Entropy is an important concept in many areas of research such as physics, probability and statistics, communication theory and economics. An early definition of a measure of uncertainty is the Shannon entropy (Shannon, 1948). The Shannon entropy of  $X$  is defined as

$$H(f) = - \int_0^\infty f(x) \log f(x) dx.$$

There is an extensive literature on estimating the Shannon entropy nonparametrically. Many researchers, such as Vasicek (1976), Van Es (1992), Ebrahimi et al. (1994), Correa (1995), Alizadeh Noughabi (2010) and others have been estimated the entropy  $H(f)$  for absolutely continuous random variables. The Vasiceks estimator has been most widely used in developing entropy-based statistical procedures among other entropy estimators discussed in the literature and is given by

$$H_{mn} = \frac{1}{n} \sum_{i=1}^n \log \left\{ \frac{n}{2m} (X_{i+m:n} - X_{i-m:n}) \right\},$$

where  $m$  is a positive integer smaller than  $n/2$  which is called window size and  $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$  are order statistics from the sample  $X_1, X_2, \dots, X_n$ .  $X_{i:n} = X_{1:n}$  if  $i < 1$  and  $X_{i:n} = X_{n:n}$  if  $i > n$ . Vasicek (1976) established the consistency of  $H_{mn}$  for the population entropy  $H(f)$  and used  $H_{mn}$  as an estimator of  $H(f)$ .

The Kullback-Leibler (KL) divergence (Kullback and Leibler, 1951 and Kullback, 1959) is a non-symmetric measure of the difference between two pdfs  $f$  and  $g$ , and is defined as

$$I(f, g) = \int_0^\infty f(x) \log \frac{f(x)}{g(x)} dx.$$

KL divergence is non-negative and the equality to zero holds iff  $f(x) = g(x)$ . Tests of fit based on KL divergence have been developed, see Ebrahimi et al. (1992), Choi et al. (2004) and Gurevich and Davidson (2008).

Considering the survival function instead of the pdf in the Shannon entropy leads to a new entropy measure named cumulative residual entropy (CRE) that Rao et al. (2004) introduced as

$$CRE(X) = - \int_0^\infty \bar{F}(x) \log \bar{F}(x) dx,$$

where  $\bar{F}(x) = 1 - F(x)$ . Baratpour and Habibirad (2012) introduced a new measure of distance between two non-negative and continuous distributions based on the CRE and called it cumulative residual KL (CRKL) as below

$$CRKL(F : G) = \int_0^\infty \bar{F}(x) \log \frac{\bar{F}(x)}{\bar{G}(x)} dx - E(X) + E(Y).$$

They used CRKL for testing exponentiality.

The KL divergence has been generalized by using a family of functions called generalized logarithm functions or  $\alpha$ -logarithm that is define as

$$\log_\alpha(x) = \frac{1}{1-\alpha}(x^{1-\alpha} - 1), \quad x > 0, \quad (1)$$

which is a power function of  $x$  with power  $1 - \alpha$ , and is the natural logarithm function as  $\alpha \rightarrow 1$ . Often, the power function (1) allows to generate more robust divergences in respect to outliers and consequently better or more flexible performance (see, for example, Eguchi and Kato, 2010). By using this type of extension, the gamma-divergence which is generalization of the KL-divergence is obtained.

In section 2, we introduce the gamma-divergence along with its basic properties. Then, we use this divergence for constructing distance between two equilibrium distributions. By the gamma-divergence, we construct test statistics for testing exponentiality based on equilibrium distribution function and Vasicek's method in section 3. In section 4, we compare the power of the proposed tests with the other tests under several alternative distributions by a simulation study. The performance of the considered tests for a real data is evaluated in section 5. Finally, some concluding remarks are given in section 6.

## 2 Gamma-divergence

Gamma-divergence proposed by Jones et al. (2001) and Fujisawa and Eguchi (2008) as a very robust similarity measure respect to outliers. This divergence is defined as (Cichocki and Amari, 2010):

$$D_G^\gamma(p \parallel q) = \frac{1}{\gamma(\gamma-1)} \log \left( \frac{\left( \int p^\gamma(x) d\mu(x) \right) \left( \int q^\gamma(x) d\mu(x) \right)^{\gamma-1}}{\left( \int p(x) q^{\gamma-1}(x) d\mu(x) \right)^\gamma} \right), \quad \gamma \in \mathbb{R} \setminus \{0, 1\},$$

where  $p(x)$  and  $q(x)$  do not need to be normalized and are positive measures (densities).

### 2.1 Gamma-divergence based on equilibrium distribution

Suppose that  $X$  is a non-negative random variable with pdf  $f$ , then the pdf of the weighted random variable  $X^*$  is given by

$$f^*(t) = \frac{w(t)f_X(t)}{E(w(X))}, \quad x \geq 0,$$

where  $w(t)$  is a non-negative weight function and  $0 < E(w(X)) < \infty$ . If  $w(t) = \frac{1}{r_X(t)}$ , where  $r_X(t) = \frac{f_X(t)}{F_X(t)}$  is failure rate function of  $X$ , then  $X^*$  is said the equilibrium random variable associate to  $X$ . The pdf of  $X^*$  is  $f^*(t) = \frac{\bar{F}_X(t)}{E(X)}$  (see Andrews and Andrews, 1962 and Nakamura, 2009). The following Theorem says that the cdf can be characterized by associated equilibrium distribution.

**Theorem 2.1.** Let  $X$  and  $Y$  be two random variables with common support  $\chi = [0; \infty)$  and expectations  $E(X) < \infty$  and  $E(Y) < \infty$ , pdfs  $f$  and  $g$  and cdfs  $F$  and  $G$ , respectively. If  $X^*$  and  $Y^*$  are the equilibrium random variables associated to  $X$  and  $Y$  respectively, then  $X \stackrel{d}{=} Y \iff X^* \stackrel{d}{=} Y^*$ , where  $\stackrel{d}{=}$  stands for equality in distribution.

Let  $f^*$  and  $g^*$  be the equilibrium pdfs associated to  $f$  and  $g$ , respectively. Then, we define the gamma-divergence based on equilibrium distributions as follows

$$\begin{aligned} D_G^\gamma(f^*, g^*) &= \frac{1}{\gamma(\gamma-1)} \log \int_0^\infty \frac{\bar{F}^\gamma(x)}{E^\gamma(X)} dx + \frac{1}{\gamma} \log \int_0^\infty \frac{\bar{G}^\gamma(x)}{E^\gamma(Y)} dx \\ &\quad - \frac{1}{(\gamma-1)} \log \int_0^\infty \frac{\bar{F}(x)}{E(X)} \frac{\bar{G}^{\gamma-1}(x)}{E^{\gamma-1}(Y)} dx. \end{aligned} \quad (2)$$

### 3 Testing procedures based on the estimating gamma-divergence

In this section, by utilizing equilibrium distribution function and Vasicek's method in the estimating gamma-divergence, we construct test statistics for testing exponentiality and then consider some competing tests to compare with the proposed tests.

#### 3.1 Test statistics

Suppose  $X_1, X_2, \dots, X_n$  is a random sample from an unknown continuous distribution  $F$ . Consider the null hypothesis  $H_0 : f(x) = f_0(x, \theta)$  where  $f_0(x, \theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}$ ,  $x > 0$ ,  $\theta > 0$ , and  $\theta$  is unknown. The alternative hypothesis is  $H_1 : f(x) \neq f_0(x, \theta)$ . To test exponentiality, we construct and evaluate a test based on the sample EDF  $F_n(x)$  by  $D_G^\gamma(f^*, g^*)$ . We let  $F(x) = F_n(x)$  and  $G(x) = F_0(x, \theta)$  in (2) and estimate the unknown parameter  $\theta$  with the maximum likelihood estimate,  $\bar{x}_n$ . After simplification, we obtain the following test statistic for testing exponentiality:

$$\begin{aligned} T_{\gamma,n} &= \frac{1}{\gamma(\gamma-1)} \log \left[ \frac{\sum_{i=0}^{n-1} \left( \frac{n-i}{n} \right)^\gamma (x_{i+1:n} - x_{i:n})}{\bar{x}_n} \right] - \frac{1}{\gamma} \log \gamma \\ &\quad - \frac{1}{\gamma-1} \log \left[ \sum_{i=0}^{n-1} \left( \frac{n-i}{n} \right) \left( e^{-\frac{x_{i:n}(\gamma-1)}{\bar{x}_n}} - e^{-\frac{x_{i+1:n}(\gamma-1)}{\bar{x}_n}} \right) \right] + \frac{1}{\gamma-1} \log(\gamma-1), \end{aligned} \quad (3)$$

where  $x_{0:n} = 0$ . The test statistic (3) is invariant with respect to the scale transformations.

Now, we wish to estimate  $D_G^\gamma(f, g)$  similar to Vasicek's method, first by changing variable  $F(x) = p$  and considering  $g(x) = f_0(x, \theta)$  in  $D_G^\gamma(f, g)$ , we have

$$\begin{aligned} D_G^\gamma(f, f_0) &= \frac{1}{\gamma(\gamma-1)} \log \left[ \int_0^1 \left( \frac{d}{dp} F^{-1}(p) \right)^{-(\gamma-1)} dp \right] + \left( \frac{1-\gamma}{\gamma} \right) \log \theta - \left( \frac{1}{\gamma} \right) \log \gamma \\ &\quad + \log \theta - \frac{1}{(\gamma-1)} \log \int_0^1 e^{-\frac{F^{-1}(p)}{\theta}(1-\alpha)} dp. \end{aligned} \quad (4)$$

The estimator of (4) can be constructed by replacing the distribution function  $F$  by the EDF  $F_n$  and using a difference operator instead of the differential operator. Therefore, we have

$$\begin{aligned} f(x) &= \left( \frac{d}{dp} F^{-1}(p) \right)^{-1} \\ &\simeq \left( \frac{n}{2m} (x_{i+m:n} - x_{i-m:n}) \right)^{-1}. \end{aligned}$$

So, we obtain another test statistic for testing exponentiality which can be written as

$$\begin{aligned} T_{\gamma,mn} &= \frac{1}{\gamma(\gamma-1)} \log \left[ \frac{\bar{x}_n^{\gamma-1}}{n} \sum_{i=1}^n \left( \frac{x_{i+m:n} - x_{i-m:n}}{\frac{2m}{n}} \right)^{(1-\gamma)} \right] \\ &- \frac{1}{\gamma-1} \log \left[ \frac{1}{n} \sum_{i=1}^n e^{-\frac{(\gamma-1)x_{i:n}}{\bar{x}_n}} \right] - \frac{1}{\gamma} \log \gamma. \end{aligned} \quad (5)$$

Note that  $T_{\gamma,mn}$  is also invariant with respect to scale transformations.

### 3.2 Competitor tests

We compare the performance of the proposed tests with some well-known EDF-based and entropy-based tests.

The Van-Soest ( $W^2$ ) (1969) and Finkelstein and Schafer ( $S^*$ ) (1971) statistics based on EDF are respectively as follows

$$\begin{aligned} W^2 &= \sum_{i=1}^n \left\{ F_0(x_{i:n}, \hat{\theta}) - \frac{2i-1}{2n} \right\}^2 + \frac{1}{12n}, \\ S^* &= \sum_{i=1}^n \max \left\{ \left| F_0(x_{i:n}, \hat{\theta}) - \frac{i}{n} \right|, \left| F_0(x_{i:n}, \hat{\theta}) - \frac{i-1}{n} \right| \right\}, \end{aligned}$$

where  $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}_n$ .

Choi et al. (2004) discussed goodness of fit tests of the exponential distribution based on KL-divergence for testing  $H_0$  against  $H_1$  as bellow

$$KLC_{mn} = \frac{\exp(C_{mn})}{\exp(\ln \bar{x}_n + 1)},$$

where  $C_{mn} = -\frac{1}{n} \sum_{i=1}^n \log \left\{ \frac{\sum_{j=i-m}^{i+m} (x_{j:n} - \bar{x}_i)(j-i)}{n \sum_{j=i-m}^{i+m} (x_{j:n} - \bar{x}_i)^2} \right\}$  and  $\bar{x}_i = \frac{\sum_{j=i-m}^{i+m} x_{j:n}}{2m+1}$ . For  $KLC_{mn}$  statistics,  $x_{j:n} = x_{1:n}$  if  $j < 1$  and  $x_{j:n} = x_{n:n}$  if  $j > n$  and the windows size  $m$  is a positive integer smaller than  $\frac{n}{2}$ .  $H_0$  is rejected for large values of  $W^2$  and  $S^*$  and for small value of  $KLC_{mn}$ .

The test statistic proposed by Baratpour and Habibirad (2012) is

$$T_n = \frac{\sum_{i=1}^n \frac{n-i}{n} (\log \frac{n-i}{n})(x_{i+1:n} - x_{i:n})}{\frac{\sum_{i=1}^n x_i^2}{2 \sum_{i=1}^n x_i}} + 1.$$

The test statistic proposed by Park et al. (2014) based on equilibrium distribution is

$$T_1 = \frac{1}{\bar{x}_n} \sum_{i=1}^n \frac{n-i}{n} \left( \log \frac{n-i}{n} \right) (x_{i+1:n} - x_{i:n}) + \frac{1}{2n\bar{x}_n^2} \sum_{i=1}^n x_i^2.$$

The  $KLC_{mn}$ ,  $T_n$  and  $T_1$  statistics are based on entropy.

## 4 Simulation study

Large values of  $T_{\gamma,n}$  and  $T_{\gamma,mn}$  indicate that the sample is from a non-exponential distribution. Because the null distribution of the test statistics  $T_{\gamma,n}$  and  $T_{\gamma,mn}$  are not available, we proceed the Mont Carlo simulation to determine critical values of the test statistics. In order to obtain the critical values of proposed test statistics, 50,000 samples of size  $n$  were generated from the standard exponential distribution for selected values  $n \leq 25$ .

The power values of the proposed tests depend on the  $\gamma$  value and type of failure rate function of alternative. Also,  $T_{\gamma,mn}$  statistic depend on the window size  $m$ . Unfortunately, there is no choice criterion of  $m$ , and in general it depends on  $n$ . We determine the optimum values of  $m$  in a fixed interval  $\gamma$  values based on maximum power of the tests for some selected values of  $n$  in Table 1.

Before we present simulations results, we control type I error using the 0.95 percentiles of the proposed test statistics. To this end, we simulated random samples from standard exponential distribution and estimated type I error based on ratio rejection of null hypothesis. The result indicated that all the assumed tests have acceptable function based on type I error.

**Table 1.** Optimal values of  $m$  for some selected values of  $n$

n	5	10	20	25
m	1	2	4	4

Now, we study the power of each test defined in Subsection 3.1 by means of Monte Carlo simulations and consider the maximum power of those tests in terms of specific values of and  $\gamma$ . Alternatives are selected according to the type of their failure rate function as follows:

- Increasing failure rate (IFR): Gamma (shape parameter 2), Weibull (shape parameter 2), Beta (2, 1), EV (shape parameter 2), CH (shape parameter 1.1).
- Decreasing failure rate (DFR): Gamma (shape parameter 0.5), Weibull (shape parameter 0.5), Log-normal (shape parameter 2), Lom (shape parameter 0.5), DL (shape parameter 0).
- Non-monotone failure rate (NFR): Log-normal (shape parameter 1.2), Beta (1, 0.5), DL (shape parameter 0.5),

where

- EV(.) is Modified Extreme Value distribution with cdf,  $F(x; \theta) = 1 - \exp(\theta^{-1}(1 - e^x))$ .

- CH(.) is Chen distribution (2000) with cdf,  $F(x; \theta) = 1 - \exp(2(1 - e^{x^\theta}))$ .
- Lom(.) is Lomax distribution with cdf,  $F(x; \theta) = 1 - (1 + x)^{-(1+\theta)}$ .
- DL(.) is Dhillon distribution (1981) with cdf,  $F(x; \theta) = 1 - \exp(-(log(x + 1))^{\theta+1})$ .

As  $\gamma$  values play a pivotal role in the power of the proposed tests, we determined  $\gamma$  values that maximized power according to the type of failure rate function. The results are presented in Table 2.

Table 3 indicates the critical values of the proposed test statistics at a 5% significance level as for values of  $\gamma$  in Table 2. The values of  $m$  are calculated based on Table 1.

Tables 4, 5 and 6 present power values of the proposed tests and the competing tests at a 5% significance level based on non-entropy and entropy classes.

**Table 2.** Optimal values of  $\gamma$  according to the type of failure rate function

Type of failure rate function	$\gamma$
IFR	$\gamma\epsilon(0, 0.5]$
DFR	$\gamma\epsilon[2, 2.5]$
NFR	$\gamma\epsilon(1, 1.5]$

The estimated powers of tests with the IFR function alternatives are shown in Table 4. The results show that, almost in most cases,  $T_{\gamma,mn}$  statistic have higher power than the other proposed test and also the competing tests in entropy and non-entropy classes, so that  $T_{\gamma,mn}$  is the best for 70% of the cases (14 out of 20).

**Table 3.** Critical values of the proposed test statistics

n	Tests	$\gamma\epsilon(0, 0.5]$	$\gamma\epsilon[2, 2.5]$	$\gamma\epsilon(1, 1.5]$
5	$T_{\gamma,n}$	1.622	0.058	0.169
	$T_{\gamma,mn}$	2.965	1.369	1.406
10	$T_{\gamma,n}$	1.772	0.031	0.110
	$T_{\gamma,mn}$	1.905	0.593	0.644
20	$T_{\gamma,n}$	0.975	0.016	0.067
	$T_{\gamma,mn}$	1.382	0.283	0.358
25	$T_{\gamma,n}$	0.889	0.012	0.056
	$T_{\gamma,mn}$	1.273	0.257	0.255

In Table 5, the alternatives have DFR function. We observe that for samples with size  $n \leq 10$ , the  $T_{\gamma,n}$  is better than all of the alternatives considered in the entropy and non-entropy classes. Furthermore,  $T_{\gamma,mn}$  performs well in competition with the entropy class for sizes  $n \geq 10$  and in non-entropy class, the difference of powers of  $T_{\gamma,mn}$  is not remarkable.

The alternatives of Table 6 have the NFR function. Table 6 shows that the powers depend on the kind of alternatives distribution. So, a comprehensive conclusion can not be proposed.

**Table 4.** Comparing powers of considered tests based on non-entropy and entropy tests separately for the alternatives with the IFR function at a 5% significance level

n	Alternative (IFR)	entropy				non - entropy		
		$T_{\gamma,n}$	$T_{\gamma,mn}$	$T_n$	$T_{park}$	$KLC_{mn}$	$W^2$	$S^*$
5	Gamma(2)	0.157	0.181	0.162	0.101	0.174	0.131	0.135
	Weibull(2)	0.322	0.370	0.344	0.226	0.366	0.288	0.301
	EV(2)	0.212	0.222	0.201	0.139	0.195	0.149	0.159
	CH(1,1)	0.131	0.137	0.117	0.076	0.116	0.096	0.084
	Beta(2,1)	0.751	0.731	0.750	0.642	0.717	0.617	0.655
10	Gamma(2)	0.241	0.329	0.261	0.132	0.324	0.244	0.252
	Weibull(2)	0.581	0.717	0.647	0.424	0.691	0.611	0.626
	EV(2)	0.411	0.438	0.397	0.254	0.366	0.288	0.301
	CH(1,1)	0.206	0.228	0.208	0.104	0.194	0.149	0.152
	Beta(2,1)	0.992	0.991	0.991	0.958	0.986	0.946	0.963
20	Gamma(2)	0.344	0.501	0.388	0.192	0.502	0.484	0.492
	Weibull(2)	0.847	0.951	0.917	0.762	0.929	0.931	0.942
	EV(2)	0.718	0.763	0.720	0.521	0.618	0.554	0.574
	CH(1,1)	0.369	0.416	0.350	0.176	0.304	0.255	0.279
	Beta(2,1)	1	1	1	1	0.999	0.999	0.999
25	Gamma(2)	0.384	0.545	0.448	0.242	0.565	0.588	0.592
	Weibull(2)	0.906	0.975	0.963	0.859	0.963	0.976	0.981
	EV(2)	0.826	0.858	0.828	0.638	0.687	0.664	0.706
	CH(1,1)	0.436	0.490	0.424	0.239	0.346	0.326	0.344
	Beta(2,1)	1	1	1	1	1	0.999	0.999

**Table 5.** Comparing powers of considered tests based on non-entropy and entropy tests separately for the alternatives with the DFR function at a 5% significance level

n	Alternative (DFR)	entropy				non - entropy		
		$T_{\gamma,n}$	$T_{\gamma,mn}$	$T_n$	$T_{park}$	$KLC_{mn}$	$W^2$	$S^*$
5	Gamma(0.5)	0.251	0.083	0.012	0.127	0.011	0.146	0.154
	Weibull(0.5)	0.437	0.130	0.005	0.262	0.009	0.308	0.314
	Log-normal(2)	0.453	0.116	0.007	0.312	0.018	0.336	0.335
	Lom(0.5)	0.625	0.307	0.005	0.481	0.117	0.507	0.511
	DL(0)	0.362	0.100	0.014	0.261	0.026	0.270	0.265
10	Gamma(0.5)	0.340	0.172	0.038	0.216	0.016	0.289	0.288
	Weibull(0.5)	0.645	0.365	0.166	0.519	0.062	0.611	0.616
	Log-normal(2)	0.699	0.353	0.298	0.617	0.137	0.654	0.652
	Lom(0.5)	0.864	0.643	0.582	0.820	0.471	0.835	0.840
	DL(0)	0.589	0.293	0.268	0.531	0.125	0.519	0.518
20	Gamma(0.5)	0.474	0.346	0.139	0.349	0.093	0.524	0.518
	Weibull(0.5)	0.860	0.715	0.541	0.781	0.444	0.901	0.899
	Log-normal(2)	0.906	0.724	0.725	0.876	0.579	0.918	0.919
	Lom(0.5)	0.986	0.934	0.931	0.972	0.892	0.984	0.985
	DL(0)	0.818	0.645	0.658	0.799	0.599	0.801	0.802
25	Gamma(0.5)	0.535	0.483	0.187	0.400	0.195	0.626	0.622
	Weibull(0.5)	0.912	0.840	0.669	0.855	0.669	0.952	0.952
	Log-normal(2)	0.948	0.833	0.836	0.928	0.769	0.961	0.964
	Lom(0.5)	0.994	0.972	0.973	0.991	0.962	0.995	0.994
	DL(0)	0.876	0.765	0.759	0.865	0.601	0.869	0.873

**Table 6.** Comparing powers of considered tests based on non-entropy and entropy tests separately for the alternatives with the NFR function at a 5% significance level

n	Alternative (NFR)	entropy				non - entropy		
		$T_{\gamma,n}$	$T_{\gamma,mn}$	$T_n$	$T_{park}$	$KLC_{mn}$	$W^2$	$S^*$
5	Log-normal(1,2)	0.077	0.068	0.065	0.073	0.069	0.072	0.071
	DL(0.5)	0.107	0.066	0.050	0.077	0.053	0.083	0.070
	Beta(1,0.5)	0.506	0.458	0.572	0.504	0.517	0.417	0.441
10	Log-normal(1,2)	0.126	0.100	0.074	0.118	0.088	0.104	0.102
	DL(0.5)	0.179	0.185	0.073	0.160	0.000	0.107	0.115
	Beta(1,0.5)	0.871	0.875	0.930	0.872	0.890	0.741	0.779
20	Log-normal(1,2)	0.204	0.176	0.131	0.189	0.129	0.154	0.153
	DL(0.5)	0.286	0.272	0.162	0.261	0.000	0.179	0.177
	Beta(1,0.5)	0.998	0.999	0.999	0.998	0.998	0.973	0.979
25	Log-normal(1,2)	0.236	0.206	0.159	0.221	0.163	0.178	0.181
	DL(0.5)	0.329	0.349	0.216	0.305	0.000	0.208	0.203
	Beta(1,0.5)	0.999	0.999	0.999	0.999	0.999	0.993	0.995

## 5 Illustrative examples

In this section, an artificial example in order to investigate the importance of the proposed tests for testing the validity of exponential distribution is presented. Since the p-value of the proposed tests depends on the  $\gamma$  values, it is calculated for specific values of  $\gamma$  in the intervals reported in Table 2 and then the average of p-values is presented.  $\gamma$  values were chosen equal to 0.35, 1.3 and 2.2.

**Example.** In this example, first a random sample of size 25 from the exponential distribution with mean 1 is produced (sample 1). In the samples 2 and 3, respectively 76% and 44% of sample 1 are selected randomly and others are produced from Beta(2,1) as outliers.

**Table 7.** Critical values, Test statistics and p-values of the proposed tests

sample 1	$T_{\gamma,n}$			$T_{\gamma,mn}$		
	$\gamma = .35$	$\gamma = 1.3$	$\gamma = 2.2$	$\gamma = .35$	$\gamma = 1.3$	$\gamma = 2.2$
Test statistic	0.541	0.249	0.005	1.837	0.182	0.279
Critical value	0.642	0.329	0.014	1.958	0.268	0.919
Mean of the p-values	0.315			0.321		
1.2pt1pt1pt sample 2						
$T_{\gamma,n}$	$T_{\gamma,n}$			$T_{\gamma,mn}$		
	$\gamma = .35$	$\gamma = 1.3$	$\gamma = 2.2$	$\gamma = .35$	$\gamma = 1.3$	$\gamma = 2.2$
Test statistic	1.467	0.036	0.008	1.919	0.242	0.195
Critical value	1.558	0.055	0.014	1.956	0.267	0.256
Mean of the p-values	0.165			0.095		
1.2pt1pt1pt sample 3						
$T_{\gamma,n}$	$T_{\gamma,n}$			$T_{\gamma,mn}$		
	$\gamma = .35$	$\gamma = 1.3$	$\gamma = 2.2$	$\gamma = .35$	$\gamma = 1.3$	$\gamma = 2.2$
Test statistic	2.264	0.117	0.032	3.032	0.597	0.501
Critical value	1.544	0.055	0.014	1.950	0.269	0.258
Mean of the p-values	0.001			0.0003		

**Table 8.** Critical values, Test statistics and p-values of the competing tests

sample 1	$T_n$	$T_{park}$	$KLC_{mn}$	$W^2$	$S^2$
Test statistic	0.113	0.087	0.735	0.200	1.867
Critical value	0.135	0.122	0.916	0.308	2.398
p-value	0.093	0.153	0.257	0.239	0.236
sample 2	$T_n$	$T_{park}$	$KLC_{mn}$	$W^2$	$S^2$
Test statistic	0.118	0.088	0.667	0.323	2.291
Critical value	0.135	0.123	0.914	0.297	2.401
p-value	0.080	0.151	0.085	0.034	0.073
sample 3	$T_n$	$T_{park}$	$KLC_{mn}$	$W^2$	$S^2$
Test statistic	0.398	0.230	0.914	1.018	4.513
Critical value	0.135	0.124	0.443	0.297	2.428
p-value	0	0.008	0.0002	0	0

Based on Tables 7 and 8, for data sample 1, as expected all the proposed test statistics indicate that the considered data set at a 5% significance level follow an exponential distribution. For data sample 2, all the proposed (Table 7) and competing tests (Table 8), except  $W^2$  statistic, confirm the assumption exponentiality of data distribution. For sample 3, the null hypothesis of exponentiality at a 5% significance level is rejected for all the proposed and competing tests as expected (see Barnett and Lewis (1994), page 68).

## 6 Conclusion

In the present paper, test of exponentiality based on estimate of the gamma-divergence has been presented. This divergence measure in the presence of outliers is very robust and consequently more flexible performance. To construct the test statistics, EDF and Vasicek entropy estimator were used. Then, we studied their performances for an exponential distribution. The power of the proposed tests were compared with competing tests. For this end, we have chosen specific values of  $\gamma$  that maximized the power of the proposed tests. In the simulation results, the proposed tests perform well in comparison with the competing tests for alternatives with IFR and DER functions.

## References

- [1] Abbasnejad, M. (2011). Some goodness of fit test based on renyi information. *Applied Mathematical Sciences*, 5, 1921-1934.
- [2] Ahmad, I. A. and Alwasel, I. (1999). A goodness of fit test for exponentiality based on the memoryless property. *Journal of the Royal Statistical Society, Ser. B*, 61, 681-689.
- [3] Alizadeh Noughabi, H. (2010). A new estimator of entropy and its application in testing normality. *Journal of Statistical Computation and Simulation*, 80(10), 1151-1162.
- [4] Alwasel, I. (2001). On goodness of fit testing for exponentiality using the memoryless property. *Journal of Nonparametric Statistics*, 13, 569-581.
- [5] Andrews, F. C. and Andrews, A. C. (1962). The form of the equilibrium distribution function. *Transactions of the Kansas Academy of Science*, 65, 247-256.
- [6] Baratpour, S. and Habibirad, A. (2012). Testing goodness of fit for exponential distribution based on cumulative residual entropy. *Communications in Statistics-Theory and Methods*, 41, 1387-1396.
- [7] Barnett, V. and Lewis, T. (1994). Outliers in Statistical Data, John Wiley.
- [8] Chen, Z. (2000). A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. *Stprobl* , 49, 155-161.
- [9] Choi, B. Kim, K. and Song, S.H. (2004). Goodness of fit test for exponentiality based on Kullback-Leibler information. *Communications in Statistics-Simulation and Computation*, 33(2), 525-536.
- [10] Cichocki, A. and Amari, S. (2010). Families of alpha- beta- and gamma-divergences: flexible and robust measures of similarities. *Entropy*, 12 , no. 6, 1532-1568.
- [11] Correa J. C. (1995). A new estimator of entropy. *Communications in Statistics-Theory and Methods*, 24, 2439-2449.
- [12] Dhillon, B. (1981). Lifetime distributions. *IEEE Trans Reliab*, 30, 457-459.
- [13] Ebrahimi, N. Habibullah, M. and Soofi, E.S. (1992). Testing exponentiality based on Kullback-Leibler information. *Journal of the Royal Statistical Society*, 54, 739-748.
- [14] Ebrahimi, N. Pflughoeft, K. and Soofi, E. (1994). Two measures of sample entropy. *Statistics and Probability Letters*, 20, 225-234.
- [15] Eguchi, S. and Kato, S. (2010). Entropy and divergence associated with power function and the statistical application. *Entropy*, 12, 262-274.
- [16] Finkelstein, J. and Schafer, R. E. (1971). Imported goodness of fit tests. *Biometrika*, 58, 641-645.

- [17] Fujisawa, H. and Eguchi, S. (2008). Robust parameter estimation with a small bias against heavy contamination. *Journal of Multivariate Analysis*, 99, 2053-2081.
- [18] Gurevich, G. and Davidson, A. (2008). Standardized form of Kullback-Leibler Information based statistics for normality and exponentiality. *Computer Modeling and New Technologies*, 12, 14-25.
- [19] Harris, C. M. (1976). A note on testing for exponentiality. *Naval Research Logistics Quarterly*, 28, 169-175.
- [20] Jones, M. C. Hjort, N. L. Harris, I. R. and Basu, A. (2001). A comparison of related density based minimum divergence estimators. *Biometrika*, 88(3), 865-873.
- [21] Kullback, S. (1959). Information theory and statistics. New York: John Wiley.
- [22] Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, 22, 79-86.
- [23] Lilliefors, H.W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association*, 64, 387-389.
- [24] Nakamura, T. K. (2009). Relativistic equilibrium distribution by relative entropy maximization. *Europhysics Letters Association*, Volume 88, Number 4.
- [25] Ooms, G. and Moore, K. (1991). A novel assay for genetic and environmental changes in the architecture of intact root systems of plants grown in vitro. *Plant cell, tissue and organ culture*, 27(2), 129-139.
- [26] Park, S. Choi, D. and Jung, S. (2014). Kullback-Leibler information of the equilibrium distribution function and its application to goodness of fit test. *Communications for Statistical Applications and Methods*, 2, 125-134.
- [27] Rao, M. Chen, Y. Vemuri, B. C. and Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE Transactions on Information Theory*, 50, 1220-1228.
- [28] Shannon C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379-423.
- [29] Stephens, M.A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of American Statistical Association*, 69, 730-737.
- [30] Van Es, B. (1992). Estimating functional related to a density by a class of statistic based on spacings. *Scandinavian Journal of Statistics*, 19, 61-72.
- [31] Van Soest, J. (1969). Some goodness of fit tests for the exponential distribution. *Statistica Neerlandica*, 23, 41-51.
- [32] Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of the Royal Statistical Society, B*, 38, 54-59.



# Stochastic comparisons of mixture of coherent systems with dependent component lifetimes

Amini-Seresht, E.<sup>1</sup>

Department of Statistics, Bu-Ali Sina University, Hamedan

## Abstract

In this paper, we are investigated that various stochastic orders between the random environments can be translated into stochastic orders between the lifetime of the systems with dependent component lifetimes sharing the different random environment variables.

**Keywords:** Stochastic orders, Coherent system, Random environment, Distorted function, Copula function.

## 1 Introduction

A system is said to be coherent system if it has no irrelevant component and the structure function of the system is monotone in each argument (that is, an improvement of a component cannot lead to a deterioration in system performance). In the context of reliability is usually assumed that the component lifetimes are homogeneous independent, but in many cases, there are some situations in which the component lifetimes of the coherent system are dependent, such as the components that work in a collaborative environment, therefore, the components of a system maybe dependent through an environment random variable.

Suppose that  $\mathbf{X} = (X_1, \dots, X_n)$  be a random vector of homogeneous dependent component lifetimes with the distribution function  $F(x_1, \dots, x_n)$ . Let  $\Theta$  be an environment

<sup>1</sup>e.amini64@yahoo.com

random variable with distribution function  $W$  and density function  $w$  that takes on values in  $\chi \subseteq \mathcal{R}^+$ . Given  $\Theta = \theta$ , then we use  $\mathbf{X}(\theta) = (X_1(\theta), \dots, X_n(\theta))$  to denote the lifetime vector of homogeneous dependent component lifetimes of a coherent system, that its distribution function depends on the value of  $\theta$ , where  $X_i(\theta), i = 1, \dots, n$  having the marginal distribution function  $F(\cdot|\theta)$ . Then, the dependence will be determined by the joint reliability function of  $\mathbf{X}(\theta)$  that is given by

$$\begin{aligned}\bar{F}(x_1, \dots, x_n|\theta) &= P(X_1(\theta) > x_1, \dots, X_n(\theta) > x_n) \\ &= \mathbb{K}(\bar{F}(x_1|\theta), \dots, \bar{F}(x_n|\theta)),\end{aligned}$$

where  $\mathbb{K}$  is a multivariate distribution function with support  $\mathcal{R}^{+n}$  and has uniform marginals on  $[0, 1]$ . In the literature,  $\mathbb{K}$  is called a reliability copula. Through this paper, we consider  $\mathbf{X}(\Theta_i) = (X_1(\Theta_i), \dots, X_n(\Theta_i))$ ,  $i = 1, 2$  as a random mapping whose distribution function is a mixture of distribution functions in  $\{F(\cdot|\theta), \theta \in \chi\}$ . Suppose that the random variable  $\tau(\mathbf{X}(\Theta_i)) = \phi(X_1(\Theta_i), \dots, X_n(\Theta_i))$  (where  $\phi$  is the coherent system life function) denotes the lifetime of a coherent system. Then, the reliability function of coherent system  $\phi$  is given by

$$\bar{F}_{\tau(\mathbf{X}(\Theta_i))}(x) = P(\tau(\mathbf{X}(\Theta_i)) > x) = \int h(\bar{F}(x|\theta))dW_i(\theta), \quad (1)$$

where  $h : [0, 1] \rightarrow [0, 1]$  is a distorted function that depends only on the structure function  $\phi$  and the reliability copula  $\mathbb{K}$ .

Various stochastic comparison results for the general mixture models have been investigated in the literature with respect to some well-known univariate/multivariate stochastic orders, see for example Khaledi and Shaked (2010), Belzunce et al. (2009), Amini-Seresht and Khaledi (2015), Misra and Misra (2012), Li and Da (2010), Jarrahiferiz et al. (2016) and Misra et al. (2009). Also, Kayid et.al. (2015) consider the usual mixture of frailty models and obtained some relative behaviour based on a stochastic order. Fernndez-Ponce et al. (2015) obtained some stochastic comparisons between generalized frailty models. The purpose of this paper is to investigate that under some conditions on the distorted function of the coherent system, when the environmental random variables are ordered with respect to the various stochastic orders then their lifetime random variable of coherent systems are also ordered.

## 2 Main results

In this section, we stochastically compare the lifetimes of  $\tau(\mathbf{X}(\Theta_1))$  and  $\tau(\mathbf{X}(\Theta_2))$  according to various stochastic orders.

**Theorem 1.** *Let  $\tau(\mathbf{X}(\Theta_1))$  and  $\tau(\mathbf{X}(\Theta_2))$  be the lifetimes of a coherent system with common dependent component lifetimes and with system's reliability function given in (2). Given  $\Theta_i = \theta_i$ ,  $i = 1, 2$  and assume that  $\theta_1 \leq \theta_2$ . If*

- $X_i(\theta_1) \leq_{st} X_i(\theta_2)$  for  $i = 1, \dots, n$ , and
- $\Theta_1 \leq_{st} \Theta_2$ .

Then, it holds that  $\tau(\mathbf{X}(\Theta_1))) \leq_{\text{st}} \tau(\mathbf{X}(\Theta_2))$ .

In the next result we consider the hazard rate order to compare the lifetimes of coherent systems with the homogeneous dependent component lifetimes.

**Theorem 2.** Let  $\tau(\mathbf{X}(\Theta_1)))$  and  $\tau(\mathbf{X}(\Theta_2))$  be the lifetimes of a coherent system with common dependent component lifetimes and with system's reliability function given in (2). Given  $\Theta_i = \theta_i$ ,  $i = 1, 2$  and assume that  $\theta_1 \leq \theta_2$ . If

- $\frac{uh'(u)}{h(u)}$  is decreasing in  $u \in (0, 1)$
- $X_i(\theta_1) \leq_{\text{hr}} X_i(\theta_2)$  for  $i = 1, \dots, n$ , and
- $\Theta_1 \leq_{\text{hr}} \Theta_2$ .

Then, it holds that  $\tau(\mathbf{X}(\Theta_1))) \leq_{\text{hr}} \tau(\mathbf{X}(\Theta_2))$ .

The next result involves the reversed hazard rate stochastic order.

**Theorem 3.** Let  $\tau(\mathbf{X}(\Theta_1)))$  and  $\tau(\mathbf{X}(\Theta_2))$  be the lifetimes of a coherent system with common dependent component lifetimes and with system's reliability function given in (2). Given  $\Theta_i = \theta_i$ ,  $i = 1, 2$  and assume that  $\theta_1 \leq \theta_2$ . If

- $\frac{(1-u)h'(u)}{1-h(u)}$  is increasing in  $u \in (0, 1)$
- $X_i(\theta_1) \leq_{\text{rh}} X_i(\theta_2)$  for  $i = 1, \dots, n$ , and
- $\Theta_1 \leq_{\text{rh}} \Theta_2$ .

Then, it holds that  $\tau(\mathbf{X}(\Theta_1))) \leq_{\text{rh}} \tau(\mathbf{X}(\Theta_2))$ .

In the next result, we obtained under some conditions on the distorted function that the likelihood ratio order between the random environments can be translated into likelihood ratio order between the lifetime of the mixture of coherent systems.

**Theorem 4.** Let  $\tau(\mathbf{X}(\Theta_1)))$  and  $\tau(\mathbf{X}(\Theta_2))$  be the lifetimes of a coherent system with common dependent component lifetimes and with system's reliability function given in (2). Given  $\Theta_i = \theta_i$ ,  $i = 1, 2$  and assume that  $\theta_1 \leq \theta_2$  and for some point  $v \in (0, 1)$ . If

- $\frac{uh''(u)}{h'(u)}$  is decreasing in  $u$  and non-negative for all  $u \in (0, v)$ ,
- $\frac{(1-u)h''(u)}{h'(u)}$  is decreasing in  $u$  and non-positive for all  $u \in (v, 1)$ ,
- $X_i(\theta_1) \leq_{\text{lr}} X_i(\theta_2)$  for  $i = 1, \dots, n$ , and
- $\Theta_1 \leq_{\text{lr}} \Theta_2$ .

Then, it holds that  $\tau(\mathbf{X}(\Theta_1))) \leq_{\text{lr}} \tau(\mathbf{X}(\Theta_2))$ .

## References

- [1] Amini-Seresht, E., and Khaledi, B-E. (2015). Multivariate stochastic comparisons of mixture models. *Metrika*. **78**, 1015-1034.
- [2] Belzunce, F., Mercader, J.A., Ruiz, J.M, and Spizzichino, F. (2009). Stochastic comparisons of multivariate mixture models *Journal of Multivariate Analysis*. **100**, 1657-1669.
- [3] Fernndez-Ponce, J.M., Pellerey, F., and Rodrguez-Griolo, M.R. (2015). Some stochastic properties of conditionally dependent frailty models. *Statistics*. **50**, 649-666.
- [4] Kayid1, M., Izadkhah, S., and Zuo, M. (2015). Some results on the relative ordering of two frailty models. *Statistical Papers*, DOI: 10.1007/s00362-015-0697-8.
- [5] Khaledi, B., Shaked, M. (2007). Ordering conditional lifetimes of coherent systems. *Journal of Statistical Planning and Inference*. **137**, 1173-1184.
- [6] Khaledi, B-E., and Shaked, M. (2010). Stochastic comparisons of multivariate mixtures. *Journal of Multivariate Analysis*. **101**, 2486-2498.
- [7] Misra, A.K., and Misra, N. (2012). Stochastic properties of conditionally independent mixture models, *Journal of Statistical Planning and Inference*. **142**, 1599-1607.



# Regression analysis of clustered interval-censored failure time data with the proportional hazards model

Amoorahimi, E.<sup>1</sup>      Kelkinnama, M.<sup>2</sup>

<sup>1,2</sup> Department of Mathematical Sciences, Isfahan University of Technology

## Abstract

The Cox proportional hazard (PH) model is the most popular regression model for analyzing the survival data. In this paper, we consider the clustered interval-censored failure time data and discuss the regression analysis under the PH model. This type of data set can occur when the desired failure times are gathered from several clusters and known only to lie in certain intervals. We propose the use of expectation-maximization (EM) algorithm for obtaining the maximum likelihood estimates of model parameters. The developed EM algorithm consists of two-stage data augmentation involving latent Poisson random variables. The performance of the proposed method is evaluated through a simulation study.

**Keywords:** Cluster interval-censored data, Within-cluster resampling, EM algorithm, Monotone splines, Proportional hazard model.

## 1 Introduction

In many survival analysis problems, we are interested in the relationship between a failure time and a vector of covariates. However, it is common that the failure times of interest may not be observed exactly but are known only to belong to certain intervals. Such data are usually referred to as interval-censored failure time data. Interval-censored data

<sup>1</sup>californium3101@gmail.com

<sup>2</sup>m.kelkinnama@cc.iut.ac.ir

often occur in many studies of epidemiology, longitudinal, or biomedical research in which subjects are followed periodically for the event of interest. In these studies, the event time  $T$  is not directly observable but may be detected in some periodic examination interval, denoted as  $(L, R]$ , where  $L$  is the left examination time and  $R$  is the right examination time [4].

Clustered interval-censored failure time data may naturally arise in periodic follow-up studies where each study subject is repeatedly measured at discrete time points and some subjects come from the same cluster or group such as sibling, family and community. These subjects usually share certain unobserved characteristics and their failure times tend to be correlated as a result. Siblings, for example, share similar genetic and environmental influences.

Regression analysis of failure time data is one of the main objectives in failure time data analysis, and in the case of right-censored data, many procedures have been developed under various regression models. Such as the proportional hazards (PH) mode for regression analysis of clustered failure time data, several methods have been proposed when only right censoring is present. For example, fitting of semi-parametric linear transformation models; estimating equations and pseudolikelihood; second order generalized estimating equations (GEE) and a copula model.

Fitting the regression models to clustered interval-censored data has been investigated by some authurs. They used both parametric and semiparametric approaches. For example [1] and [3] considered fitting the semiparametric additive hazards model to clustered interval-censored. For most of the existing semiparametric methods, it is assumed that the failure time of interest follows the proportional hazards model. This model is one of the most commonly used models for regression analysis of failure time data.

For inference, in theory, the maximum likelihood approach applies to any model and is the most efficient method. However, it may be complicated both in terms of investigation of its properties and its implementation. Given the omnipresent nature of clustered interval-censored data, there exists a necessary need to develop accurate, flexible, easy-to-implement and computationally efficient statistical methods for regression analysis of this form of data. In this paper we develop an EM algorithm to analyse clustered interval-censored failure time data. Our method reduces the analysis of clustered interval-censored data to that of interval-censored failure time data. The algorithm is implemented in two steps. First, we apply a within-cluster resampling (WCR) procedure to sample a single subject from each cluster and obtain independent interval-censored failure time data. In other words, one observation is randomly sampled with replacement from each of the  $n$  clusters and hence independent interval-censored dataset of size  $n$  is formed. Then an EM algorithm can be used to find the maximum likelihood estimates of all unknown parameters, and closed-form expressions of the asymptotic variance estimates. For this end, we adopted the idea of [5]. The basic idea of their approach is the use of monotone splines for approximating the cumulative baseline hazard function in the PH and a two-stage data augmentation process that leads to the development of an EM algorithm. If the resampling process is repeated a large number of times, say  $Q$ , where each of the  $Q$  analysis provides a consistent estimator of parameter of interest, one can obtain the WCR estimate by taking the average of the  $Q$  resampling-based estimates.

In the following,in Section 2 we will briefly introduce the proportional hazards model

and some notations. The proposed estimation procedure is then presented in Section 3, details include the use of monotone splines and two-stage data augmentation process that leads to the development of an EM algorithm. Results from an a simulation study are reported in Section 4 for assessing the performance of our proposed approach.

## 2 Data and model

Now we consider a survival study that includes  $n$  independent clusters with  $n_i$  denoting the size of the  $i$ th cluster, where  $i = 1, \dots, n$ . Let  $T_{ij}$ , denote the failure time of interest for subject  $j$  in cluster  $i$ ,  $j = 1, \dots, n_i$ . Suppose that one observes clustered interval-censored failure time data on the  $T_{ij}$ 's denoted by  $\{(L_{ij}, R_{ij}], x_{ij}, i = 1, \dots, n; j = 1, \dots, n_i\}$ , where  $L_{ij} \leq R_{ij}$ . Here  $L_{ij} = 0$  represents a left-censored observation,  $R_{ij} = \infty$  corresponds to a right-censored one, and  $x_{ij}$  is a possibly time-independent covariate vector associated with subject  $j$  in cluster  $i$ . Note that for interval-censored data, all available information about  $T_{ij}$  is the interval  $(L_{ij}, R_{ij}]$  with  $T_{ij} \in (L_{ij}, R_{ij}]$ . We assume that  $L_{ij}$  and  $R_{ij}$  are independent of  $T_{ij}$  given the covariates  $x_{ij}$ . Throughout this paper the  $T_{ij}$ 's are considered to be independent for subjects in different clusters but could be dependent for subjects within the same cluster. To describe the association within each cluster, we will assume that there exists a cluster-specific latent variable  $b_i$  with mean zero and unknown variance. Furthermore suppose that given covariates and  $b_i$ , the failure times within each cluster are independent and  $T_{ij}$  follows the proportional hazards model

$$\lambda_{ij}(t|X_{ij}, b_i) = \lambda_0(t) \exp\{\boldsymbol{\beta}' X_{ij}\} \xi_i \quad (1)$$

where  $\lambda_0(t)$  denotes the unknown baseline function and  $\boldsymbol{\beta}$  is the vector of unknown regression coefficients [2] and  $\xi_i$  is random variable follow a positive-value distribution. In this paper we use Lognormal distribution for  $\xi_i$ . This model is commonly called frailty model in the context. Our main goal is to estimate  $\boldsymbol{\beta}$ .

## 3 The proposed method

### 3.1 Within-cluster resampling (WCR) method

In the within-cluster resampling method [1], we randomly select one data point  $\{(L_i, R_i], \mathbf{x}_i\}$  with replacement from each cluster and generate a set of interval-censored failure time data. Let  $Q$  be a positive integer. We repeat the resampling process  $Q$  times and generate  $Q$  sets of independent samples of interval-censored failure time data. For each set of the interval-censored failure time data, we propose the EM algorithm in Section(3.4) to estimate  $\boldsymbol{\beta}$ .

For this data let  $F(\cdot|\mathbf{x})$  denote the cumulative distribution function (CDF) of the failure time of interest given the covariate vector  $\mathbf{x}$ . Under the PH model (1) the failure time distribution for individuals with covariates  $\mathbf{x}_i$  is given by

$$F(t|\mathbf{x}_i) = 1 - \exp\{-\Lambda_0(t) \exp(\boldsymbol{\beta}' \mathbf{x}_i)\xi_i\}, \quad (2)$$

where  $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})'$  is a  $p \times 1$  vector of time-independent covariates,  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$  is the corresponding vector of regression coefficients, and  $\Lambda_0(\cdot)$  is the cumulative baseline hazard function.

The likelihood function based on the observed data  $\mathcal{D} = \{(L_i, R_i], \mathbf{x}_i, i = 1, \dots, n\}$  is

$$\mathcal{L}_{obs} = \prod_{i=1}^n \{F(R_i|\mathbf{x}_i) - F(L_i|\mathbf{x}_i)\}.$$

Let  $\delta_{i1}$ ,  $\delta_{i2}$  and  $\delta_{i3}$  be censoring indicators for the  $i$ th subject denoting left-, interval-, and right-censoring, respectively. Then the observed data likelihood can be rewritten as the following form

$$\mathcal{L}_{obs} = \prod_{i=1}^n F(R_i|\mathbf{x}_i)^{\delta_{i1}} \{F(R_i|\mathbf{x}_i) - F(L_i|\mathbf{x}_i)\}^{\delta_{i2}} F(L_i|\mathbf{x}_i)^{\delta_{i3}}, \quad (3)$$

where  $\delta_{i1} + \delta_{i2} + \delta_{i3} = 1$ .

### 3.2 Modeling $\Lambda_0(\cdot)$ with monotone splines

The observed data likelihood in Equation (3) is a function of the regression parameter  $\beta$  and the cumulative baseline hazard function. Some methods, such as partial likelihood consistently estimate  $\boldsymbol{\beta}$ , without having to estimate  $\Lambda_0(\cdot)$ , under the PH model. Estimating  $\Lambda_0(\cdot)$  is challenging from both a theoretical and computational perspective because of its infinite dimension.

In the survival literature, approximating unknown functions with splines is commonly done. The use of splines bridges the gap between nonparametric and parametric modeling approaches because, while providing adequate flexibility, only a finite number of parameters need to be estimated. In this paper,  $\Lambda_0(\cdot)$  is modeled using I-splines. This approach leads to the following representation

$$\Lambda_0(\cdot) = \sum_{l=1}^k \gamma_l b_l(\cdot) \quad (4)$$

where the  $b_l(\cdot)$ 's are integrated spline basis functions, each of which is nondecreasing from 0 to 1 and  $\gamma_l$ 's are nonnegative coefficients which ensure that  $\Lambda_0(\cdot)$  is nondecreasing. This approximation is valid within a finite interval formed by the minimum and maximum of the censoring times in the observed data. To construct these functions, one needs only to specify an increasing sequence of points as knots and to choose a degree for the splines. The degree controls the smoothness of the splines and takes on values 1, 2, and 3 for linear, quadratic, and cubic functions, respectively. The placement of knots plays an important role in specifying monotone splines. In [5] it is recommended to use only a few knots, e.g., at the median or at the three quartiles. Once the degree and placement of knots are specified, the  $k$  spline basis functions are fully determined, where  $k$  is equal to the number of interior knots plus the degree and using 10 – 30 equally-spaced knots.

### 3.3 Data augmentation

In this paper, we develop an EM algorithm for finding the maximum likelihood estimate of  $\boldsymbol{\theta}$ , where  $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}')$  and  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_k)'$ . For derivation of the algorithm we use two layers of Poisson latent variables. Consider a non-homogeneous Poisson process  $N(t)$  with a cumulative intensity function  $\Lambda_0(t) \exp\{\boldsymbol{\beta}' \mathbf{x}\} \xi$ . Let  $T$  denote the first time that the counting process is start; i.e.,  $T = \inf\{t : N(t) > 0\}$ . Note that since  $N(t)$  is a Poisson random variable with mean parameter  $\Lambda_0(t) \exp\{\boldsymbol{\beta}' \mathbf{x}\} \xi$ , then for the random variable  $T$  we have

$$Pr(T > t) = Pr\{N(t) = 0\} = \exp\{-\Lambda_0(t) \exp(\boldsymbol{\beta}' \mathbf{x}) \xi\} = 1 - F(t|\mathbf{x})$$

This means that,  $T$  follows the proportional hazard model (1).

Using this fact, we construct an augmented data likelihood, using a latent nonhomogeneous Poisson process. Suppose that  $N_i(t)$  denote the latent Poisson process for subject  $i$ , which has cumulative intensity function  $\Lambda_0(t) \exp(\boldsymbol{\beta}' \mathbf{x}_i) b$  for  $i = 1, \dots, n$ . Let  $t_{i1} = R_i I_{(\delta_{i1}=1)} + L_i I_{(\delta_{i1}=0)}$  and define  $Z_i = N_i(t_{i1})$ . Similarly, when  $\delta_{i1} = 0$  define  $W_i = N_i(t_{i2}) - N_i(t_{i1})$ , where  $t_{i2} = R_i I_{(\delta_{i2}=1)} + L_i I_{(\delta_{i3}=1)}$ . Under the assumptions, the random variables  $Z_i$  and  $W_i$  have Poisson distributions with mean parameters  $\Lambda_0(t_{i1}) \exp(\boldsymbol{\beta}' \mathbf{x}_i) \xi_i$  and  $[\Lambda_0(t_{i2}) - \Lambda_0(t_{i1})] \exp(\boldsymbol{\beta}' \mathbf{x}_i) \xi_i$ , respectively. In addition,  $Z_i$  and  $W_i$  are independent when  $\delta_{i1} = 0$ .

The augmented likelihood function based on the latent  $Z_i$ 's,  $W_i$ 's and the observed data is given by

$$\begin{aligned} \mathcal{L}_{aug}(\theta) = & \prod_{i=1}^n P_{Z_i}(Z_i) [P_{W_i}(W_i)]^{\delta_{i2}+\delta_{i3}} \{ \delta_{i1} I_{(Z_i>0)} \\ & + \delta_{i2} I_{(Z_i=0, W_i>0)} + \delta_{i3} I_{(Z_i=0, W_i=0)} \} \end{aligned} \quad (5)$$

where  $P_A(\cdot)$  denotes the probability mass function of the random variable  $A$ .

In the second layer, we again exploit the Poisson distribution and also the monotone spline representation of  $\Lambda_0(t)$ . Note that  $Z_i$  and  $W_i$  can be decomposed as sum of  $k$  independent Poisson random variables,  $Z_i = \sum_{l=1}^k Z_{il}$  and  $W_i = \sum_{l=1}^k W_{il}$ , where  $Z_{il}$  and  $W_{il}$ , for  $l = 1, \dots, k$ , are Poisson random variables with mean parameters  $\gamma_l b_l(t_{i1}) \exp(\boldsymbol{\beta}' \mathbf{x}_i) \xi_i$  and  $\gamma_l [b_l(t_{i2}) - b_l(t_{i1})] \exp(\boldsymbol{\beta}' \mathbf{x}_i) \xi_i$ , respectively. At this layer, what we call the complete data likelihood is given by

$$\mathcal{L}_c(\theta) = \prod_{i=1}^n \prod_{l=1}^k P_{Z_{il}}(Z_{il}) [P_{W_{il}}(W_{il})]^{\delta_{i2}+\delta_{i3}}. \quad (6)$$

To see the connection between the complete data likelihood and the observed data likelihood, one can integrate the  $Z_{il}$ 's and  $W_{il}$ 's out of (6) and get the (3).

### 3.4 The EM algorithm

We now describe the EM algorithm for the PH model. The E-step of the algorithm involves taking the expectation of  $\log \mathcal{L}_c(\theta)$  with respect to all of the latent variables

$(Z'_i s, Z'_{il} s, W'_i s$  and  $W'_{il} s)$ , conditional on the observed data  $\mathcal{D}$  and the current parameter  $\boldsymbol{\theta}^{(d)} = (\boldsymbol{\beta}^{(d)\prime}, \boldsymbol{\gamma}^{(d)\prime})'$ . This yields the function  $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)}) = \mathbb{E}[\log \mathcal{L}_c(\theta) | \mathcal{D}, \theta^{(d)}]$  which can be expressed as

$$\begin{aligned} \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)}) &= \sum_{i=1}^n \sum_{l=1}^k \left[ \{\mathbb{E}(Z_{il} | \mathcal{D}, \boldsymbol{\theta}^{(d)}) + (\delta_{i2} + \delta_{i3})\mathbb{E}(W_{il} | \mathcal{D}, \boldsymbol{\theta}^{(d)})\} \{\log(\gamma_l) + \boldsymbol{\beta}' \mathbf{x}_i \right. \\ &\quad \left. + \log(\xi_i)\} - \gamma_l \exp(\boldsymbol{\beta}' \mathbf{x}_i) \xi_i \{(\delta_{i1} + \delta_{i2})b_l(R_i) + \delta_{i3}b_l(L_i)\} + L(\boldsymbol{\theta}^{(d)}) \right]. \end{aligned} \quad (7)$$

where  $L(\boldsymbol{\theta}^{(d)})$  is a function of  $\boldsymbol{\theta}^{(d)}$  but is free of  $\boldsymbol{\theta}$ .

Noting that the conditional distribution of  $Z_{il}(W_{il})$  given  $Z_i(W_i)$  is binomial and by applying the law of iterated expectations, one can obtain the following conditional expectations

$$\mathbb{E}(Z_{il} | \mathcal{D}, \theta^{(d)}) = \frac{\gamma_l^{(d)} b_l(R_i)}{\Lambda_0^{(d)}(R_i)} \mathbb{E}(Z_i | \mathcal{D}, \theta^{(d)})$$

and

$$\mathbb{E}(W_{il} | \mathcal{D}, \theta^{(d)}) = \frac{\gamma_l^{(d)}(b_l(R_i) - b_l(L_i))}{\Lambda_0^{(d)}(R_i) - \Lambda_0^{(d)}(L_i)} \mathbb{E}(W_i | \mathcal{D}, \theta^{(d)})$$

where,  $\Lambda_0^{(d)}(.) = \sum_{l=1}^k \gamma_l^{(d)} b_l(.)$ . It can be concluded from augmented likelihood (5) that  $Z_i(W_i)$  conditionally has a truncated Poisson distribution given the observed data. Therefore, the respective expected values of  $Z_i$  and  $W_i$ , given  $\mathcal{D}$  and  $\boldsymbol{\theta}^{(d)}$ , can be expressed as

$$\mathbb{E}(Z_i | \mathcal{D}, \theta^{(d)}) = \frac{[\Lambda_0^{(d)}(R_i) e^{(\boldsymbol{\beta}^{(d)\prime} \mathbf{x}_i) \xi_i}] \delta_{i1}}{1 - \exp \left\{ -[\Lambda_0^{(d)}(R_i) e^{(\boldsymbol{\beta}^{(d)\prime} \mathbf{x}_i) \xi_i}] \right\}}$$

and

$$\mathbb{E}(W_i | \mathcal{D}, \boldsymbol{\theta}^{(d)}) = \frac{\left\{ [\Lambda_0^{(d)}(R_i) - \Lambda_0^{(d)}(L_i)] e^{(\boldsymbol{\beta}^{(d)\prime} \mathbf{x}_i) \xi_i} \right\} \delta_{i2}}{1 - \exp \left\{ -[\Lambda_0^{(d)}(R_i) - \Lambda_0^{(d)}(L_i)] e^{(\boldsymbol{\beta}^{(d)\prime} \mathbf{x}_i) \xi_i} \right\}}.$$

Note  $\delta_{i3} \mathbb{E}(W_{il} | \mathcal{D}, \boldsymbol{\theta}^{(d)}) = 0$  for all  $i$  and  $l$ , and these terms are therefore ignored henceforth.

The M-step at EM algorithm is devoted to obtaining  $\boldsymbol{\theta}^{(d+1)} = \arg \max_{\boldsymbol{\theta}} \mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ . For this, we need the partial derivatives of  $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$  with respect to  $\boldsymbol{\theta}$ , which are

$$\begin{aligned} \frac{\partial \mathcal{Q}}{\partial \gamma_l} &= \sum_{i=1}^n \left\{ \gamma_l^{-1} [\mathbb{E}(Z_{il} | \mathcal{D}, \boldsymbol{\theta}^{(d)}) + \delta_{i2} \mathbb{E}(W_{il} | \mathcal{D}, \boldsymbol{\theta}^{(d)})] \right. \\ &\quad \left. - [(\delta_{i1} + \delta_{i2})b_l(R_i) + \delta_{i3}b_l(L_i)] e^{(\boldsymbol{\beta}' \mathbf{x}_i) \xi_i} \right\} \end{aligned}$$

$$\begin{aligned} \frac{\partial \mathcal{Q}}{\partial \boldsymbol{\beta}} &= \sum_{i=1}^n \left\{ [\mathbb{E}(Z_i | \mathcal{D}, \boldsymbol{\theta}^{(d)}) + \delta_{i2} \mathbb{E}(W_i | \mathcal{D}, \boldsymbol{\theta}^{(d)})] \right. \\ &\quad \left. - [(\delta_{i1} + \delta_{i2})\Lambda_0(R_i) + \delta_{i3}\Lambda_0(L_i)] e^{(\boldsymbol{\beta}' \mathbf{x}_i) \xi_i} \right\} \mathbf{x}_i \end{aligned}$$

Setting these derivatives equal to zero and solving the resulting system of equations for  $\boldsymbol{\theta}$  yields  $\boldsymbol{\theta}^{(d+1)}$ . Towards this end, note that solving  $\frac{\partial \mathcal{Q}}{\partial \gamma_l} = 0$  leads to an explicit form of  $\gamma_l^{(d+1)}$  which itself is a function of  $\boldsymbol{\beta}^{(d+1)}$ . For implementation purposes, we now summarize our proposed EM algorithm succinctly for the PH model. First, initialize  $\boldsymbol{\theta}^{(d)} = (\boldsymbol{\beta}^{(d)\prime}, \gamma^{(d)\prime})'$  for  $d = 0$  and then simply repeat the following two steps until convergence

1. Obtain  $\boldsymbol{\beta}^{(d+1)}$  by solving the following system of  $p$  equations

$$\begin{aligned} & \sum_{i=1}^n [\mathbb{E}(Z_i|\mathcal{D}, \boldsymbol{\theta}^{(d)}) + \delta_{i2}\mathbb{E}(W_i|\mathcal{D}, \boldsymbol{\theta}^{(d)})] \\ &= \sum_{i=1}^n \sum_{l=1}^k [(\delta_{i1} + \delta_{i2})b_l(R_i) + \delta_{i3}b_l(L_i)]\gamma_l^{\star(d)}(\boldsymbol{\beta})e^{(\boldsymbol{\beta}'\mathbf{x}_i)}\xi_i \end{aligned}$$

where

$$\gamma_l^{\star(d)}(\boldsymbol{\beta}) = \frac{\mathbb{E}(Z_{il}|\mathcal{D}, \boldsymbol{\theta}^{(d)}) + \delta_{i2}\mathbb{E}(W_{il}|\mathcal{D}, \boldsymbol{\theta}^{(d)})}{[(\delta_{i1} + \delta_{i2})b_l(R_i) + \delta_{i3}b_l(L_i)]e^{(\boldsymbol{\beta}'\mathbf{x}_i)}\xi_i}$$

2. Let  $\gamma_l^{(d+1)} = \gamma_l^{\star(d)}(\boldsymbol{\beta}^{d+1})$  and update  $d = d + 1$ .

Solving the system of equations in the first step of the algorithm can be done using standard root finding functions, available in some statistical softwares. The function used for model fitting required inputs include the censoring indicators, the observed intervals, the covariate matrix, the number of knots, the degree of the monotone splines, a tolerance value that specifies the convergence criterion, and the initial values for the unknown parameters. The output includes the point estimates of the regression and spline coefficients.

It can be proven that  $\boldsymbol{\theta}^{(d+1)}$  is the unique global maximizer of  $\mathcal{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(d)})$ . Under the standard regularity conditions, for the maximum likelihood estimators we have, as  $n \rightarrow \infty$

$$n^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \text{Normal}(0, \mathcal{I}^{-1}(\boldsymbol{\theta})), \quad (8)$$

where,  $\mathcal{I}(\boldsymbol{\theta})$  is the Fisher information matrix.

Let  $\hat{\boldsymbol{\beta}}_q^{(d+1)}$  and  $\hat{\Sigma}_q^{(d+1)}$  denote the final estimate of  $\boldsymbol{\beta}$  and covariance matrix  $\boldsymbol{\Sigma}$ , for the  $q$ th resampled interval-censored failure time data with  $q = 1, \dots, Q$ . We estimate  $\boldsymbol{\beta}$  by

$$\hat{\boldsymbol{\beta}} = \frac{1}{Q} \sum_{q=1}^Q \hat{\boldsymbol{\beta}}_q^{(d+1)}$$

and  $\boldsymbol{\Sigma}$  by

$$\hat{\Sigma} = \frac{1}{Q} \sum_{q=1}^Q \hat{\Sigma}_q^{(d+1)} - \frac{1}{Q} \sum_{q=1}^Q (\hat{\boldsymbol{\beta}}_q^{(d+1)} - \hat{\boldsymbol{\beta}})(\hat{\boldsymbol{\beta}}_q^{(d+1)} - \hat{\boldsymbol{\beta}})'$$

## 4 A simulation study

An extensive simulation study is conducted to assess the performance of the EM algorithm presented in the previous sections. To generate the failure time of interest, we take

$\lambda_0(t) = 2$  in model (1) therefore the hazard function for the underlying failure time  $T$  was taken to be  $2 \exp\{\boldsymbol{\beta}' \mathbf{x}_{ij}\} \xi_i$  and generate the  $x_{ij}$ s from the Bernoulli distribution with success probability  $p = 0.5$ . The latent variable  $\xi_i$  is assumed to follow a lognormal distribution  $LN(0, 1/4)$ . For the generation of covariate-independent censoring intervals, we first generate multiple examination times for each subject to mimic many medical follow-up studies. For example, this would be the case if each subject is repeatedly monitored at discrete time points and some subjects are siblings or from the same family. More specifically, it is supposed that there are potentially ten examination times for each subject and the length or period between two consecutive examination times, denoted as  $len$ , was assumed to be a uniform random variable over  $(0, 0.2)$ . Suppose that  $\tau_{1ij}$  is the first random examination time generated from the uniform distribution  $(0, 0.2)$ . Then the following seven examination times were calculated as  $\tau_{kij} = \tau_{1ij} + (k - 1)len$ ,  $k = 2, 3, \dots, 10$ . It is assumed that a subject may miss the scheduled examination times with probability 0.5. Then the observed interval  $(L_{ij}, R_{ij}]$  was defined with  $L_{ij} = \tau_{sij}$  and  $R_{ij} = \tau_{qij}$ , where  $\tau_{sij}$  and  $\tau_{qij}$  with  $0 \leq s < q \leq 10$  are the two real adjacent observation times such that  $(\tau_{sij}, \tau_{qij})$  contains the true failure time  $T_{ij}$ .

Table (1) presents the simulation results for estimation of  $\boldsymbol{\beta}$  with the true value, denoted by  $\boldsymbol{\beta}_0$ , of  $\boldsymbol{\beta}$  taken to be 0.25, 0 and -0.25 with covariate-independent censoring intervals. Finally, the cluster size  $n_i$  was assumed to follow the uniform distribution 2, 3, 4 and  $n = 100$  and 200, and the results include the bias of  $\boldsymbol{\beta}$  (BIAS), the means of the estimated standard deviations (ESD), the sample standard errors of the estimated  $\boldsymbol{\beta}$  (SSE), and the 95% empirical converge probabilities (CP) of confidence interval obtain based on (8).

Table 1: Estimation of  $\boldsymbol{\beta}$

$\boldsymbol{\beta}_0$	$n$	The proposed estimator			
		BIAS	ESD	SSE	CP
-0.25	100	0.0135	0.2435	0.2515	0.937
0	100	0.0098	0.2582	0.2598	0.939
0.25	100	0.0140	0.2710	0.2915	0.940
-0.25	200	0.0038	0.1748	0.1795	0.948
0	200	0.0060	0.1840	0.1770	0.955
0.25	200	0.0031	0.1935	0.2025	0.944

It can be seen from the table that the EM algorithm performs well as the estimator has negligible bias and the estimated standard deviation is close to the empirical standard error with proper coverage probability. For the simulation results given above, we used  $Q = 1000$  for resampling. To investigate the effect of  $Q$  on the performance of the procedure, we used  $Q = 2000$  and obtained similar results.

## References

- [1] Chen, L., Sun, J. and Xiong, Ch. (2016), A multiple imputation approach to the analysis of clustered interval-censored failure time data with the additive hazards model, *Computational Statistics and Data Analysis*, **103**, 242-249.

- [2] Lam, K.F., Xu, Y. and Cheung, T.L. (2010), A multiple imputation approach for clustered interval-censored survival data, *Stat. Med.*, **29**, 680-693.
- [3] Li, J., Wang, Ch. and Sun, J (2012), Regression analysis of clustered interval-censored failure time data with the additive hazards model, *Journal of Nonparametric Statistics*, **24**, 1041-1050.
- [4] Sun, J. (2006), *The Statistical Analysis of Interval Censored Failure Time Data*, NewYork: Springer.
- [5] Wang, L., McMahan, Ch. S., Hudgens, M. G., and Qureshi, Z. P., (2016), A flexible, computationally efficient method for fitting the proportional hazards model to interval-censored data, *Biometrics*, **72**, 222-231.



Forth Seminar on  
**Reliability Theory and its Applications**  
April 25-26, 2018



## Three-state systems with different types of components

Ashrafi, S.<sup>1</sup>

Department of Statistics, University of Isfahan

### Abstract

In this paper, we consider a three-state system consisting of  $n$  binary components of two different types. We assume that component lifetimes of the same type are exchangeable and component lifetimes of two different types are independent. A mixture representation is obtained for the joint reliability function of the state lifetimes of the system. For this purpose, we generalize the concept of survival signature to the three-state systems and call it bivariate survival signature. The bivariate survival signature is computed for several systems composed of two different types of independent modules.

**Keywords:** Survival signature, Redundancy system, Reliability.

## 1 Introduction

Consider a three-state system consisting of  $n$  binary components having states  $K = 0, 1, 2$  where  $K = 2$  denotes the up state,  $K = 1$  corresponds to the partial performance state and  $K = 0$  represents the down state. Let  $T_1$  denote the lifetime of the system at state  $K = 2$  and  $T_2$  denote the lifetime of the system. Gertsbakh and Shpungin [5] showed that the joint reliability function of  $T_1$  and  $T_2$  can be obtained as

$$P(T_1 > t_1, T_2 > t_2) = \sum_{i=1}^n \sum_{j=i+1}^n s_{i,j} P(X_{i:n} > t_1, X_{j:n} > t_2), \quad t_1 < t_2, \quad (1)$$

---

<sup>1</sup>s.ashrafi@sci.ui.ac.ir

where  $s_{i,j} = P(T_1 = X_{i:n}, T_2 = X_{j:n})$  is two-dimensional signature and  $X_{1:n}, \dots, X_{n:n}$  are the ordered lifetimes of the components. The representation (1) is true when  $X_1, \dots, X_n$  are independent and identical or exchangeable.

Da and Hu [2] presented a new method to obtain two-dimensional signature. Consider  $P$  and  $L$  as the subsets of  $\{1, 2, \dots, n\}$  such that when components in  $P$  ( $L$ ) are working the system is not at the down state (is at state  $K = 2$ ).  $P$  ( $L$ ) is called path set (perfection path set). Let  $\mathbf{P}_j$  ( $\mathbf{L}_i$ ) denote the set of all path sets of size  $j$  (perfection path sets of size  $i$ ). Define  $r_{i,j}$  as the number of  $(L, P)$  such that  $L$  is of size  $i$  and  $P$  is of size  $j$  and  $P \subseteq L$  that is

$$r_{i,j} = \#\{(L, P) | L \in \mathbf{L}_i, P \in \mathbf{P}_j, P \subseteq L\}.$$

Assume that  $p_{i,j} = \frac{r_{i,j}}{\binom{n}{i-j}}$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, i$ . Da and Hu [2] showed if component lifetimes are i.i.d. or exchangeable then  $p_{i,j} = \bar{S}_{n-i, n-j}$  where  $\bar{S}_{i,j} = \sum_{k=i+1}^n \sum_{l=\max\{k,j\}+1}^n s_{i,j}$ .

In a recent work, Coolen and Coolen-Maturi [1], considered a binary system consisting of  $M$  types of different components that their lifetimes are independent. They introduced the notion of survival signature and obtained a mixture representation for the reliability function of the system. Eryilmaz [3] studied the reliability function of a binary system having two types of dependent components. Eryilmaz [4] investigated the case that system has  $M$  types of dependent components. In this paper, we consider a three-state system composed of two different types of components. We assume that the lifetimes of the components of the same type are exchangeable and the lifetimes of the components of two different types are independent. In Section 2, the notion of bivariate survival signature, that is a generalization of the survival signature for three-state systems, is introduced. A mixture representation is obtained for the joint reliability function of state lifetimes of the system based on bivariate survival signature. In Section 3, bivariate survival signature for some systems consisting of two different types of modules that are independent are computed.

## 2 Bivariate survival signature

Consider a three-state system composed of  $n$  binary components. Assume that the system has two types of different components,  $n_1$  components of type  $I$  and  $n_2$  components of type  $II$ ,  $n = n_1 + n_2$ . Suppose that the component lifetimes of the same type are exchangeable while component lifetimes of different types are independent. Suppose that  $\mathbf{X} = (X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2})$  denotes the component lifetimes of the system such that  $\mathbf{X} = (X_1, \dots, X_{n_1})$  and  $\mathbf{Y} = (Y_1, \dots, Y_{n_2})$  represent the component lifetimes of type  $I$  and  $II$ , respectively.

Let  $L$  and  $P$  be the perfection path set and path set for the system, respectively. Consider  $r_{i_1, j_1, i_2, j_2}$  as the number of pairs  $(L, P)$  such that  $L$  ( $P$ ) is composed of  $i_1$  ( $j_1$ ) components of type  $I$  and  $i_2$  ( $j_2$ ) components of type  $II$  and  $P \subseteq L$ . In such a situation the bivariate survival signature is defined as

$$p_{i_1, j_1, i_2, j_2} = \frac{r_{i_1, j_1, i_2, j_2}}{\binom{n_1}{i_1-j_1, j_1} \binom{n_2}{i_2-j_2, j_2}},$$

where  $i_1 = 0, \dots, n_1$ ,  $j_1 = 0, \dots, i_1$ ,  $i_2 = 0, \dots, n_2$ ,  $j_2 = 0, \dots, i_2$ .

Assume that  $S_{n_1}^{(1)}(t)$  ( $S_{n_2}^{(2)}(t)$ ) denote the number of working components of type  $I$  ( $II$ ) until time  $t$ . The joint reliability function of  $T_1$  and  $T_2$  can be obtained as

$$P(T_1 > t_1, T_2 > t_2) = \sum_{i_1=0}^{n_1} \sum_{i_2=0}^{n_2} \sum_{j_1=0}^{i_1} \sum_{j_2=0}^{i_2} p_{i_1, j_1, i_2, j_2} \\ \times P(S_{n_1}^{(1)}(t_1) = i_1, S_{n_1}^{(1)}(t_2) = j_1) P(S_{n_2}^{(2)}(t_1) = i_2, S_{n_2}^{(2)}(t_2) = j_2), \quad t_1 < t_2.$$

Let the component lifetimes of type  $I$  ( $II$ ) are independent from distribution function  $F_1$  ( $F_2$ ). Then,

$$P(S_{n_k}^{(k)}(t_1) = i_k, S_{n_k}^{(k)}(t_2) = j_k) \\ = \binom{n_k}{i_k - j_k, j_k} F_k^{n_k - i_k}(t_1) (F_k(t_2) - F_k(t_1))^{i_k - j_k} \bar{F}_k^{j_k}(t_2), \quad t_1 < t_2, \quad k = 1, 2.$$

**Example 2.1.** Consider the bridge network shown in Figure 1 with two types of links, type 1 and type 2 and four nodes,  $a, b, c, d$ . Assume that links are subject to failure and nodes  $a, c, d$  are terminals. The network is at state  $K = 2$ , if all terminals are connected,  $K = 1$ , if two terminals are connected and  $K = 0$ , if all terminals are disconnected. Table 1 presents the non-zero values of bivariate survival signature for this network. Figure 2 presents the plot of  $\bar{F}(t_1, t_2)$  where  $\bar{F}_1(t) = e^{-t}$  and  $\bar{F}_2(t) = e^{-t^2}$  and link lifetimes are independent.

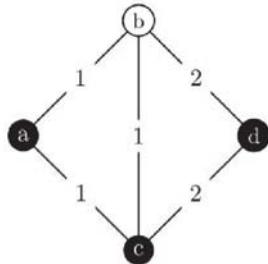


Figure 1: The bridge network

### 3 The bivariate survival signature of systems composed of two different types of modules

In this section, we compute the bivariate survival signature for different classes of three-state systems consisting of two independent three-state modules such that lifetimes of components in each module are exchangeable. We, also assume that module  $I$  is composed of  $n_1$  components of type  $I$  and module  $II$  is composed of  $n_2$  components of type  $II$ . Let  $T_1^{(k)}$  and  $T_2^{(k)}$  denote the lifetime of the  $k$ th module at state  $K = 2$  and the its lifetime,  $k = 1, 2$ . Let  $r_{i_1, j_1}^{(1)}$  ( $r_{i_2, j_2}^{(2)}$ ) denote the number of pairs  $(L_1, P_1)$  ( $(L_2, P_2)$ ) where  $L_1$  ( $L_2$ ) is perfection path set of size  $i_1$  ( $i_2$ ) and  $P_1$  ( $P_2$ ) is the path set of size  $j_1$  ( $j_2$ ) and  $P_1 \subseteq L_1$

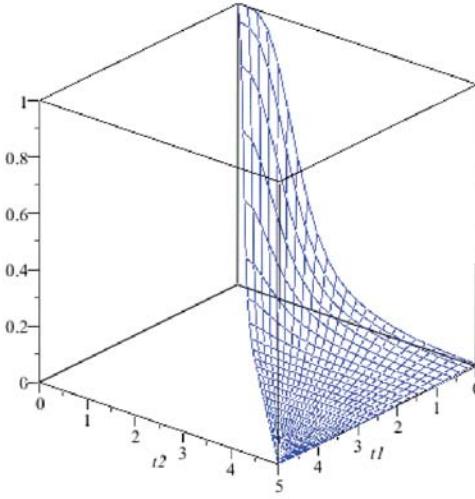


Figure 2: The joint reliability function of  $(T_1, T_2)$

$(P_2 \subseteq L_2)$  for the first module (second module). Assume that  $p_{i_1,j_1}^{(1)} = \frac{r_{i_1,j_1}^{(1)}}{\binom{n_1}{i_1-j_1,j_1}}$  and  $p_{i_2,j_2}^{(2)} = \frac{r_{i_2,j_2}^{(2)}}{\binom{n_2}{i_2-j_2,j_2}}$ . Let  $\bar{S}_{i_1,j_1}^{(1)}$  and  $\bar{S}_{i_2,j_2}^{(2)}$  denote the survival function corresponding to the two-dimensional signature of the first module and second module, respectively. Due to the result of Da and Hu [2],  $p_{i_1,j_1}^{(1)} = \bar{S}_{n_1-i_1,n_1-j_1}^{(1)}$  and  $p_{i_2,j_2}^{(2)} = \bar{S}_{n_2-i_2,n_2-j_2}^{(2)}$ .

First, we assume that that two three-state modules are connected in series. It is clear that the lifetime of system at state  $K = 2$ , denoted by  $T_{S,1}$ , and the lifetime of the system, denoted by  $T_{S,2}$ , are given as

$$T_{S,1} = \min\{T_1^{(1)}, T_1^{(2)}\}, \quad T_{S,2} = \min\{T_2^{(1)}, T_2^{(2)}\}.$$

It can be seen that the bivariate survival signature for this system is obtained as

$$\begin{aligned} p_{i_1,j_1,i_2,j_2} &= \frac{r_{i_1,j_1}^{(1)} r_{i_2,j_2}^{(2)}}{\binom{n_1}{i_1-j_1,j_1} \binom{n_2}{i_2-j_2,j_2}} \\ &= p_{i_1,j_1}^{(1)} p_{i_2,j_2}^{(2)} \\ &= \bar{S}_{n_1-i_1,n_1-j_1}^{(1)} \bar{S}_{n_2-i_2,n_2-j_2}^{(2)}. \end{aligned}$$

Now, we assume that two three-state modules are connected in parallel. It is clear that the lifetime of system at state  $K = 2$ , denoted by  $T_{P,1}$ , and the lifetime of the system, denoted by  $T_{P,2}$ , are given as

$$T_{P,1} = \max\{T_1^{(1)}, T_1^{(2)}\}, \quad T_{P,2} = \max\{T_2^{(1)}, T_2^{(2)}\}.$$

Table 1: Bivariate survival signature for bridge network in Figure 1

$i_1$	$j_1$	$i_2$	$j_2$	$p_{i_1,j_1,i_2,j_2}$	$i_1$	$j_1$	$i_2$	$j_2$	$p_{i_1,j_1,i_2,j_2}$
1	0	1	1	1/6	2	2	2	2	1
1	0	2	1	2/6	3	0	1	1	1/2
1	0	2	2	2/3	3	0	2	1	1/2
1	1	1	0	1/6	3	0	2	2	1
1	1	1	1	1/6	3	1	1	0	2/6
1	1	2	0	1/3	3	1	1	1	1
1	1	2	1	4/6	3	1	2	0	1/3
1	1	2	2	2/3	3	1	2	1	1
2	0	1	1	3/6	3	1	2	2	1
2	0	2	1	3/6	3	2	1	0	1
2	0	2	2	1	3	2	1	1	1
2	1	1	0	4/12	3	2	2	0	1
2	1	1	1	12/12	3	2	2	1	1
2	1	2	0	2/6	3	2	2	2	1
2	1	2	1	1	3	3	1	0	1
2	1	2	2	1	3	3	1	1	1
2	2	1	0	1	3	3	2	0	1
2	2	1	1	1	3	3	2	1	1
2	2	2	0	1	3	3	2	2	1
2	2	2	1	1					

It can be seen that the bivariate survival signature for this system is obtained as

$$\begin{aligned}
 p_{i_1,j_1,i_2,j_2} &= \frac{r_{i_1,j_1}^{(1)}}{\binom{n_1}{i_1-j_1,j_1}} + \frac{r_{i_2,j_2}^{(2)}}{\binom{n_2}{i_2-j_2,j_2}} - \frac{r_{i_1,j_1}^{(1)} r_{i_2,j_2}^{(2)}}{\binom{n_1}{i_1-j_1,j_1} \binom{n_2}{i_2-j_2,j_2}} \\
 &= p_{i_1,j_1}^{(1)} + p_{i_2,j_2}^{(2)} - p_{i_1,j_1}^{(1)} p_{i_2,j_2}^{(2)} \\
 &= \bar{S}_{n_1-i_1,n_1-j_1}^{(1)} + \bar{S}_{n_2-i_2,n_2-j_2}^{(2)} - \bar{S}_{n_1-i_1,n_1-j_1}^{(1)} \bar{S}_{n_2-i_2,n_2-j_2}^{(2)}.
 \end{aligned}$$

In the following, we consider component-wise redundancy in three-state systems. Consider a three-state system with  $n$  components. Suppose that for each component of the original system a spares is placed in parallel. Assume that components of two systems (original system and system composed of spares components) have two different distributions. Let component lifetimes of two systems are independent and component lifetimes in each system are exchangeable. Assume that  $r_{i_1,j_1}$  denotes the number of pairs  $(L_1, P_1)$  where  $L_1$  is perfection path set of size  $i_1$  and  $P_1$  is the path set of size  $j_1$  and  $P_1 \subseteq L_1$  for the original system. In the following theorem, we compute the bivariate survival signature in terms of  $p_{i_1,j_1}$ .

**Theorem 3.1.** *Consider a three-state system with component-wise redundancy structure. For  $0 \leq j_1 \leq i_1 \leq n$  and  $0 \leq j_2 \leq i_2 \leq n$ , we have*

$$p_{i_1,j_1,i_2,j_2} = \frac{r_{i_1,j_1,i_2,j_2}}{\binom{n}{i_1-j_1,j_1} \binom{n}{i_2-j_2,j_2}}$$

where

$$r_{i_1, j_1, i_2, j_2} = \sum_{l=l^*}^{i \wedge n} \sum_{m=m^*}^{l \wedge j} \sum_{k_2=k_2^*}^{m-j_2} \sum_{k_1=k_1^*}^{k_1^{**}} \binom{m}{j_1} \binom{j_1}{j-m} \binom{m-j_1}{k_1} \binom{m-j_2}{k_2} \\ \times \binom{l-m}{i_1-j_1-k_1} \binom{i_1-j_1-k_1}{i-j-k_1-k_2-l+m} \binom{n}{l-m, m} p_{l,m}$$

where

$$l^* = \lfloor \frac{i-1}{2} \rfloor + 1, \quad m^* = (j-i+l) \vee (\lfloor \frac{j-1}{2} \rfloor + 1), \\ k_1^* = i_1 - j_1 - l + m, \quad k_1^{**} = (m-j_1) \wedge (i-j-l+m-k_2), \\ k_2^* = i_2 - j_2 - l + m, \quad i = i_1 + i_2, \quad j = j_1 + j_2.$$

*Proof.* Consider the redundancy system as a system consists of  $n$  modules where each module is composed of a original component and a sparse component that are placed in parallel. Let  $N_{l,m;i_1,j_1,i_2,j_2}$  denote the number of pairs  $(L, P)$  such that  $L$  ( $P$ ) is composed of  $i_1$  ( $j_1$ ) original components and  $i_2$  ( $j_2$ ) sparse components that  $i = i_1 + i_2$  ( $j = j_1 + j_2$ ) components belong to  $l$  ( $m$ ) different modules and  $P \subseteq L$ . It is clear that  $l \in \{l^*, \dots, i \wedge n\}$  and  $m \in \{m^*, \dots, l \wedge j\}$  such that  $l^* = \lfloor \frac{i-1}{2} \rfloor + 1$  and  $m^* = (j-i+l) \vee (\lfloor \frac{j-1}{2} \rfloor + 1)$ . We compute the  $N_{l,m;i_1,j_1,i_2,j_2}$  at two steps. At the first step, we choose  $j_1$  original components and  $j_2$  spares components such that from each of  $m$  modules at least one component is chosen. The number of all ways that can be chosen is

$$N_1 = \binom{m}{j_1} \binom{j_1}{j-m}.$$

At the second step, we choose  $i_1 - j_1$  original components and  $i_2 - j_2$  sparse components such that from each of  $l - m$  modules at least one component is chosen. To this end, we choose  $k_1$  original components and  $k_2$  sparse components from  $m$  modules where  $k_1 \in \{i_1 - j_1 - l + m, \dots, (m - j_1) \wedge (i - j - l + m - k_2)\}$  and  $k_2 \in \{i_2 - j_2 - l + m, \dots, m - j_2\}$ . Then, we choose  $(i_1 - j_1 - k_1)$  original components and  $(i_2 - j_2 - k_2)$  sparse components such that from each of  $l - m$  modules at least one component is chosen. The number of all ways that can be chosen is

$$N_2 = \sum_{k_2=k_2^*}^{m-j_2} \sum_{k_1=k_1^*}^{k_1^{**}} \binom{m-j_1}{k_1} \binom{m-j_2}{k_2} \binom{l-m}{i_1-j_1-k_1} \binom{i_1-j_1-k_1}{i-j-k_1-k_2-l+m},$$

where  $k_1^* = i_1 - j_1 - l + m$  and  $k_1^{**} = (m - j_1) \wedge (i - j - l + m - k_2)$ ,  $k_2^* = i_2 - j_2 - l + m$ . Therefore,  $N_{l,m;i_1,j_1,i_2,j_2} = N_1 N_2$ .

Let  $r_{i_1, j_1, i_2, j_2}$  denote the number of  $(L, P)$  such that  $L$  ( $P$ ) is a perfection path set (path set) that is composed of  $i_1$  ( $j_1$ ) original components and  $i_2$  ( $j_2$ ) sparse components and  $P \subseteq L$ . For every pair  $(L, P)$  there is a unique  $l \in \{l^*, \dots, i \wedge n\}$  and a unique  $m \in \{m^*, \dots, l \wedge j\}$  satisfying

- (I)  $L$  is composed of components of  $l$  various modules, and original components in  $l$  modules form a perfection path set for the original system.

- (II)  $P$  is composed of components of  $m$  various modules, and original components in  $m$  modules form a path set for the original system.
- (III) The path set of size  $m$  is a subset of perfection path set of size  $l$  for the original system.

Since, the collection of all pairs  $(L, P)$  can be written as

$$\bigcup_{l=l^*}^{i \wedge n} \bigcup_{m=m^*}^{l \wedge j} W_{l,m,i_1,j_1,i_2,j_2}$$

where  $W_{l,m,i_1,j_1,i_2,j_2}$  denotes the collection of all pairs of  $(L, P)$  that satisfy (I) and (II), (III). It can be seen that

$$|W_{l,m,i_1,j_1,i_2,j_2}| = N_{l,m;i_1,j_1,i_2,j_2} r_{l,m},$$

where  $r_{l,m}$  denotes the number of pairs  $(L_1, P_1)$  such that  $L_1$  ( $P_1$ ) is a perfection path set (path set) that is composed of  $l$  ( $m$ ) original components and  $P_1 \subseteq L_1$ . Hence,

$$\begin{aligned} r_{i_1,i_2,j_1,j_2} &= \sum_{l=l^*}^{i \wedge n} \sum_{m=m^*}^{l \wedge j} |W_{l,m,i_1,j_1,i_2,j_2}| \\ &= \sum_{l=l^*}^{i \wedge n} \sum_{m=m^*}^{l \wedge j} N_{l,m;i_1,j_1,i_2,j_2} r_{l,m}. \end{aligned}$$

Therefore, the required result follows from the fact that  $p_{l,m} = \frac{r_{l,m}}{\binom{n}{l-m}}$ .  $\square$

In the following, we consider two binary systems with  $n_1$  and  $n_2$  components and structure functions  $\phi_1(x_1, \dots, x_{n_1})$  and  $\phi_2(y_1, \dots, y_{n_2})$ , respectively. Assume that the component lifetimes of each system are exchangeable. Suppose that component lifetimes of two systems are independent from two different distributions. Consider a three-state system with structure function

$$\varphi(x_1, \dots, x_{n_1}; y_1, \dots, y_{n_2}) = \phi_1(x_1, \dots, x_{n_1}) + \phi_2(y_1, \dots, y_{n_2}). \quad (2)$$

Let  $T_1$  and  $T_2$  denote the lifetime of the system at state  $K = 2$  and the lifetime of the system, respectively. It is clear that

$$T_1 = \min\{T_1^{(1)}, T_1^{(2)}\}, \quad T_2 = \max\{T_2^{(1)}, T_2^{(2)}\}.$$

In the following theorem, we compute bivariate survival signature for this system.

**Theorem 3.2.** *Consider the three-state with structure function defined in (2). The bivariate survival signature is given as*

$$p_{i_1,j_1,i_2,j_2} = p_{1,j_1} p_{2,i_2} + p_{1,i_1} p_{2,j_2} - p_{1,j_1} p_{2,j_2}$$

where  $p_{1,i}$  and  $p_{2,j}$  are the survival signatures of the systems with structure functions  $\phi_1$  and  $\phi_2$ , respectively.

*Proof.* Let  $r_{i_1,j_1,i_2,j_2}$  denote the number of  $(L, P)$  such that  $L$  ( $P$ ) is a perfection path set (path set) that is composed of  $i_1$  ( $j_1$ ) components of the system with structure function  $\phi_1$  and  $i_2$  ( $j_2$ ) components of the system with structure function  $\phi_2$  and  $P \subseteq L$ . Consider the following events:

$A_1$  : The system with structure  $\phi_1$  works while  $i_1$  components are active.

$A_2$  : The system with structure  $\phi_2$  works while  $i_2$  components are active.

$B_1$  : The system with structure  $\phi_1$  works while  $j_1$  components are active.

$B_2$  : The system with structure  $\phi_2$  works while  $j_2$  components are active.

It can be seen that the bivariate survival signature is obtained as

$$\begin{aligned}
& p_{i_1,j_1,i_2,j_2} \\
&= P\{(A_1 \cap A_2) \cap [(B_1 \cap B'_2) \cup (B'_1 \cap B_2) \cup (B_1 \cap B_2)]\} \\
&= P(A_1 \cap B_1)P(A_2 \cap B'_2) + P(A_2 \cap B_2)P(A_1 \cap B'_1) + P(A_1 \cap B_1)P(A_2 \cap B_2) \\
&= P(B_1)[P(A_2) - P(B_2)] + P(B_2)[P(A_1) - P(B_1)] + P(B_1)P(B_2) \\
&= p_{1,j_1}p_{2,i_2} + p_{1,i_1}p_{2,j_2} - p_{1,j_1}p_{2,j_2}
\end{aligned}$$

□

## References

- [1] Coolen, F. P. and Coolen-Maturi, T. (2013). Generalizing the signature to systems with multiple types of components. *Complex systems and dependability*, 115-130.
- [2] Da, G. and Hu, T. (2013). On bivariate signatures for systems with independent modules. *Stochastic Orders in Reliability and Risk*, 143-166.
- [3] Eryilmaz, S. (2015). Systems composed of two types of nonidentical and dependent components. *Naval Research Logistics* **62**, 388-394.
- [4] Eryilmaz, S. (2016). Reliability of systems with multiple types of dependent components. *IEEE Transactions on Reliability* **65**, 1022-1029.
- [5] Gertsbakh, I. and Shpungin, Y. (2012) Stochastic models of network survivability. *QTQM* **9**, 45-58.



*Forth Seminar on*  
**Reliability Theory and its Applications**

*April 25-26, 2018*



## Planning of step-stress accelerated degradation test with tampered failure rate model under an inverse Gaussian process

Azizi, F.<sup>1</sup>      Mosayebi Omshi, E.<sup>2</sup>      Haghghi, F.<sup>3</sup>      Shemehsavar, S.<sup>4</sup>

<sup>1,2,3,4</sup> School of Mathematics, Statistics and Computer Sciences, University of Tehran

### Abstract

In this article, we consider the planning of simple step-stress accelerated degradation test (SSADT) when the degradation follows an inverse Gaussian (IG) process. We apply the idea of the TFR model to describe the effect of the changing stress on the degradation and obtain maximum likelihood estimates (MLEs) of the model parameters. Under the constraint that the total experimental cost does not exceed a pre-specified budget, the optimal settings such as sample size, measurement frequency, and number of measurements at each stress level are obtained using variance optimality. Finally, a real-world example is analyzed to illustrate the application of the proposed methods.

**Keywords:** Step-Stress Accelerated Degradation Test, Tampered Failure Rate Model, Inverse Gaussian Process, Optimal Test Plan.

## 1 Introduction

Most modern products are designed to operate without failure for years. Thus few units will fail or degrade during a traditional life test. In such situations, accelerated life tests (ALTs) are commonly used. In ALTs, products are subjected to severe conditions to obtain reliability information sooner. For this reason, ALTs have become important

---

<sup>1</sup>fariba.azizi@khayam.ut.ac.ir

technical methods in reliability. There are several statistical models which can be used to extrapolate the results of step-stress ALTs: the tampered random variable (TRV) model described by [2], the cumulative exposure (CE) model illustrated to [4], and the tampered failure rate (TFR) model proposed by [1].

Moreover, in some reliability studies, it is possible to measure the amount of degradation of a unit over time. Such measures often provide more information than failure time data for purposes of assessing and improving a product's reliability. Degradation data are generally modeled by deterministic models such as the general path models or stochastic models using the Wiener process, Gamma process, or Inverse Gaussian (IG) process. We can name some research works such as [2], [5], [6] and [10] where the degradation path follows a Wiener process or Gamma process.

It is known that the Wiener and Gamma processes cannot handle all degradation data in real-world problems. For example, [7] found that neither Wiener nor Gamma processes fit some GaAs laser degradation data, while the IG process model performs well in fitting these data. The IG process has a monotone increasing behavior and has good properties when dealing with covariates and random effects. Regarding an IG process as a degradation model, [9] considered the problem of finding an optimal constant-stress accelerated degradation test (CSADT) under the constraint that the experiment cost does not exceed a pre-specified budget. Recently, [7] have presented an optimal SSADT plan for degradation data based on the IG process. They adopted the CE model to express the effect of cumulative degradation.

In this paper, we apply the idea of the TFR model to describe the effect of the change of stress on the degradation, that is to say, the stress change has a multiplicative effect on the degradation path. This model is an alternative for the model proposed by [7]. In comparison with the model proposed by them, our model results in closed form MLEs for majority of the model parameters are consequently easier to fit. Then, under the constraint that the total experimental cost does not exceed a predetermined budget, the optimal settings of an SSADT are obtained by minimizing the asymptotic variance of the estimated  $p$ -quantile of the products lifetime distribution.

The remainder of the paper is organized as follows. In Section 2, we describe the IG process, the SSADT and the required assumptions. The estimation problem is discussed in Section 3. In Section 4, an optimal test plan is presented. Finally, a real-world example is analyzed to illustrate the application of the proposed methods in Section 5.

## 2 Degradation modeling and assumptions

**Assumption 1:** Let  $Z(t|S_0)$  denote the degradation path of the product under normal stress ( $S_0$ ). The product is assumed to fail when the degradation path  $Z(t|S_0)$  for the first time crosses a pre-specific threshold level  $\omega$ . The first passage time,  $T_\omega$ , is called the product's lifetime. In this paper,  $Z(t|S_0)$  is modeled as an IG degradation process, having the properties:

- $Z(t|S_0)$  has independent increments; that is,  $Z(t_{i+1}|S_0) - Z(t_i|S_0)$  and  $Z(t_{j+1}|S_0) - Z(t_j|S_0)$  are independent  $\forall t_{j+1} > t_j \geq t_{i+1} > t_i$ .

- Each increment,  $\Delta Z(t|S_0) = Z(t + \Delta t|S_0) - Z(t|S_0)$ , follows an Inverse Gaussian distribution  $IG(\mu_0\Delta t, \lambda_0\Delta t^2)$  with the probability density function (PDF) given by:

$$f_{\Delta Z}(y) = \sqrt{\frac{\lambda_0\Delta t^2}{2\pi y^3}} \exp\left(-\frac{\lambda_0(y - \mu_0\Delta t)^2}{2y\mu_0^2}\right), \quad y > 0,$$

where  $\mu_0\Delta t$  and  $\lambda_0\Delta t^2$  are the mean and scale parameter for fixed  $\Delta t$ , respectively.

The cumulative distribution function (CDF) of  $T_\omega = \inf\{t|Z(t|S_0) > \omega\}$  can be simply obtained by:

$$\begin{aligned} F_{T_\omega} &= P(T_\omega < t) = P(Z(t|S_0) > \omega) \\ &= \Phi[\sqrt{\lambda_0/\mu_0}(t - \omega/\mu_0)] - e^{2\lambda_0 t/\mu_0} \Phi[-\sqrt{\lambda_0/\mu_0}(t + \omega/\mu_0)], \end{aligned} \quad (1)$$

where  $\Phi$  is the standard normal CDF. Based on the CDF in Eq. 1, there is no closed form for the  $p$ -quantile of lifetime distribution and [9] noted that the CDF of  $T_\omega$  can be approximated as:

$$F_{T_\omega} = 1 - \Phi\left(\frac{\omega - \mu_0 t}{\sqrt{\mu_0^3 t / \lambda_0}}\right) = \Phi\left(\sqrt{\lambda_0 t / \mu_0} - \omega \sqrt{\frac{\lambda_0 / \mu_0^3}{t}}\right).$$

Using this approximation, the  $p$ -quantile of  $T_\omega$  is:

$$q_p = \frac{\mu_0}{4\lambda_0} \left( z_p + \sqrt{z_p^2 + 4\omega\lambda_0/\mu_0^2} \right)^2 \quad (2)$$

where  $z_p$  is the standard normal  $p$ -quantile.

**Assumption 2:** A simple step-stress ADT with two stress levels  $S_1$  and  $S_2$  ( $S_1 < S_2$ ), is considered where  $S_1$  and  $S_2$  are assumed to be pre-specified. The test is conducted as follows. Suppose the test units are initially placed on stress  $S_1$  and the degradation data are recorded with a measurement frequency per  $f$  units of time. The stress level remains at  $S_1$  until stress change time  $\tau_1$ . Next, stress is increased to higher stress level  $S_2$  and held constant until the test termination time  $\tau_2$ . Therefore, we have  $\tau_1 = \ell_1 \times f$  and  $\tau_2 = (\ell_1 + \ell_2) \times f$  where  $\ell_1$  and  $\ell_2$  are the numbers of measurements under stress  $S_1$  and  $S_2$ , respectively.

**Assumption 3:** Based on the idea of [1], we represent a statistical model for simple step-stress ADT in which a change of stress has a multiplicative effect on the degradation path. According to this assumption, the degradation path of the simple step-stress ADT of the product at time  $t$ ,  $Z_{SS}(t)$ , can be expressed as:

$$Z_{SS}(t) = \begin{cases} Z(t|S_1) & 0 < t \leq \tau_1 \\ \alpha Z(t|S_1) & \tau_1 < t \leq \tau_2, \end{cases} \quad (3)$$

where  $Z(t|S_1)$  is the degradation path of the product under the stress  $S_1$  and  $\alpha \geq 1$  is the acceleration factor which is an unknown parameter of the model. Further, we assume that the independent increments of the degradation path at the first stress level,  $S_1$ , follow an IG distribution with the mean  $\mu_1\Delta t$  and the scale parameter  $\lambda_1\Delta t^2$  for fixed  $\Delta t$ . Then, according to the properties of the IG distribution, the distribution of the independent increments of the degradation path at the second stress level,  $S_2$ , will be an IG distribution with mean  $\alpha\mu_1\Delta t$  and scale parameter  $\alpha\lambda_1\Delta t^2$  for fixed  $\Delta t$ .

### 3 Maximum likelihood estimation

In this section, we deal with the estimation of the unknown parameters based on the maximum likelihood and Bayesian methods. Suppose  $n$  units are placed on a step-stress ADT and let  $Z_{SS}^{(i)}(t)$  denote the degradation value of the  $i$ -th unit at time  $t$ . For  $0 < t_{j-1} < t_j \leq \tau_1$ , define the degradation increment as follows:

$$\begin{aligned} Y_{ij} &= Z_{SS}^{(i)}(t_j) - Z_{SS}^{(i)}(t_{j-1}) \\ &= Z^{(i)}(t_j|S_1) - Z^{(i)}(t_{j-1}|S_1), \end{aligned}$$

for  $i = 1, \dots, n$  and  $j = 1, \dots, \ell_1$ . Then, from the above assumptions, we have:

$$Y_{ij} \sim IG(\mu_1 \Delta t_j, \lambda_1 \Delta t_j^2),$$

while for  $\tau_1 < t_{j-1} < t_j \leq \tau_2$ , we know that  $Y_{ij} \sim IG(\alpha \mu_1 \Delta t_j, \alpha \lambda_1 \Delta t_j^2)$ , for  $i = 1, \dots, n$  and  $j = \ell_1 + 1, \dots, \ell_1 + \ell_2$ . Given the observed degradation values  $\{Z_{SS}^{(i)}(t_j)\}_{j=1}^{\ell_1+\ell_2}$ ,  $i = 1, \dots, n$ , the likelihood function based on the simple step-stress ADT model for an IG degradation process is obtained as:

$$\begin{aligned} L(\mu_1, \lambda_1, \alpha | \underline{y}) &= \prod_{i=1}^n \left\{ \prod_{j=1}^{\ell_1} \left( \frac{\lambda_1 \Delta t_j^2}{2\pi y_{ij}^3} \right)^{\frac{1}{2}} \exp\left(-\frac{\lambda_1(y_{ij} - \mu_1 \Delta t_j)^2}{2\mu_1^2 y_{ij}}\right) \right. \\ &\quad \times \left. \prod_{j=\ell_1+1}^{\ell_1+\ell_2} \left( \frac{\alpha \lambda_1 \Delta t_j^2}{2\pi y_{ij}^3} \right)^{\frac{1}{2}} \exp\left(-\frac{\lambda_1(y_{ij} - \alpha \mu_1 \Delta t_j)^2}{2\alpha \mu_1^2 y_{ij}}\right) \right\}. \end{aligned} \quad (4)$$

Now, the MLEs of the unknown parameters can be obtained by maximizing the logarithm of the likelihood function above with respect to the unknown parameters. For fixed  $\alpha$ , we can find an explicit form for the MLEs of  $\mu_1$  and  $\lambda_1$  as:

$$\hat{\mu}_1(\alpha) = \frac{\sum_{i=1}^n \sum_{j=1}^{\ell_1} y_{ij} + \frac{1}{\alpha} \sum_{i=1}^n \sum_{j=\ell_1+1}^{\ell_1+\ell_2} y_{ij}}{\sum_{i=1}^n \sum_{j=1}^{\ell_1+\ell_2} \Delta t_j}, \quad (5)$$

and

$$\hat{\lambda}_1(\alpha, \hat{\mu}_1(\alpha)) = \frac{n(\ell_1 + \ell_2)}{2 \sum_{i=1}^n \left[ \sum_{j=1}^{\ell_1} \frac{(y_{ij} - \hat{\mu}_1 \Delta t_j)^2}{2\hat{\mu}_1^2 y_{ij}} + \sum_{j=\ell_1+1}^{\ell_1+\ell_2} \frac{(y_{ij} - \alpha \hat{\mu}_1 \Delta t_j)^2}{2\alpha \hat{\mu}_1^2 y_{ij}} \right]}. \quad (6)$$

The likelihood equation for  $\alpha$  is obtained as:

$$\frac{n\ell_2}{2\alpha} + \hat{\lambda}_1 \sum_{i=1}^n \sum_{j=\ell_1+1}^{\ell_1+\ell_2} \frac{y_{ij}^2 - \alpha^2 \hat{\mu}_1^2 \Delta t_j^2}{2\alpha^2 \hat{\mu}_1^2 y_{ij}} = 0. \quad (7)$$

Therefore, the MLE of  $\alpha$  can be derived by substituting  $\hat{\mu}_1(\alpha)$  and  $\hat{\lambda}_1(\alpha, \hat{\mu}_1(\alpha))$  into Eq. 7 and solving this equation for the unknown parameter  $\alpha$ . Once  $\hat{\alpha}$  is obtained, the MLE of  $\mu_1$  and  $\lambda_1$  can be obtained as  $\hat{\mu}_1(\hat{\alpha})$  and  $\hat{\lambda}_1(\hat{\alpha}, \hat{\mu}_1(\hat{\alpha}))$ , respectively.

## 4 Optimization

To find the optimal plan of the SSADT, we treat the sample size  $n$ , measurement frequency  $f$ , and the number of measurements  $\ell_i$  for the stress level  $S_i$ ,  $i = 1, 2$ , as the design variables. Let  $\hat{q}_p$  be the estimated  $p$ -quantile of the lifetime distribution, which can be obtained by substituting the estimated parameters in Eq. (2). The asymptotic variance of  $\hat{q}_p$ , i.e.  $\text{AVar}(\hat{q}_p|n, f, \ell_1, \ell_2)$ , is chosen to be the objective function. Assume that  $\text{TC}(n, f, \ell_1, \ell_2)$  denotes the total experiment cost that should not exceed a pre-specified budget  $C_b$ . Clearly, both the asymptotic variance of  $\hat{q}_p$  and the total experiment cost are affected by design variables  $n, f, \ell_1, \ell_2$ . Then a typical problem of interest can be formulated as:

$$\text{Minimize: } \text{AVar}(\hat{q}_p|n, f, \ell_1, \ell_2),$$

$$\text{subject to: } \text{TC}(n, f, \ell_1, \ell_2) \leq C_b,$$

where  $n, f, \ell_1, \ell_2 \in \mathbb{N}$ . The optimal design variables  $n^*, f^*, \ell_1^*, \ell_2^*$  can be obtained using the Genetic algorithm in the Matlab software. In the following subsections, we will discuss the asymptotic variance of the estimated  $p$ -quantile of the lifetime distribution and develop a formula for the total experiment cost.

### 4.1 Computation of $\text{AVar}(\hat{q}_p|n, f, \ell_1, \ell_2)$

Herein, we employ a reparameterization as follows:

$$\eta_1 = \mu_1, \quad \eta_2 = \alpha\mu_1, \quad \phi = \frac{\lambda_1}{\mu_1}.$$

Also, we use the notation  $\eta_0$  to refer to the degradation rate in the normal use condition; that is,  $\eta_0 = \mu_0$ . We suppose that the relationship between the mean rate of degradation and the corresponding stress level is given by the following log-linear formula:

$$\log \eta_i = \beta_0 + \beta_1 S_i, \quad i = 0, 1, 2.$$

The asymptotic variance of  $\hat{q}_p$  can then be derived using the delta method as:

$$\text{AVar}(\hat{q}_p) = C(\Theta)' I^{-1}(\Theta) C(\Theta),$$

where  $C(\Theta)$  is the first derivative of  $q_p$  with respect to  $\Theta = (\beta_0, \beta_1, \phi)$ ,

$$C(\Theta)' = \left( \frac{\partial q_p}{\partial \beta_0}, \frac{\partial q_p}{\partial \beta_1}, \frac{\partial q_p}{\partial \phi} \right), \quad (8)$$

and  $I^{-1}(\Theta)$  is the inverse of the Fisher information matrix  $I(\Theta)$  which can be computed by:

$$I(\Theta) = \begin{pmatrix} E\left(-\frac{\partial^2 l(\Theta)}{\partial \beta_0^2}\right) & E\left(-\frac{\partial^2 l(\Theta)}{\partial \beta_0 \partial \beta_1}\right) & E\left(-\frac{\partial^2 l(\Theta)}{\partial \beta_0 \partial \phi}\right) \\ E\left(-\frac{\partial^2 l(\Theta)}{\partial \beta_1 \partial \beta_0}\right) & E\left(-\frac{\partial^2 l(\Theta)}{\partial \beta_1^2}\right) & E\left(-\frac{\partial^2 l(\Theta)}{\partial \beta_1 \partial \phi}\right) \\ E\left(-\frac{\partial^2 l(\Theta)}{\partial \phi \partial \beta_0}\right) & E\left(-\frac{\partial^2 l(\Theta)}{\partial \phi \partial \beta_1}\right) & E\left(-\frac{\partial^2 l(\Theta)}{\partial \phi^2}\right) \end{pmatrix}. \quad (9)$$

With the reparameterization  $\Theta = (\beta_0, \beta_1, \phi)$  the  $p$ -quantile of the lifetime distribution can be rewritten as:

$$q_p = \frac{1}{4\phi} \left( z_p + (z_p + 4\omega\phi e^{-(\beta_0+\beta_1 S_0)})^{\frac{1}{2}} \right)^2.$$

Then, the elements of the vector  $C(\Theta)'$  can be expressed as:

$$\begin{aligned} \frac{\partial q_p}{\partial \beta_0} &= -\omega e^{-(\beta_0+\beta_1 S_0)} [1 + z_p(z_p + 4\omega\phi e^{-(\beta_0+\beta_1 S_0)})^{-\frac{1}{2}}], \\ \frac{\partial q_p}{\partial \beta_1} &= -\omega S_0 e^{-(\beta_0+\beta_1 S_0)} [1 + z_p(z_p + 4\omega\phi e^{-(\beta_0+\beta_1 S_0)})^{-\frac{1}{2}}], \\ \frac{\partial q_p}{\partial \phi} &= -\frac{1}{4\phi^2} \left( z_p + (z_p + 4\omega\phi e^{-(\beta_0+\beta_1 S_0)})^{\frac{1}{2}} \right)^2 \\ &\quad + \frac{\omega}{\phi} e^{-(\beta_0+\beta_1 S_0)} [1 + z_p(z_p + 4\omega\phi e^{-(\beta_0+\beta_1 S_0)})^{-\frac{1}{2}}]. \end{aligned}$$

Taking into account that  $t_j = jf$ , in the likelihood function with parameters  $\Theta = (\beta_0, \beta_1, \phi)$ , the elements of the Fisher information  $I(\Theta)$  in Eq. (9) can be obtained:

$$\begin{aligned} E\left(-\frac{\partial^2 l}{\partial \beta_0^2}\right) &= \frac{n}{2}(\ell_1 + \ell_2)(1 + 2\phi f), & E\left(-\frac{\partial^2 l}{\partial \beta_1^2}\right) &= \frac{n}{2}(\ell_1 S_1^2 + \ell_2 S_2^2)(1 + 2\phi f), \\ E\left(-\frac{\partial^2 l}{\partial \phi^2}\right) &= \frac{n}{2\phi^2}(\ell_1 + \ell_2), & E\left(-\frac{\partial^2 l}{\partial \beta_0 \partial \beta_1}\right) &= \frac{n}{2}(\ell_1 S_1 + \ell_2 S_2)(1 + 2\phi f), \\ E\left(-\frac{\partial^2 l}{\partial \beta_0 \partial \phi}\right) &= \frac{n}{2\phi}(\ell_1 + \ell_2), & E\left(-\frac{\partial^2 l}{\partial \beta_1 \partial \phi}\right) &= \frac{n}{2\phi}(\ell_1 S_1 + \ell_2 S_2). \end{aligned}$$

## 4.2 The total cost of experimentation

The total cost of the experiment  $TC(n, f, \ell_1, \ell_2)$  consists of three parts:

- the cost of conducting an experiment is  $C_{op}f \sum_{i=1}^2 \ell_i$ ;
- the cost of measurement is  $C_{mea}n \sum_{i=1}^2 \ell_i$ ; and
- the cost of testing devices is  $nC_{it}$ ;

where  $C_{op}$ ,  $C_{mea}$ , and  $C_{it}$  are the unit cost of operation, the unit cost of measurement, and the unit cost of an item, respectively. Hence, the total cost of the experiment can be formulated as:

$$TC(n, f, \ell_1, \ell_2) = C_{op}f \sum_{i=1}^2 \ell_i + C_{mea}n \sum_{i=1}^2 \ell_i + nC_{it}.$$

## 5 An illustrative example

In this section, we illustrate our proposed method through a numerical example based on the data from the stress relaxation problem described by [8]. The stress relaxation data

has been collected based on the CSADT. So, we adopted it for our work and computed the parameter estimates using Eq. (5), (6) and (7). As a result, parameters  $\alpha$ ,  $\mu_1$ , and  $\lambda_1$  are estimated as 1.4995, 0.0034, and  $5.25 \times e^{-05}$ , respectively. To obtain the optimal SSADT plan, we consider the following cost configurations:

$$\begin{aligned} C_{op} &= \$2.5/\text{per unit time}, \\ C_{mea} &= \$1.5/\text{per measurement, and} \\ C_{it} &= \$50/\text{per unit}. \end{aligned}$$

Also, we assume the two stress levels are  $S_1 = 65^\circ\text{C}$  and  $S_2 = 95^\circ\text{C}$ .

Based on the proposed method, the optimal test plan for various budgets  $C_b$  are obtained in Table 1. For example, when  $C_b = 1500$ , the optimal test plan is  $(n^*, l_1^*, l_2^*, f^*) = (11, 16, 10, 8)$ ; that is, the optimal sample size is 11, the optimal measurement frequency is 8, and the corresponding optimal numbers of measurements for stress  $S_1$  and  $S_2$  are 16 and 10, respectively. This means that, based on the optimal plan, the total test time for the SSADT experiment is  $8 \times (16 + 10) = 208$  units of time, the total cost is \$1499, and the corresponding approximate standard deviation is 3004.896 .

Table 1: Optimal test plan under different  $C_b$ .

$C_b$	$n^*$	$l_1^*$	$l_2^*$	$f^*$	total cost	$std(\hat{q}_{0.1})$
1500	11	16	10	8	1499	3004.896
2000	15	17	8	11	2000	2381.155
2500	18	20	8	12	2496	2006.689
3000	22	18	8	16	2998	1735.425

## References

- [1] Bhattacharyya, G. and Soejoeti, Z. (1989), A tampered failure rate model for step-stress accelerated life test. Communications in statistics- Theory and methods, vol. 18, no. 5, pp. 1627-1643.
- [2] DeGroot, M. H. and Goel, P. K. (1979), Bayesian estimation and optimal designs in partially accelerated life testing. Naval Research Logistics, 26(2), 223-235.
- [3] Hu, H., Lee, M. Y. and Tang, J. (2015), Optimum step-stress accelerated degradation test for Wiener degradation process under constraints, European Journal of Operational Research, 241, 412-421.
- [4] Nelson, W. (1980), Accelerated life testing-step-stress models and data analyses, IEEE transactions on reliability, 29(2), 103-108.

- [5] Sung, S. L. and Yum, B. J. (2016), Optimal design of step-stress accelerated degradation tests based on the Wiener degradation process. *Quality Technology and Quantitative Management*, 13(4), 367-393.
- [6] Tsai, C. C., Tseng, S. T. and Balakrishnan, N. (2012), Optimal Design for Degradation Tests Based on Gamma Processes With Random Effects, *IEEE Transactions on Reliability*, 61(2), 604-613.
- [7] Wang, X. and Xu, D. (2010), An inverse Gaussian process model for degradation data, *Technometrics*, vol. 52, no. 2, pp. 188-197.
- [8] Wang, H., Wang, G. J. and Duan, F. J. (2016), Planning of step-stress accelerated degradation test based on the inverse Gaussian process, *Reliability Engineering and System Safety*, 154, 97-105.
- [9] Yang, G. (2007), *Life Cycle Reliability Engineering*. Hoboken, NJ, USA: Wiley.
- [10] Ye, Z. S., Chen, L. P., Tang, L. C. and Xie, M. (2014), Accelerated degradation test planning using the inverse Gaussian process, *IEEE Trans Reliab*, 63(3), 750-63.
- [11] Zhang, C., Lu, X., Tan, Y. and Wang, Y. (2015), Reliability demonstration methodology for products with Gamma Process by optimal accelerated degradation testing, *Reliability Engineering and System Safety*, 142, 369-377.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## Cumulative residual inaccuracy for minimum and maximum of order statistics

Daneshi, S. <sup>1</sup>      Nezakati, A. <sup>2</sup>      Tahmasebi, S. <sup>3</sup>

<sup>1,2</sup> Faculty of Mathematics, Shahrood University of Technology

<sup>3</sup> Department of Statistics, Persian Gulf University, Bushehr

### Abstract

In this paper, we propose a measure of cumulative residual inaccuracy between survival function of the first-order statistic and parent survival function  $\bar{F}$ . We also consider cumulative inaccuracy measure between distribution of the last-order statistic and parent distribution  $F$ . We discuss some reliability properties of the proposed measures.

**Keywords:** Cumulative inaccuracy, Order statistics, Empirical approach.

## 1 Introduction

Let  $X$  denote the lifetime of a device or, a system with probability density function (pdf)  $f$  and cumulative distribution function (cdf)  $F$ , respectively. Then, the differential entropy known as Shannon entropy, is defined by [6] as follows:

$$H(X) = - \int_0^{+\infty} f(x) \log f(x) dx, \quad (1)$$

---

<sup>1</sup>sdaneshi445@gmail.com

<sup>2</sup>nezakati@shahroodut.ac.ir

<sup>3</sup>tahmasebi@pgu.ac.ir

where, by convention,  $0 \log 0 = 0$ . Recently, new measures of information are proposed in literatures: replacing the pdf by the survival function  $\bar{F} = 1 - F$  in Shannon entropy, the cumulative residual entropy(CRE) is defined by [5] as

$$\mathcal{E}(X) = \int_0^{+\infty} \bar{F}(x) \Lambda(x) dx,$$

where  $\Lambda(x) = -\log \bar{F}(x)$ . A new information measure similar to CRE has been proposed by [3] as follows:

$$\mathcal{CE}(X) = \int_0^{+\infty} F(x) \tilde{\Lambda}(x) dx, \quad (2)$$

where  $\tilde{\Lambda}(x) = -\log F(x)$ . Suppose that  $X$  and  $Y$  are two non-negative random variables with reliability functions  $\bar{F}(x)$ ,  $\bar{G}(x)$ , respectively. If  $\bar{F}(x)$  is the actual survival function corresponding to the observations and  $\bar{G}(x)$  is the survival function assigned by the experimenter, then [4] defined the cumulative residual inaccuracy based on  $\bar{F}(x)$  and  $\bar{G}(x)$  as follows:

$$I(\bar{F}, \bar{G}) = - \int_0^{+\infty} \bar{F}(x) \log \bar{G}(x) dx. \quad (3)$$

In analogy with (1), a measure of cumulative inaccuracy associated with  $F$  and  $G$  is given by

$$\tilde{I}(F, G) = - \int_0^{+\infty} F(x) \log G(x) dx. \quad (4)$$

Order statistics play an important role in problems such as industrial stress testing, meteorological analysis, hydrology, economics and other related fields. Different order statistics can be used in different applications; for example, the maximum is of interest in the study of floods and other meteorological phenomena while the minimum is often used in reliability and survival analysis, etc. For more details about order statistics and their applications, one may refer to [1]. Let  $X_1, X_2, \dots, X_n$  be a random sample of size  $n$  from an absolutely continuous cumulative distribution function  $F(x)$ . If  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$  represent the order statistics of the sample  $X_1, X_2, \dots, X_n$ . Then the empirical measure of  $F(x)$  is defined as

$$\hat{F}_n(x) = \begin{cases} 0, & x < X_{(1)}, \\ \frac{k}{n}, & X_{(k)} \leq x \leq X_{(k+1)}, \quad k = 1, 2, \dots, n-1 \\ 1, & x > X_{(n)}. \end{cases}$$

Recently [7] have introduced the measure of residual inaccuracy of order statistics and prove a characterization result for it. In this paper we consider a measure of cumulative residual inaccuracy between  $\bar{F}_{X_{(1:n)}}$  and  $\bar{F}$  and study its characterization results. We also propose the cumulative inaccuracy measure between  $F_{X_{(n:n)}}$  and  $F$  and obtain an estimator of cumulative inaccuracy using empirical approach .

## 2 Cumulative residual inaccuracy for min order statistics

In this section, we propose the cumulative residual measure of inaccuracy between  $\bar{F}_{X_{(1:n)}}$  and  $\bar{F}$  as follows:

$$I(\bar{F}_{X_{(1:n)}}, \bar{F}) = - \int_0^{+\infty} \bar{F}_{X_{(1:n)}}(x) \log(\bar{F}(x)) dx = \frac{1}{n} \mathcal{E}(X_{(1:n)}). \quad (5)$$

In analogy with (5), a measure of cumulative residual inaccuracy associated with  $\bar{F}$  and  $\bar{F}_{X_{(1:n)}}$  is given by

$$I(\bar{F}, \bar{F}_{X_{(1:n)}}) = - \int_0^{+\infty} \bar{F}(x) \log(\bar{F}_{X_{(1:n)}}(x)) dx = n \mathcal{E}(X). \quad (6)$$

Hereafter we present some properties of  $I(\bar{F}_{X_{(1:n)}}, \bar{F})$ .

**Proposition 2.1.** *Let  $X$  be an absolutely continuous nonnegative random variable with  $I(\bar{F}_{X_{(1:n)}}, \bar{F}) < \infty$ , for  $n \geq 1$ . Then, we have*

$$\begin{aligned} I(\bar{F}_{X_{(1:n)}}, \bar{F}) &= \int_0^{+\infty} [\bar{F}(x)]^n \left( \int_0^x \frac{f(z)}{\bar{F}(z)} dz \right) dx \\ &= \int_0^{+\infty} \int_z^{+\infty} \lambda(z) [\bar{F}(x)]^n dx dz, \end{aligned} \quad (7)$$

where  $\lambda(\cdot)$  is the failure rate function.

**Proposition 2.2.** *Let  $X$  be an absolutely continuous nonnegative random variable with  $I(\bar{F}_{X_{(1:n)}}, \bar{F}) < \infty$ , for  $n \geq 1$ . Then, we have*

$$I(\bar{F}_{X_{(1:n)}}, \bar{F}) = E \left( m_{X_{(1:n)}}(Z) (\bar{F}(Z))^{n-1} \right), \quad (8)$$

where

$$m_{X_{(1:n)}}(z) = E[X_{(1:n)} - z | X_{(1:n)} > z] = \frac{1}{(\bar{F}(z))^n} \int_z^{+\infty} (\bar{F}(x))^n$$

is the mean residual lifetime(mrl) of  $X_{(1:n)}$ .

*Proof.* From (3), we have

$$\begin{aligned} I(\bar{F}_{X_{(1:n)}}, \bar{F}) &= \int_0^{+\infty} \int_z^{+\infty} \lambda(z) [\bar{F}(x)]^n dx dz \\ &= \int_0^{+\infty} \frac{f(z)}{\bar{F}(z)} dz \left[ \int_z^{+\infty} [\bar{F}(x)]^n dx \right] \\ &= \int_0^{+\infty} \frac{f(z)}{\bar{F}(z)} [\bar{F}(z)]^n m_{X_{(1:n)}}(z) dz = \int_0^{+\infty} f(z) [\bar{F}(z)]^{n-1} m_{X_{(1:n)}}(z) dz. \end{aligned} \quad (9)$$

So, the proof is complete.  $\square$

**Proposition 2.3.** Let  $a, b > 0$ . For  $n = 1, 2, \dots$  it holds that

$$I(\bar{F}_{aX_{(1:n)}+b}, \bar{F}_{aX+b}) = aI(\bar{F}_{X_{(1:n)}}, \bar{F}).$$

The next propositions give some lower and upper bounds for  $I(\bar{F}_{X_{(1:n)}}, \bar{F})$ .

**Proposition 2.4.** For a nonnegative random variable  $X$  and  $n \geq 1$ , it holds that

$$I(\bar{F}_{X_{(1:n)}}, \bar{F}) \geq m_{X_{(1:n)}}(t) |\log \bar{F}(t)| [\bar{F}(t)]^2, \quad (10)$$

where  $m_{X_{(1:n)}}(t)$  is the mrl of  $X_{(1:n)}$ .

*Proof.* The proof follows from [2]. □

**Proposition 2.5.** Let  $X$  be an absolutely continuous nonnegative random variable with  $I(\bar{F}_{X_{(1:n)}}, \bar{F}) < \infty$ , for  $n \geq 1$ . Then, we have

$$I(\bar{F}_{X_{(1:n)}}, \bar{F}) \geq \int_0^{+\infty} (\bar{F}(x))^n F(x) dx. \quad (11)$$

*Proof.* Recalling that  $-\log \bar{F}(x) \geq F(x)$ , the proof then finally follows. □

**Proposition 2.6.** Let  $X$  be an absolutely continuous nonnegative random variable with  $I(\bar{F}_{X_{(1:n)}}, \bar{F}) < \infty$ , for  $n \geq 1$ . Then, we have

$$I(\bar{F}_{X_{(1:n)}}, \bar{F}) \leq \mathcal{E}(X). \quad (12)$$

*Proof.* Since  $\bar{F}(x) \geq [\bar{F}(x)]^n$ ,  $x \geq 0$ , when  $n \geq 1$ , the proof then finally follows. □

**Proposition 2.7.** If  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$  represent the order statistics of the sample  $X_1, X_2, \dots, X_n$ . Then, the empirical measure of  $I(\bar{F}_{X_{(1:n)}}, \bar{F})$  is obtained as

$$\begin{aligned} \hat{I}(\bar{F}_{X_{(1:n)}}, \bar{F}) &= - \int_0^{+\infty} [\hat{\bar{F}}_n(x)]^n \log \hat{\bar{F}}_n(x) dx \\ &= - \sum_{k=1}^{n-1} \int_{X_{(k)}}^{X_{(k+1)}} \left(1 - \frac{k}{n}\right)^n \log \left(1 - \frac{k}{n}\right) dx \\ &= - \sum_{k=1}^{n-1} U_k \left(1 - \frac{k}{n}\right)^n \log \left(1 - \frac{k}{n}\right), \end{aligned} \quad (13)$$

where  $U_k = X_{(k+1)} - X_{(k)}$ ,  $k = 1, 2, \dots, n-1$ .

### 3 Cumulative inaccuracy for max order statistics

We consider the cumulative measure of inaccuracy between  $F_{X_{(n:n)}}$  and  $F$  as follows:

$$\tilde{I}(F_{X_{(n:n)}}, F) = - \int_0^{+\infty} F_{X_{(n:n)}}(x) \log(F(x)) dx = \frac{1}{n} \mathcal{CE}(X_{(n:n)}). \quad (14)$$

In analogy with (2), a measure of cumulative inaccuracy associated with  $F$  and  $F_{X_{(n:n)}}$  is given by

$$\tilde{I}(F, F_{X_{(n:n)}}) = - \int_0^{+\infty} F(x) \log(F_{X_{(n:n)}}(x)) dx = n \mathcal{CE}(X). \quad (15)$$

Hereafter we consider some properties of  $\tilde{I}(F_{X_{(n:n)}}, F)$ .

**Proposition 3.1.** *Let  $X$  be an absolutely continuous nonnegative random variable with  $\tilde{I}(F_{X_{(n:n)}}, F) < \infty$ , for  $n \geq 1$ . Then, we have*

$$\begin{aligned} \tilde{I}(F_{X_{(n:n)}}, F) &= \int_0^{+\infty} [F(x)]^n \left( \int_x^\infty \frac{f(z)}{F(z)} dz \right) dx \\ &= \int_0^{+\infty} \int_0^z \tilde{\lambda}(z) [F(x)]^n dx dz, \end{aligned} \quad (16)$$

where  $\tilde{\lambda}(\cdot)$  is the reversed failure rate function.

**Proposition 3.2.** *Let  $X$  be an absolutely continuous nonnegative random variable with  $\tilde{I}(F_{X_{(n:n)}}, F) < \infty$ , for  $n \geq 1$ . Then, we have*

$$\tilde{I}(F_{X_{(n:n)}}, F) = E \left( \tilde{m}_{X_{(n:n)}}(Z) (F(Z))^{n-1} \right), \quad (17)$$

where

$$\tilde{m}_{X_{(n:n)}}(z) = E[z - X_{(n:n)} | z > X_{(n:n)}] = \frac{1}{(F(z))^n} \int_0^z (F(x))^n$$

is the mean inactivity time of  $X_{(n:n)}$ .

*Proof.* By (7), we obtain

$$\begin{aligned} \tilde{I}(F_{X_{(n:n)}}, F) &= \int_0^{+\infty} \int_0^z \tilde{\lambda}(z) [F(x)]^n dx dz \\ &= \int_0^{+\infty} \frac{f(z)}{F(z)} dz \left[ \int_0^z [F(x)]^n dx \right] \\ &= \int_0^{+\infty} \frac{f(z)}{F(z)} [F(z)]^n \tilde{m}_{X_{(n:n)}}(z) dz = \int_0^{+\infty} f(z) [F(z)]^{n-1} \tilde{m}_{X_{(n:n)}}(z) dz. \end{aligned} \quad (18)$$

Thus, the proof is complete.  $\square$

**Proposition 3.3.** Let  $a, b > 0$ . For  $n = 1, 2, \dots$  it holds that

$$\tilde{I}(F_{aX_{(n:n)}+b}, F_{aX+b}) = a\tilde{I}(F_{X_{(n:n)}}, F).$$

The next propositions give some lower and upper bounds for  $\tilde{I}(F_{X_{(n:n)}}, F)$ .

**Proposition 3.4.** For a nonnegative random variable  $X$  and  $n \geq 1$ , it holds that

$$\tilde{I}(F_{X_{(n:n)}}, F) \geq \tilde{m}_{X_{(n:n)}}(t)|\log F(t)|[F(t)]^2, \quad (19)$$

where  $\tilde{m}_{X_{(n:n)}}(t)$  is the mean inactivity time of  $X_{(n:n)}$ .

*Proof.* The proof follows from [2]. □

**Proposition 3.5.** Let  $X$  be an absolutely continuous nonnegative random variable with  $\tilde{I}(F_{X_{(n:n)}}, F) < \infty$ , for  $n \geq 1$ . Then, we have

$$I(F_{X_{(n:n)}}, F) \geq \int_0^{+\infty} (F(x))^n \bar{F}(x) dx. \quad (20)$$

*Proof.* Recalling that  $-\log F(x) \geq \bar{F}(x)$ , the proof then finally follows. □

**Proposition 3.6.** Let  $X$  be an absolutely continuous nonnegative random variable with  $\tilde{I}(F_{X_{(n:n)}}, F) < \infty$ , for  $n \geq 1$ . Then, we have

$$\tilde{I}(F_{X_{(n:n)}}, F) \leq \mathcal{CE}(X). \quad (21)$$

*Proof.* Since  $F(x) \geq [F(x)]^n$ ,  $x \geq 0$ , when  $n \geq 1$ , the proof then finally follows. □

**Proposition 3.7.** Let  $X$  be a symmetric random variable with respect to the finite mean  $\mu = E(X)$ , i.e.  $F(x + \mu) = 1 - F(\mu - x)$  for all  $x \in \mathbb{R}$ . Then

$$\tilde{I}(F_{X_{(n:n)}}, F) = I(\bar{F}_{X_{(1:n)}}, \bar{F}).$$

**Proposition 3.8.** If  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$  denote the order statistics of the sample  $X_1, X_2, \dots, X_n$ . Then, the empirical measure of  $\tilde{I}(F_{X_{(n:n)}}, F)$  is given by

$$\begin{aligned} \hat{\tilde{I}}(F_{X_{(n:n)}}, F) &= - \int_0^{+\infty} [\hat{F}_n(x)]^n \log \hat{F}_n(x) dx \\ &= - \sum_{k=1}^{n-1} \int_{X_{(k)}}^{X_{(k+1)}} \left(\frac{k}{n}\right)^n \log \left(\frac{k}{n}\right) dx \\ &= \frac{-1}{n^n} \sum_{k=1}^{n-1} k^n U_k \log \left(\frac{k}{n}\right), \end{aligned} \quad (22)$$

where  $U_k = X_{(k+1)} - X_{(k)}$ ,  $k = 1, 2, \dots, n-1$ .

In the following example, we calculate  $\tilde{I}(F_{X_{(n:n)}}, F)$  and  $I(\bar{F}_{X_{(1:n)}}, \bar{F})$  for some specific lifetime distributions which are widely used in reliability theory and life testing.

**Example 3.1.** (a) If  $X$  is uniformly distributed in  $[0, \theta]$ , then it is easy to see that  $\tilde{I}(F_{X_{(n:n)}}, F) = I(\bar{F}_{X_{(1:n)}}, \bar{F}) = \frac{\theta}{(n+1)^2}$ , for all integers  $n \geq 1$ . Note that  $\tilde{I}(F_{X_{(n:n)}}, F)$  is a decreasing function of  $n$ .

(b) If  $X$  has a Weibull distribution with survival function  $\bar{F}(x) = e^{-\lambda^q x^q}$ ,  $x > 0$ ,  $\lambda, q > 0$ , then for all integers  $n \geq 1$ , we obtain  $I(\bar{F}_{X_{(1:n)}}, \bar{F}) = \frac{1}{\lambda q^2 n^{\frac{q+1}{q}}} \Gamma(\frac{1}{q})$ .

(c) If  $X$  has a Pareto distribution with pdf  $f(x) = \frac{\alpha \beta^\alpha}{x^{\alpha+1}}$ ,  $x \geq \beta$ ,  $\beta > 0$ ,  $\alpha > 0$ , then  $I(\bar{F}_{X_{(1:n)}}, \bar{F}) = \frac{\alpha \beta}{(n \alpha - 1)^2}$ , for all integers  $n > \frac{1}{\alpha}$ . Note that  $I(\bar{F}_{X_{(1:n)}}, \bar{F})$  is a decreasing function of  $n$  for all  $\alpha > \frac{1}{n}$ .

(d) Let  $X$  be an exponential distribution with mean  $\frac{1}{\lambda}$ , then  $\tilde{I}(F_{X_{(n:n)}}, F) = \frac{1}{\lambda} [\frac{\pi^2}{6} - \sum_{j=1}^n \frac{n}{j^2}]$  and  $I(\bar{F}_{X_{(1:n)}}, \bar{F}) = \frac{1}{n^2 \lambda}$ . Note that  $I(\bar{F}_{X_{(1:n)}}, \bar{F})$  is a decreasing function of  $n$ .

(e) Let  $X$  be a nonnegative random variable which has an inverse Weibull distribution with the cdf  $F(x) = \exp(-(\frac{x}{\alpha})^\beta)$ ,  $x > 0$ , then for all integers  $n \geq 1$ , we obtain  $\tilde{I}(F_{X_{(n:n)}}, F) = \frac{\alpha(n)}{\beta} \frac{1-\beta}{\beta} \Gamma(\frac{\beta-1}{\beta})$ .

## References

- [1] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N.(1992), *A first course in order statistics*, John Wiley, New York.
- [2] Baratpour, S.(2010), *Characterizations based on cumulative residual entropy of first-order statistics*. *Communications in Statistics-Theory and Methods*, **39**, 3645-3651.
- [3] Di Crescenzo, A. and Longobardi, M.( 2009), On cumulative entropies. *Journal of Statistical Planning and Inference*, **139**, 4072-4087.
- [4] Kumar, V. and Taneja, H.C.(2015), Dynamic cumulative residual and past inaccuracy measures. *J. Stat. Theory Appl.*, **14**(4), 399-412.
- [5] Rao, M., Chen, Y., Vemuri, B.C. and Wang, F.(2004), Cumulative Residual Entropy: A New Measure of Information. *IEEE Trans Inf Theory*, **50**(6), 1220-1228.
- [6] Shannon, C. E.(1948), A mathematical theory of communication. *Bell Syst. Tec. J.*, **37**, 379-432.
- [7] Thapliyal, R. and Taneja, H.C.(2015), On residual inaccuracy of order statistics. *Statist. Probab. Lett.*, **97**, 125-131.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## P ( $Y < X$ ) using generalized order statistics and concomitant

Domma, F.<sup>1</sup>      Eftekharian, A.<sup>2</sup>      Razmkhah, M.<sup>3</sup>

<sup>1</sup> Department of Economics, Statistics and Finance, University of Calabria, Via Pietro Bucci,  
Cubo 0C, 87036 Arcavacata di Rende (CS)-Italy

<sup>2</sup> Department of Statistics, Faculty of Science, University of Hormozgan, Bandar Abbas

<sup>3</sup> Department of Statistics, Ferdowsi University of Mashhad

### Abstract

The stress-strength model is proposed based on the m-generalized order statistics and corresponding concomitant. For dependency between m-generalized order statistics and its concomitant, a bivariate copula expansion is considered and the stress-strength model is obtained for two special cases of order statistics and upper record values. Basis on the order statistics and record values, two estimators of stress-strength are presented using a similar procedure to the inference functions for margins.

**Keywords:** Copula function, Order statistics, stress-strength.

## 1 Introduction

In the stress-strength reliability, one usually considers two random variables in such that one of them is denoted as stress ( $X$ ) and another is called strength ( $Y$ ) of a component or a system. If the stress exceeds the strength the component fails, while the component works whenever the stress does not exceed the strength. The reliability is defined as the probability that the component works, i.e.  $R = P(Y > X)$ .

---

<sup>1</sup>filippo.domma@unical.it

<sup>2</sup>eftekharian@hormozgan.ac.ir

A much smaller number of papers addressed the stress-strength problem when the random variables are dependent. The evaluation and the estimation of  $R$  were discussed in the literature when  $(X, Y)$  follows the bivariate normal distribution by Gupta and Subramanian [10], the bivariate Pareto distribution by Hanagal [11] and Jeevanand [12], the bivariate beta distribution by Nadarajah [14], the bivariate lognormal distribution by Gupta and Gupta [9] and the bivariate Rayleigh distribution by Pak et al. [15]. Recently, the stress-strength model in the case of dependence using the copula approach was considered by Domma and Giordano [5, 6].

The special attention shown by many authors in regard to the stress-strength model is due to the wide applicability in various fields of science. In fact, introduced in the reliability context, it has been adapted and applied in engineering, medicine, economics, biology and psychology. For example, in a clinical study,  $P(Y > X)$  measures the effectiveness of the treatment when  $X$  and  $Y$  are the responses of a control group and treatment group, respectively. Other known applications in this context, concern the evaluation of the area under the ROC curve for diagnostic tests with continuous outcomes [1]. In economics, it has been used to evaluate the distance between the income distributions [4] and, more recently, as a measure of household of financial fragility which occurs whenever expenses exceed a household yearly income for dependence case [5].

Some authors have placed their attention on concomitants of generalized order statistics (GOSs) proposed by Kamps [13] as a unified approach containing several models of ordered random variables as special cases, e.g. order statistics, the  $k$ th and Pfeifers record values, progressively type-II censored order statistics, among others. Beg and Ahsanullah [3] have examined concomitants of GOSs from Farlie-Gumbel-Morgenstern (*FGM*) distribution, Tahmasebi et al. [16] have studied the concomitants of Dual-GOSs from Morgenstern type bivariate generalized exponential distribution, Domma and Giordano [7] have considered the concomitants of m-GOSs from Generalized Farlie-Gumbel-Morgenstern (*GFGM*) distribution. It should be noted that since between ordered data and the corresponding concomitant exists a dependence structure, it needs to use a dependence structure such as copula function.

In this paper, we introduce the above problem from a general viewpoint of studying the stress-strength model in the case of the concomitant based on the GOSs. In addition, we use the *PRHF* as marginal distributions and the copula function to model the dependence structure. Relevant special cases of concomitant from order statistics and record values are studied in depth by specifying a particular copula function and selecting a particular member of the *PRHF* as marginal distributions.

## 2 Concomitant of $m$ -GOSs and stress-strength model

The GOSs was first proposed by Kamps [13] as a unified approach containing several models of ordered random variables as special cases, e.g. order statistics, the  $k$ th and Pfeifers record values, progressively type-II censored order statistics, among others.

Suppose that  $F$  is an absolutely continuous cumulative distribution function (*cdf*) with survival function  $\bar{F} = 1 - F$ , and *pdf*  $f$ . Assume that  $n \in \mathbb{N}$ ,  $\tilde{m} = (m_1, \dots, m_{n-1}) \in \mathbb{R}^{n-1}$ , and  $k > 0$ , are parameters such that  $\gamma_i = k + n - i + M_i > 0$ , for  $i = 1, 2, \dots, n-1$ , where

$M_i = \sum_{j=i}^{n-1} m_j$ . We denote the random variables of GOSs by  $X_{(r,n,\tilde{m},k)}$ ,  $r = 1, \dots, n$ , and their *joint pdf* is given by

$$f_{X_{(1,n,\tilde{m},k)}, \dots, X_{(n,n,\tilde{m},k)}}(x_1, \dots, x_n) = k \prod_{i=1}^n \gamma_i \left( \prod_{j=1}^{n-1} \bar{F}^{m_j}(x_j) f(x_j) \right) \bar{F}^{k-1}(x_n) f(x_n), \quad (1)$$

on the cone  $F^{-1}(0) < x_1 \leq \dots \leq x_n < F^{-1}(1)$ , where  $F^{-1}(\cdot)$  is the inverse function of  $u = F(x)$ , i.e. quantile function. In the particular case, when  $m_1 = m_2 = \dots = m_{n-1} = m$ , and  $\gamma_i = k + (n-i)(m+1)$ , for  $i = 1, \dots, n-1$ , the random variable  $X_{(r,n,\tilde{m},k)}$ , is called  $m$ -GOSs and is denoted by  $X_{(r,n,m,k)}$ ,  $r = 1, \dots, n$ . Using (1) the marginal *pdf* of the  $r$ th  $m$ -GOSs is as follows

$$f_{(r,n,m,k)}(x) = \frac{\prod_{i=1}^r \gamma_i}{(r-1)!} \bar{F}^{\gamma_r-1}(x) f(x) t_m^{r-1}(F(x)), \quad (2)$$

where

$$t_m(F(x)) = \begin{cases} \frac{1}{m+1} [1 - \bar{F}^{m+1}(x)] & m \neq -1, \\ -\log \bar{F}(x) & m = -1, F(x) \in [0, 1]. \end{cases}$$

Special cases of  $m$ -GOSs can be obtained with appropriate choices of the parameters  $m$  and  $k$ . For example, it is easy to verify that for  $m = 0$  and  $k = 1$ , the  $m$ -GOSs becomes the  $r$ th order statistic, whereas for  $m = 1$  and  $k = 1$ ,  $X_{(r,n,m,k)}$  becomes the  $r$ th upper record.

In order to obtain the stress-strength based on the  $n$ th  $m$ -GOSs and its concomitant, it is necessary to derive the *joint pdf* between  $X_{(n,n,m,k)}$  and  $Y_{[n,n,m,k]}$ . It is possible to deduce that the *joint pdf* is given by

$$h_{X_{(n,n,m,k)}, Y_{[n,n,m,k]}}(x, y) = h_{X,Y}(x, y) \frac{\prod_{i=1}^n \gamma_i}{(n-1)!} \bar{F}^{\gamma_n-1}(x) t_m^{n-1}(F(x)). \quad (3)$$

Now, using (3) we can determine the probability that  $n$ th  $m$ -GOS,  $X_{(n,n,m,k)}$ , is smaller than its concomitant,  $Y_{[n,n,m,k]}$ , *i.e.*

$$\begin{aligned} R := P(Y_{[n,n,m,k]} > X_{(n,n,m,k)}) &= \int \int_{y>x} g(y|x) f(x) \frac{\prod_{i=1}^n \gamma_i}{(n-1)!} \bar{F}^{\gamma_n-1}(x) t_m^{n-1}(F(x)) dy dx \\ &= \int_{-\infty}^{\infty} \frac{\prod_{i=1}^n \gamma_i}{(n-1)!} \bar{F}^{\gamma_n-1}(x) t_m^{n-1}(F(x)) f(x) \int_x^{\infty} g(y|x) dy dx \\ &= E_F \left[ \frac{\prod_{i=1}^n \gamma_i}{(n-1)!} \bar{F}^{\gamma_n-1}(X) t_m^{n-1}(F(X)) \{1 - G(X|X)\} \right], \end{aligned} \quad (4)$$

where  $G(\cdot|\cdot)$  is the conditional *cdf* of  $Y$  given  $X$  and  $E_F(\cdot)$  is the expectation with respect to  $F(x)$ .

Let  $X$  and  $Y$  be two continuous random variables with joint distribution function  $H_{X,Y}(x, y)$  and marginal distribution functions  $F(x)$  and  $G(y)$ , respectively. Sklar's theorem states that any bivariate distribution function can be written as  $H_{X,Y}(x, y) =$

$C(F(x), G(y))$ , where  $C(\cdot, \cdot)$  is a unique copula function. It is easy to verify that the *joint pdf* is given by  $h_{X,Y}(x, y) = c(F(x), G(y))f(x)g(y)$  where  $c(F(x), G(y)) = \frac{\partial^2 C(F(x), G(y))}{\partial F(x)\partial G(y)}$  is a copula density.

Using the copula function, we can write the stress-strength of  $m$ -GOSs and its concomitant as

$$R = \int \int_{y>x} c(F(x), G(y))g(y)f(x) \frac{\prod_{i=1}^n \gamma_i}{(n-1)!} \bar{F}^{\gamma_n-1}(x) t_m^{n-1}(F(x)) dy dx. \quad (5)$$

The latter equation shows that using the copula function, we can consider various cases for  $X$  and  $Y$ . For example,  $X$  and  $Y$  are either independent,  $C(F(x), G(y)) = F(x)G(y)$ , or dependent random variables. Also, we can assume that they have either the same distribution,  $X \stackrel{d}{=} Y$ , or different distributions.

The *GFGM* distribution family introduced by Farlie [8] is the most general form of the *FGM* family, defined by

$$H(x, y) = F(x)G(y) \{1 + \theta A(F(x)) B(G(y))\}, \quad (6)$$

where  $A(\cdot)$  and  $B(\cdot)$  are differentiable functions on the unit interval and  $A(t) \rightarrow 0$  and  $B(t) \rightarrow 0$  as  $t \rightarrow 1$  and  $\theta$  is the dependence parameter. In special case, when  $\theta = 0$ , then  $X$  and  $Y$  are independent. The bivariate *pdf* of the *GFGM* is denoted by

$$h(x, y) = \{1 + \theta [A(F(x)) + F(x)A'(F(x))] [B(G(y)) + G(y)B'(G(y))]\} F(x)G(y),$$

where  $A'(F(x)) = \frac{\partial A(F(x))}{\partial F(x)}$  and  $B'(G(y)) = \frac{\partial B(G(y))}{\partial G(y)}$ . Evidently, the *FGM* distribution is a special case of *GFGM* distribution with dependence parameter  $\theta$ , when  $A(F(x)) = 1 - F(x)$  and  $B(G(y)) = 1 - G(y)$ , and the *cdf* of *FGM* is as follows

$$H(x, y) = F(x)G(y) \{1 + \theta (1 - F(x))(1 - G(y))\},$$

with the corresponding bivariate *pdf*

$$h(x, y) = f(x)g(y) \{1 + \theta (2F(x) - 1)(2G(y) - 1)\}. \quad (7)$$

If we consider  $A(F(x)) = [1 - F^{m_1}(x)]^{p_1}$  and  $B(G(y)) = [1 - G^{m_2}(y)]^{p_2}$  in equation (6), then we can obtain the following generalization of the *FGM* distribution as

$$H(x, y) = F(x)G(y) \{1 + \theta [1 - F^{m_1}(x)]^{p_1} [1 - G^{m_2}(y)]^{p_2}\}, \quad (8)$$

with  $m_1, m_2, p_1, p_2 > 0$  (see, [2]).

It can be shown that the stress-strength model of  $m$ -GOS and its concomitant in this case is

$$\begin{aligned} R^{(GFGM)} &= P(Y_{[n,n,m,k]} > X_{(n,n,m,k)}) = \frac{\prod_{i=1}^n \gamma_i}{(n-1)!} E_F [\bar{F}^{\gamma_n-1}(X) t_m^{n-1}(F(X)) \{1 - G(X)\}] \\ &\quad + \theta \sum_{r=0}^{p_1-1} \sum_{t=0}^{p_2-1} \xi_{p_1-1,r} \xi_{p_2-1,t} \times \frac{\prod_{i=1}^n \gamma_i}{(n-1)!} E_F \left[ \bar{F}^{\gamma_n-1}(X) t_m^{n-1}(F(X)) F^{m_1 r}(X) \right. \\ &\quad \left. \times \{1 - \eta_1 F^{m_1}(X)\} \{\eta_{2,t} (1 - G^{m_2 t+1}(X)) - \eta_{3,t} (1 - G^{m_2(t+1)+1}(X))\} \right], \end{aligned} \quad (9)$$

where  $\eta_1 = 1 + m_1 p_1$ ,  $\eta_{2,t} = \frac{1}{m_2 t + 1}$  and  $\eta_{3,t} = \frac{1+m_2 p_2}{m_2(t+1)+1}$ . As was done in the previous section, to use the equation (9), we have to specify the function  $t(\cdot)$  and the marginal distribution functions  $F(\cdot)$  and  $G(\cdot)$ .

Now, we can consider the stress-strength model of order statistics and upper record values, special cases of  $m$ -GOS, using *GFGM* bivariate distribution as folloing

## 1. Order statistics

$$R_1^{(GFGM)} = nE_F [F^{n-1}(X)(1-G(X))] + n\theta \sum_{r=0}^{p_1-1} \sum_{t=0}^{p_2-1} \xi_{p_1-1,r} \xi_{p_2-1,t} \\ \times E_F \left[ F^{n+m_1 r-1}(X) \{1 - \eta_1 F^{m_1}(X)\} \left\{ \begin{aligned} & \eta_{2,t} (1 - G^{m_2 t+1}(X)) \\ & - \eta_{3,t} (1 - G^{m_2(t+1)+1}(X)) \end{aligned} \right\} \right]. \quad (10)$$

## 2. Upper record

$$R_2^{(GFGM)} = 1 - \frac{1}{(n-1)!} \sum_{i=0}^{\infty} c_i \int_0^{\infty} (F(x))^{n+i-1} G(x) f(x) dx + \theta \sum_{r=0}^{p_1-1} \sum_{t=0}^{p_2-1} \xi_{p_1-1,r} \xi_{p_2-1,t} \\ \times \left\{ \begin{aligned} & \frac{(\eta_{2,t} - \eta_{3,t})}{(n-1)!} \sum_{i=0}^{\infty} \frac{c_i}{m_1 r + n + i} + \frac{\eta_1 (\eta_{3,t} - \eta_{2,t})}{(n-1)!} \sum_{i=0}^{\infty} \frac{c_i}{m_1(r+1) + n + i} \\ & - \frac{1}{(n-1)!} \sum_{i=0}^{\infty} c_i \int_0^{\infty} \left[ \eta_{2,t} (F(x))^{m_1 r + n + i - 1} (G(x))^{m_2 t + 1} - \eta_{3,t} (F(x))^{m_1 r + n + i - 1} \right. \\ & \times (G(x))^{m_2(t+1)+1} - \eta_1 \eta_{2,t} (F(x))^{m_1(r+1)+n+i-1} (G(x))^{m_2 t + 1} \\ & \left. + \eta_1 \eta_{3,t} (F(x))^{m_1(r+1)+n+i-1} (G(x))^{m_2(t+1)+1} \right] f(x) dx \end{aligned} \right\}. \quad (11)$$

$R_1^{(GFGM)}$  and  $R_2^{(GFGM)}$  can be estimated using some methods such as MLE.

## 3 Simulation study

The simulation study helps us to evaluate the performance of introduced estimators of the stress-strength model for order statistics and record values. Toward this end, we apply two criteria mean squared error (MSE) and average of the relative estimates (ARE) for evaluating the performance of  $\hat{\theta}$ ,  $\hat{R}_1^{(GFGM)}$  and  $\hat{R}_2^{(GFGM)}$  in which they are defined on the basis of the  $B$  iterations as follows

$$MSE_n(\hat{T}) = \sum_{j=1}^B (\hat{T}_j - T_0)^2, \quad ARE_n(\hat{T}) = \sum_{j=1}^B \frac{\hat{T}_j}{BT_0},$$

Table 1: The ARE and MSE of  $\hat{R}_1^{(GFGM)}$ ,  $\hat{R}_2^{(GFGM)}$  and  $\hat{\theta}$  for different combinations of  $(m_1, m_2, p_1, p_2)$ .

$m_1$	$m_2$	$p_1$	$p_2$	$\theta_0$	$n$	$ARE(\hat{R}_1^{(GFGM)})$	$MSE(\hat{R}_1^{(GFGM)})$	$ARE(\hat{R}_2^{(GFGM)})$	$MSE(\hat{R}_2^{(GFGM)})$	$ARE(\hat{\theta})$	$MSE(\hat{\theta})$	
3	2	3	2	-.855	5	1.17581	0.02284	1.05045	0.00094	1.00474	2.50868	
					6	1.15644	0.01421	1.01267	0.00017	1.01030	1.87495	
					7	1.15710	0.00943	1.00374	3.59283e-05	0.99856	1.54660	
					8	1.13096	0.00575	1.00115	7.52363e-06	0.99860	1.23249	
					9	1.11392	0.00387	1.00038	1.42016e-06	1.00665	1.04358	
					10	1.10856	0.00268	1.00014	2.85406e-07	0.99039	0.91220	
					0.425	5	1.08260	0.02318	1.03635	0.00075	0.99100	2.73333
					6	1.08083	0.01478	1.01016	0.00013	0.99388	2.08866	
					7	1.08276	0.00968	1.00285	2.57740e-05	1.01174	1.66481	
					8	1.08038	0.00638	1.00084	4.90315e-06	1.03810	1.35129	
5	5	4	4	-0.217	9	1.07746	0.00453	1.00028	9.33792e-07	1.00408	1.17564	
					10	1.06709	0.00316	1.00009	1.76982e-07	0.98681	1.01576	
					5	1.13623	0.03132	1.04259	0.00107	0.96907	0.78532	
					6	1.11247	0.01954	1.01082	0.00016	1.02122	0.59966	
					7	1.09818	0.01300	1.00307	2.98309e-05	1.05114	0.47672	
					8	1.10185	0.00844	1.00094	5.84757e-06	0.97144	0.39710	
					9	1.11311	0.00634	1.00037	1.29379e-06	0.97798	0.33637	
					10	1.08576	0.00403	1.00012	2.30822e-07	0.97487	0.29640	
					0.415	5	1.08693	0.03429	1.04089	0.00102	0.99240	0.76323
					6	1.07540	0.02220	1.01048	0.00015	1.00555	0.57971	
3	4	5	4	-0.634	7	1.11148	0.01293	1.00319	3.12329e-05	1.00234	0.46797	
					8	1.07948	0.01051	1.00076	4.58857e-06	1.02411	0.37184	
					9	1.07465	0.00734	1.00027	9.00991e-07	0.98244	0.32326	
					10	1.06330	0.00504	1.00008	1.55180e-07	0.99077	0.27938	
					5	1.15737	0.02130	1.05024	0.00081	0.99292	3.17357	
					6	1.15911	0.01243	1.01401	0.00016	0.96648	2.35162	
					7	1.12484	0.00705	1.00351	3.17829e-05	0.98135	1.84516	
					8	1.12564	0.00471	1.00120	7.46519e-06	0.97523	1.51453	
					9	1.10579	0.00305	1.00037	1.37963e-06	0.94758	1.29789	
					10	1.08691	0.00204	1.00013	2.51885e-07	0.99947	1.09205	
0.473	5	4	4	-0.634	5	1.08505	0.02251	1.03610	0.00057	0.99494	3.27339	
					6	1.08922	0.01306	1.00979	0.00010	1.02428	2.47461	
					7	1.07828	0.00807	1.00251	2.18929e-05	0.99739	2.00666	
					8	1.07132	0.00503	1.00077	4.56301e-06	0.99401	1.64287	
					9	1.07272	0.00335	1.00025	8.60785e-07	1.00408	1.42765	
					10	1.06833	0.00231	1.00009	1.66416e-07	0.97998	1.20997	

where  $\hat{T}_j$  and  $T_0$  are the value of the estimator in the  $j$ th iteration and fixed value of the parameter, respectively.

Based on these stages, we are performed the simulation for some combinations of  $(m_1, m_2, p_1, p_2)$  with  $B = 10000$  iterations and by increasing the sample size, i.e.  $n = 5, 6, 7, 8, 9, 10$ . Since the range of  $\theta$  is a function of  $(m_1, m_2, p_1, p_2)$ , so we are considered two arbitrary different values of  $\theta$ , both positive and negative, for each combination of  $(m_1, m_2, p_1, p_2)$ . Also, we have used Dagum marginal distributions  $Da(1.5, 1, 3)$  and  $Da(0.5, 1, 3)$  to generate the data according to the aforementioned procedure. It should be noted that the obtained results do not change remarkably in comparison with other choices of the Dagum parameters. The ARE criterion indicates the behavior of each estimator with respect to chosen fixed value of that parameter.

Table 1 contains the simulation results of MSE and ARE for  $\hat{R}_1^{(GFGM)}$ ,  $\hat{R}_2^{(GFGM)}$  and  $\hat{\theta}$ . There exist 3 different combinations of  $(m_1, m_2, p_1, p_2)$  in Table 1 and it is easy to provide quite similar outcomes for other combinations.

On the basis of the obtained results in Table 1, the estimators  $\hat{R}_1^{(GFGM)}$  and  $\hat{R}_2^{(GFGM)}$  carry out quite well. Their estimated values are really near to the chosen fixed ones of them because the ARE of two estimators are very close to 1 and it is desirable. As expected, the bias is decreasing when the sample size increases and the MSE tends decreasingly to zero as the sample size rises. This trend be able to observe for two estimators  $\hat{R}_1^{(GFGM)}$  and  $\hat{R}_2^{(GFGM)}$  and also  $\hat{\theta}$  in all cases such as different combinations of  $(m_1, m_2, p_1, p_2)$  and various values of  $\theta_0$ . On the other hand, the values of AREs and MSEs for estimators are

not remarkably sensitive with respect to changes of  $(m_1, m_2, p_1, p_2)$ , because by changing the combination  $(m_1, m_2, p_1, p_2)$ , the values of AREs and MSEs have not really changed.

By comparison the estimators  $\hat{R}_1^{(GFGM)}$  and  $\hat{R}_2^{(GFGM)}$  it can be found that both estimators decrease with respect to  $n$ , but the MSE of  $\hat{R}_2^{(GFGM)}$  is always less than that of  $\hat{R}_1^{(GFGM)}$  in each row of Table 1. Also, the ARE of  $\hat{R}_2^{(GFGM)}$  is nearer than to 1 than the ARE of  $\hat{R}_1^{(GFGM)}$ . These evidences show that the estimator  $\hat{R}_2^{(GFGM)}$  has better performance than estimator  $\hat{R}_1^{(GFGM)}$  to estimate the stress-strength parameter. In another word, based on the results of Table 1, record data with its concomitant could be better estimated the stress-strength parameter than order data and its concomitant.

## References

- [1] Gianfranco Adimari and Monica Chiogna, Partially parametric interval estimation of  $P\{Y > X\}$ , Computational statistics data analysis 51 (2006), no. 3, 1875-1891.
- [2] I Bairamov, S Kotz, and M Bekci, New generalized farlie-gumbel-morgenstern distributions and concomitants of order statistics, Journal of Applied Statistics 28 (2001), no. 5, 521-536.
- [3] MI Beg and M Ahsanullah, Concomitants of generalized order statistics from farlie-gumbel-morgenstern distributions, Statistical Methodology 5 (2008), no. 1, 1-20.
- [4] Camilo Dagum, The generation and distribution of income, the lorentz curve and the gini ratio, Economie Appliquee 33 (1980), no. 2, 327-367.
- [5] Filippo Domma and Sabrina Giordano, A stress-strength model with dependent variables to measure household financial fragility, Statistical Methods Applications 21 (2012), no. 3, 375-389.
- [6] , A copula-based approach to account for dependence in stress-strength models, Statistical Papers 54 (2013), no. 3, 807-826.
- [7] , Concomitants of m-generalized order statistics from generalized farlie-gumbel-morgenstern distribution family, Journal of Computational and Applied Mathematics 294 (2016), 413-435.
- [8] Dennis JG Farlie, The performance of some correlation coefficients for a general bivariate distribution, Biometrika 47 (1960), no. 3/4, 307-323.
- [9] Pushpa L Gupta and Ramesh C Gupta, Some properties of the bivariate lognormal distribution for reliability applications, Applied Stochastic Models in Business and Industry 28 (2012), no. 6, 598-606.
- [10] Ramesh C. Gupta and Sundarraman Subramanian, Estimation of reliability in a bivariate normal distribution with equal coefficients of variation, Communications in Statistics - Simulation and Computation 27 (1998), no. 3, 675-698.

- [11] David D. Hanagal, Estimation of reliability when stress is censored at strength, *Communications in Statistics - Theory and Methods* 26 (1997), no. 4, 911-919.
- [12] E. S. Jeevanand, Bayes estimation of  $P\{X_2 > X_1\}$  for a bivariate pareto distribution, *Journal of the Royal Statistical Society: Series D (The Statistician)* 46 (1997), no. 1, 93-99.
- [13] Udo Kamps, A concept of generalized order statistics, *Journal of Statistical Planning and Inference* 48 (1995), no. 1, 1-23.
- [14] Saralees Nadarajah, Reliability for some bivariate beta distributions, *Mathematical Problems in Engineering* 2005 (2005), no. 1, 101-111.
- [15] Abbas Pak, Nayereh Bagheri Khoolenjani, and Ali Akbar Jafari, Inference on  $P\{Y > X\}$  in bivariate rayleigh distribution, *Communications in Statistics-Theory and Methods* 43 (2014), no. 22, 4881-4892.
- [16] Saeid Tahmasebi, Ali Akbar Jafari, Mahmoud Afshari, Amit K Baral, MZ Anis, JA Jimenez, V Arunachalam, GM Serna, AKPC Swain, Mosisa Aga, et al., Concomitants of dual generalized order statistics from morgenstern type bivariate generalized exponential distribution, *Journal of Statistical Theory and Applications* 14 (2015), no. 1, 1-12.



# An imputation approach to Cox proportional hazards model in the presence of interval-censored data

Fani, S.<sup>1</sup>

Department of Statistics, Faculty of Mathematical Sciences and Computer, Allameh Tabataba'i University, Tehran

## Abstract

Due to the complexity of analyzing interval-censored data, the imputation approaches can be employed in order to reduce this problem to that of analyzing right-censored data. Then, classical nonparametric methods in the case of right-censored data can be used for imputed interval-censored observations. Four different imputation approaches are studied in this paper for computing the regression coefficients within the Cox proportional hazards model as one of the most popular methods of reliability analysis in the presence of covariates. Based on the proposed approach, the regression coefficients in the Cox model can be estimated by the traditional partial likelihood method. A numerical study is conducted to perform the proposed methods using a real data set.

**Keywords:** Cox Proportional Hazards Model, Interval-censored Data, Imputation.

## 1 Introduction

The primary concern in reliability analysis is to find the distribution of the event times of interest. Many statistical approaches study the direct estimation of the density or the reliability function. Parametric survival models, such as exponential, Weibull, and Gompertz distributions, are commonly used to model lifetimes. For reviews, see, e.g., [9],

<sup>1</sup>s.fani@atu.ac.ir

[10], [14], [17], and [13]. Although the advantages of parametric models include the ease in their computation, interpretation and prediction, the main limitation of parametric methods is the necessary model assumptions.

On the other hand, nonparametric methods require fewer restrictive assumptions of the models and are more flexible and robust to model misidentification. The most popular nonparametric technique for modelling survival data is the Kaplan-Meier (KM) or product limit estimator for a survival function; see [11]. One can derive an alternative of the estimator for the survival function based on the Nelson-Aalen (NA) estimator for the cumulative hazard rate function; see [15, 1].

In failure time data analysis, the subjects under study might have some additional characteristics that can effect the failure times. Therefore, one is often tempted to analyze how different features of subjects such as age, gender, smoking history, physical activity level, heart rate, or treatment indicator influence the distribution of the event of interest. These features are generally referred to as covariates, which are time-independent and usually recorded at the start of the study. The most renowned method of investigating the effects of covariates on lifetime distribution is the Cox proportional hazards model; see [5]. Its appeal resides mainly in its framework that enables us to efficiently estimate the regression coefficients; whereas, the baseline distribution is completely unknown. The proposed estimator was shown by [6] to be a maximum partial likelihood, and its asymptotic properties were extensively studied by [19]. [3] also proposed a different approach that yields the same maximum partial likelihood estimator along with an estimator of the baseline cumulative hazard function.

Obviously, survival times can be observed exactly. However, a major challenge in reliability analysis is to deal with a peculiar characteristic of time-to-event data, which is the presence of censoring, such as right censoring, left censoring, and interval censoring. By censoring, it implies that the information about an observation on a survival time of interest is incompletely known due to the temporal limits of the observation interval or cost. The more challenging censoring mechanism that predominantly takes place in medical, clinical trials and the longitudinal studies that entail periodic follow-ups is interval censoring. Interval-censored data imply that the failure time can not be directly observed, but is only known to have occurred within a time window or interval, e.g., two adjacent examination times in a sequence of clinical visits.

For the case of right-censored data, the crucial feature of the partial likelihood method is that one can yield the estimates of the regression coefficients without involving the estimation of the underlying baseline function; see e.g., [10] and [12]. In contrast, the estimates of the regression coefficients and the derivation of its asymptotic properties for the case of interval-censored data are a more challenging issue to be investigated due to the fact that we are not able to remove the baseline hazard function. Regression analysis for interval-censored data was first studied by [7], who considered a parametric method for approximating the baseline hazard distribution together with regression coefficients by maximizing the full likelihood of the observed data. Others who discussed the proportional hazards model in this case include [8], [16], [2], and [4].

Due to the difficulty of computing the nonparametric estimation of a failure rate function in the case of interval-censored data, a common approach is to utilize imputation technique in order to reduce the problem of analysing interval-censored data to that of

analysing right-censored data. By substituting an interval-censored observation with one or more points from that interval, the classical nonparametric approaches can be used for imputed interval-censored observations. For this purpose, four different imputation approaches, namely left end point, middle point, right end point and randomly uniform, are proposed for computing the nonparametric estimation (KM) of a reliability function for interval-censored data and also the regression coefficients within the Cox model. This research study assess the performance of the proposed imputation approaches on a real data set.

The remainder of the paper is structured as follows. Section 2 describes briefly the structure of censoring data. In Section 3, estimating the regression coefficients in the Cox proportional hazards model for interval-censored data is presented based on four single imputation approaches. The KM estimator of a reliability function and also the proposed approach are assessed via a real interval-censored failure time data set in Section 4.

## 2 Censoring types in survival analysis

Let  $T$  denote a nonnegative continuous random variable representing the failure time of an individual in a homogeneous population, which is the survival variable of interest. For a continuous random variable  $T$ , let  $f$  be the probability density function and  $F = \Pr(T \leq t) = \int_0^t f(u) du$  the corresponding cumulative distribution function. One of the first basic quantity used to characterize the distribution of  $T$  is the reliability function and is defined as

$$R(t) = 1 - F(t) = \Pr(T > t) = \int_t^\infty f(u) du, \quad 0 < t < \infty,$$

which is the probability of an individual to survive to time  $t$  or beyond. Another major function that describes the distribution of failure time  $T$  is the hazard rate function, which is given by

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}.$$

The hazard function can be interpreted as the instantaneous failure or death rate at time  $t$ , given that the individual has not failed before time  $t$ . Assume that the distribution of  $T$  is absolutely continuous. Therefore, the hazard function can also be represented as

$$h(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} = -\frac{d \log S(t)}{dt}. \quad (1)$$

Another feature of the event time distribution is the cumulative hazard function, which is defined as

$$H(t) = \int_0^t h(u) du = -\log S(t) = -\log[1 - F(t)].$$

Thus, it can be shown that

$$S(t) = 1 - F(t) = \exp[-H(t)] = \exp\left[-\int_0^t h(u) du\right]. \quad (2)$$

Therefore, there is a one-to-one relationship between functions  $h, S, H, F$  and  $f$ . Eventually, it follows immediately from (1) and (2) that  $f(t) = h(t) \exp[-H(t)]$ .

## 2.1 Right censoring

The most common type of censored data is right-censored data. Right censoring exists when the event of interest has not occurred before the study ends at a predetermined point of time or a subject under study is lost to follow-up before an event happens. Within the right censoring model, let us assume that there exists a lifetime  $T$  and a fixed censoring time  $C_r$ . Suppose that  $T_1, T_2, \dots, T_n$  be independent identically distributed (i.i.d.) event times from an absolutely continuous cumulative distribution function  $F_T$  with density  $f$  and the cumulative distribution function  $F_{C_r}$  of the i.i.d.censoring times  $C_{r1}, C_{r2}, \dots, C_{rn}$  is absolutely continuous with density  $g$ . The event times  $T_i$  and the censoring times  $C_{ri}$  are usually assumed to be independent.

## 2.2 Left censoring

The other kind of censoring data, which is relatively rare, is left censoring. By left-censored data, we mean that the event of interest is only known to have already occurred before the start of the study.

## 2.3 Interval censoring

Let  $T$  may not be directly observable, but instead may be censored by an interval of time between two values  $L$  and  $R$  such that  $T \in (L, R]$ , where  $L \leq R$ . Therefore, the observed event times are defined by

$$O_i = (L_i, R_i] \subset [0, \infty), \quad i = 1, \dots, n.$$

Interval-censored failure time data can be regarded as a generalization of exact, right-censored, and left-censored observations. That is, if  $L = R$ , we have an exact observation, whereas  $R = \infty$  gives a right-censored observation, and  $L = 0$  represents a left-censored observation. For illustration purposes, consider an example consists of four subjects under study. As shown in Figure 1, subject A represents a left-censored event time with its interval  $O_A = (0, 1]$ , whereas subject B is an example of exact observation since  $O_B = [2, 2] = \{2\}$ . Moreover, subject C illustrates a right-censored event time with  $O_C = (4, \infty]$  and subject D presents an example of interval-censored observation as  $O_D = (2, 4]$ .

## 3 A single point imputation approach

Censored or interval-censored failure time data differ from missing data since the former provides some incomplete information about failure variables of interest, while the latter does not give any information regarding the failure times. In the case of interval-censored, one can treat the purely interval-censored failure times as missing and replace them by using some imputed times conditional on the observed information.

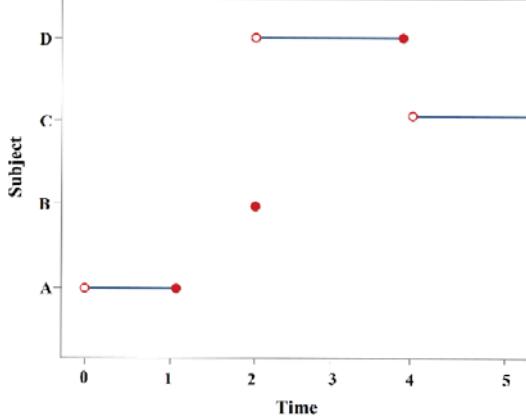


Figure 1: Illustration example of different types of censoring.

The imputation approach is to assume that for subject  $i$  under study, the underlying true failure time  $T_i$  can be a value within the observed interval  $(L_i, R_i]$ ,  $i = 1, \dots, n$ . Four different single point imputation methods for a finite interval or purely interval-censored observation, namely middle point imputation, right end point imputation, left end point imputation and randomly uniform imputation, are presented to reduce the problem of analysing interval-censored data to that of right-censored data as follow:

1. The middle point imputation approach is to let  $T_i$  be the middle point of the interval.
2. The left end point imputation approach is to take  $T_i$  to be  $L_i$ .
3. The right-end point imputation method is to take  $T_i$  to be  $R_i$ .
4. The randomly uniform imputation method is to select randomly a value based on the uniform distribution over the observed interval.

The original observations for intervals with  $R_i = \infty$  or right-censored observations are remained intact. Therefore, in this situation we can analyse a set of right-censored failure time data.

The proportional hazards or Cox model is one of the most popular methods for analyzing (right) censored survival data in the presence of covariates. Basically, the Cox model is semiparametric since it explores the relationship between the survival time and the covariates through an arbitrary baseline hazard function and the exponential of a regression function of the covariates. Within the Cox model, the hazard function of the continuously distributed survival time  $T$  with covariate vector  $Z \in \mathbb{P}$  has the form

$$h(t; Z) = h_0(t) \exp(\beta^\top Z) \quad (3)$$

where  $h_0(t)$  is an unspecified baseline hazard function and  $\beta \in \mathbb{P}$  is the vector of the unknown regression parameters. The Cox's partial likelihood (PL) approach was developed by [5] to estimate the regression coefficients of proportional hazard models, where the baseline hazard function is not involved. The main advantage of the PL approach is due to its ability to not deal with the estimation of the baseline hazard function  $h_0$  when the main interest is to estimate the regression coefficients  $\beta$ . Moreover, the asymptotic properties of the maximum partial likelihood  $\beta$  were well established by [19].

Let  $T_i$  denotes the event time with the corresponding censoring time  $C_{ri}$ . Then, the observed data consist of i.i.d. samples of the triple  $(X_i, \delta_i, Z_i)$  for  $i = 1, \dots, n$ , where  $X_i = \min(T_i, C_{ri})$  is the follow-up time,  $\delta_i = I(T_i \leq C_{ri})$ , and  $Z_i \in \mathbb{P}^p$  denotes the covariate vector. Also, the event time  $T$  and the censoring time  $C$  are assumed to be conditionally independent given  $Z = z$ . The covariate vector  $Z$  is assumed to be time fixed. We also assume that the survival time  $T$ , conditionally on  $Z = z$ , is continuous with density  $f(t | z)$ . The simplified likelihood function based on the baseline hazard  $h_0$  and the effect parameters  $\beta$  is given by

$$L(h_0, \beta) = \prod_{i=1}^n \left\{ h_0(T_i | Z_i) e^{(\beta^\top Z_i)} \right\}^{\delta_i} \exp \left\{ -e^{(\beta^\top Z_i)} H_0(T_i) \right\},$$

where  $H_0(t) = \int_0^t h_0(u) du$  denotes the baseline cumulative hazard function. Therefore, the log-likelihood function can be written as

$$\ell(h_0, \beta) = \sum_{i=1}^n \left\{ \delta_i \log h_0(T_i) + \delta_i \beta^\top Z_i - e^{(\beta^\top Z_i)} H_0(T_i) \right\}.$$

Let  $T_{(1)} < T_{(2)} < \dots < T_{(D)}$  denote the ordered distinct event times and  $Z_{(j)}$  the covariate vector of the subject whose failure time is equal to  $T_{(j)}$ . Also let  $R(T_{(j)})$  define the risk set of individuals at time  $T_{(j)}$ , where all individuals who are still under study at a time just prior to  $T_{(j)}$ . Assume that there are no ties between the event times. [6] showed that the  $\hat{\beta}$  is the maximizer of the partial likelihood function

$$\mathcal{L}_p(\beta) = \prod_{j=1}^D \frac{\exp(\beta^\top Z_{(j)})}{\sum_{i \in R(T_{(j)})} \exp(\beta^\top Z_{(i)})}.$$

In the case of tied failure times, the partial likelihood is given by

$$\mathcal{L}_p(\beta) = \prod_{j=1}^D \frac{\exp(\beta^\top Z_{(j)})}{\left\{ \sum_{i \in R(T_{(j)})} \exp(\beta^\top Z_{(i)}) \right\}^{d_j}},$$

where  $d_j$  is the number of individuals whose failure times are equal to  $T_{(j)}$ .

## 4 Numerical illustration with a real data

In this section, the analysis of the breast cancer data, observed intervals in months for times to breast retraction of early breast cancer patients, presented in [18] is discussed by using imputation approach. The estimation of the survival functions corresponding with the two different treatments, namely radiation therapy alone (RT) and radiation therapy plus adjuvant chemotherapy (RCT), are considered based on observed data within each treatment group. Figure 2 shows estimates of the two survival functions given by the middle point, left end point, right end point, and randomly uniform imputation approaches, respectively. The figure indicates that the patients in the RT treatment seem to have lower risk to develop breast retraction than those in the RCT treatment.

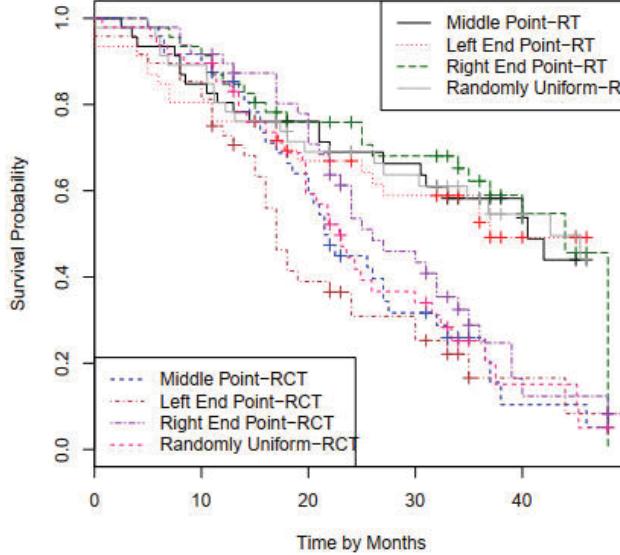


Figure 2: Estimated survival functions using single imputation approaches.

For the purpose of treatment comparison, the time to breast retraction can consider to follows the Cox model (3), where the covariate  $Z_i = 0$  indicates the patient take radiation therapy alone and  $Z_i = 1$  otherwise. In addition,  $\beta$  represents the treatment difference. Table 1 presents the results given by the single imputation approach. All of the different imputation methods give similar results about the treatment comparison and suggest that the adjuvant chemotherapy remarkably increases the risk of breast retraction.

Table 1: Estimated effects of adjuvant chemotherapy on time to breast retraction

Method	$\hat{\beta}$	SE( $\hat{\beta}$ )	p-value
Middle Point Imputation	0.8969	0.2853	0.001
Left End Point Imputation	0.9168	0.2869	0.001
Right End Point Imputation	0.8831	0.2884	0.002
Randomly Uniform Point Imputation	0.7700	0.2860	0.006

## References

- [1] O. O. Aalen, Nonparametric inference for a family of counting processes, *Annals of Statistics* 6 (1978), no. 4, 701-726.
- [2] R. A. Betensky, J. C. Lindsey, L. M. Ryan, and M. P. Wand, A local likelihood proportional hazards model for interval censored data, *Statistics in Medicine* 21 (2002), no. 2, 263-275.
- [3] Norman E Breslow, Contribution to the discussion of the paper by D. R. Cox, *Journal of the Royal Statistical Society, Series B* 34 (1972), no. 2, 216-217.
- [4] T. Cai and R. A. Betensky, Hazard regression for interval-censored data with penalized spline, *Biometrics* 59 (2003), no. 3, 570-579.

- [5] D. R. Cox, Regression models and life-tables, *Journal of the Royal Statistical Society. Series B* 34 (1972), no. 2, 187-220 (English).
- [6] , Partial likelihood, *Biometrika* 62 (1975), no. 2, 269-276.
- [7] D. M. Finkelstein, A proportional hazards model for interval-censored failure time data, *Biometrics* 42 (1986), no. 4, 845-854.
- [8] Jian Huang and Jon A. Wellner, Interval censored survival data: A review of recent progress, *Proceedings of the First Seattle Symposium in Biostatistics* (D.Y. Lin and T.R. Fleming, eds.), *Lecture Notes in Statistics*, vol. 123, Springer US, 1997, pp. 123-169.
- [9] D. A. Juckett and B. Rosenberg, Comparison of the Gompertz and Weibull functions as descriptors for human mortality distributions and their intersections, *Mechanisms of Ageing and Development* 69 (1993), no. 1-2, 1-31.
- [10] J. D. Kalbfleisch and R. L. Prentice, *The statistical analysis of failure time data*, 2nd ed., John Wiley: New York, 2002.
- [11] E. L. Kaplan and Paul Meier, Nonparametric estimation from incomplete observations, *Journal of the American Statistical Association* 53 (1958), no. 282, 457-481.
- [12] J. P. Klein and M. L. Moeschberger, *Survival analysis: Techniques for censored and truncated data*, Statistics for Biology and Health, Springer, 2003.
- [13] C. D. Lai, *Generalized weibull distributions*, Springer, 2013.
- [14] D. N. Prabhakar Murthy, Min Xie, and Renyan Jiang, *Weibull models*, Wiley Series in Probability and Statistics, Wiley, 2004.
- [15] W. Nelson, Theory and applications of hazard plotting for censored failure data, *Technometrics* 14 (1972), 945-965.
- [16] W. Pan, A multiple imputation approach to Cox regression with interval-censored data, *Biometrics* 56 (2000), no. 1, 199-203.
- [17] H. Rinne, *The weibull distribution: A handbook*, Taylor Francis, 2008.
- [18] J. Sun, *The statistical analysis of interval-censored failure time data*, Statistics for Biology and Health Series, Springer, 2006.
- [19] Anastasios A. Tsiatis, A large sample study of Coxs regression model, *Annals of Statistics* 9 (1981), no. 1, 93-108 (English).



# Statistical inferences for stress-strength in the Marshal-Olkin models based on progressively type-II censored samples

Ghanbari, S.<sup>1</sup> Rezaei, A.<sup>2</sup> Salehi, M.<sup>3</sup>

<sup>1,2</sup> Department of Statistics, Ordered and Spatial Data Center of Excellence, Ferdowsi University of mashhad

<sup>3</sup> Department of Mathematics and Statistics, Neyshabur University

## Abstract

In this article, we are mainly interested in estimating the parameter  $\mathcal{R} = P(X > Y)$  when the parent distribution follows the well-known Marshall-Olkin model and the accessible data have the form of the progressively Type-II censored sample. In this case, the parameter  $\mathcal{R}$  is free of base distribution. So, we use exponential distribution for simplicity. Also, The model is evaluated as a proportional odds. Maximum likelihood estimator, asymptotic confidence interval and bootstrap confidence interval for  $\mathcal{R}$  are derived.

**Keywords:** Stress-Strength, Marshall-Olkin distribution, Maximum likelihood estimator, Bootstrap confidence interval, Proportional odds.

## 1 Introduction

Let  $T$  be a lifetime random variable with distribution function  $F$  and survival function  $\bar{F} = 1 - F$ . The famous Cox proportional hazard model in terms of hazard rate function  $h(t) = \frac{f(t)}{\bar{F}(t)}$  is given as

$$h(t; \mathbf{x}) = h_0(t) \exp(\mathbf{b}' \mathbf{X}), \quad t \geq 0,$$

<sup>1</sup>saraghanbari5@gmail.com

<sup>2</sup>rezaei@um.ac.ir

<sup>3</sup>salehi2sms@gmail.com

where  $\mathbf{x} = (x_1, \dots, x_m)'$  is a vector of covariates (factors),  $\mathbf{b} = (b_1, \dots, b_m)'$  is an unknown vector of regression parameters and  $h_0(t)$  is an unknown hazard function called the baseline hazard function. The Cox proportional hazard model is a semi-parametric model. Cox proportional hazard model is one of the most useful models for fitting survival data based on assumptions: homogeneity of society, independence, and distribution of survival data. In this model, only some of the most famous lifetime distributions such as exponential and weibull distributions can be used and some other lifetime distributions such as Log-Logistic are not applicable. Bennett (1983) introduced the Proportional Odds (PO) model as an alternative to the classical Cox proportional hazard model. Let  $T$  represents the lifetime in a population with a vector of p-covariates  $\mathbf{z} = (z_1, \dots, z_p)'$ . In a PO model, the odds function  $\lambda_G(t; z)$  corresponding to survival function  $\bar{G}$  in presence of covariates  $\mathbf{z}$  is given by

$$\lambda_G(t; \mathbf{z}) = \lambda_F(t)[k(\mathbf{z})]^{-1}, \quad (1)$$

where  $\lambda_F(t) = \frac{F(t)}{\bar{F}(t)}$  represents an arbitrary baseline odds function w.r.t  $\bar{F}(t)$  and  $k(\mathbf{z})$  is a non-negative function of  $\mathbf{z}$  independent of time  $t$ .

From the PO model in (1) we get

$$\bar{G}(t; \mathbf{z}) = \frac{k(\mathbf{z})\bar{F}(t)}{1 - (1 - k(\mathbf{z}))\bar{F}(t)}. \quad (2)$$

In the case  $k(\mathbf{z}) = \alpha$  where  $\alpha$  is any positive constant, then we have

$$\bar{G}(t) = \frac{\alpha\bar{F}(t)}{1 - \bar{\alpha}\bar{F}(t)}, \quad (3)$$

where  $\bar{\alpha} = 1 - \alpha$ . Also, it is the survival function of the well-known Marshall-Olkin(MO) model with the base survival function  $\bar{F}(t)$ . The cumulative distribution function (cdf) and the probability density function (pdf) associated with  $\bar{G}(t)$  are given by  $G(t) = \frac{F(t)}{1 - \bar{\alpha}\bar{F}(t)}$  and  $g(t) = \frac{dG(t)}{dt} = \frac{\alpha f(t)}{[1 - \bar{\alpha}\bar{F}(t)]^2}$ , respectively. see Marshall-Olkin (1997), Murphy et al. (1997), Hunter and Lange (2002), Banerjee and Dey (2005) and Zeng et al. (2005).

Stress-Strength model is well-known in lifetime experiments. The concepts of stress-strength model was introduced by Birnbaum (1956). The stress-strength parameter in reliability system with one component is given by  $\mathcal{R} = P(X < Y)$  where  $X$  and  $Y$  are stress and strength random variables respectively (for more information see Kotz et al., 2003). There are several works on the inference methods for  $\mathcal{R}$  based on complete and incomplete data from  $X$  and  $Y$  samples. Many researchers estimated  $\mathcal{R}$  on the assumption that  $X$  and  $Y$  are distributed from a parametric model. Kundu and Gupta (2005) used two independent generalized exponential distributions with different shape parameters but having the same scale parameter. Kundu and Raqab (2009) used a three-parameter weibull distribution and Erylmaz (2010) considered an exponential distribution for a general coherent system. Also, Salehi and Ahmadi (2015) considered stress-strength reliability based on upper record ranked set sampling (RRSS) from one-parameter exponential distribution. Basirat et al. (2015) studied stress-strength reliability based the

proportional hazard rate (PHR) model when the available data have the form of progressively Type-II censored sample.

In lifetime studies, progressively Type-II censoring is of great importance. It permits one to remove live units from the test at different stages during the experiment. This method is optimal since it saves time and money while observing extreme data. Suppose we have  $n$  units implanted on a lifetime test with  $m$  failures ( $m < n$ ). When the first failure occurs,  $R_1$  of  $n - 1$  surviving units are randomly censored from the lifetime test. For the second failure  $R_2$  of the  $n - 2 - R_1$  surviving units are censored, and so on. At last, at the time of  $m$ th failure, all the remaining  $R_m = n - m - R_1 - \dots - R_{m-1}$  surviving units are censored. Observe that when  $R_1 = R_2 = \dots = R_{m-1} = 0$  then  $R_m = n - m$ , this scheme becomes to Type-II right censoring scheme. Clearly, if all  $R_i$ 's are equal to zero, we conclude the usual order statistics will be observed [see, Aggarwala (1996), Balakrishnan and Aggarwala (2000) and Balakrishnan (2007)].

The main purpose of this paper is to compare two independent MO distribution with tail parameters  $\alpha_1$  and  $\alpha_2$  while the only available data are taken from progressively Type-II censored samples. Since Gupta et al. (2010) proved that the parameter  $\mathcal{R}$  does not depend on the base distribution in the MO model, for simplicity it is often assumed that base distribution  $F_X(t)$  and  $F_Y(t)$  are exponential, with the parameters  $\lambda_1$ ,  $\lambda_2$ , respectively, with the form

$$F(t) = 1 - e^{-\lambda t}, \quad t > 0.$$

In this case, MO distribution is referred to as MO extended exponential (MOEE) distributions with survival function given by

$$\bar{G}(t) = \frac{\alpha e^{-\lambda t}}{1 - \bar{\alpha} e^{-\lambda t}},$$

The rest of this paper is organized as follows: In Section 2, we calculate  $\mathcal{R}$  in the MOEE distribution and then estimating it. Moreover, we obtain the maximum likelihood estimates (MLE) of the parameters. Also in Section 3, obtain the asymptotic confidence interval (ACI) as well as bootstrap CI of  $\mathcal{R}$ . A simulation study is done in Section 4, in order check the biase and mean squared error (MSE) of the simulated MLE of  $\mathcal{R}$  as well as the lengths and coverage probability (C.P.) of the simulated CIs of  $\mathcal{R}$ .

## 2 Point estimation

Let  $X$  and  $Y$  are random variables satisfying the PO model (3) with the tails parameters  $\alpha_1$  and  $\alpha_2$ , respectively, i.e.

$$X \sim \bar{G}_1(x; \alpha_1) = \frac{\alpha_1 \bar{F}(x)}{1 - \bar{\alpha}_1 \bar{F}(x)},$$

and

$$Y \sim \bar{G}_2(x; \alpha_2) = \frac{\alpha_2 \bar{F}(y)}{1 - \bar{\alpha}_2 \bar{F}(y)}.$$

Then, according to Gupta et al. (2010)

$$\mathcal{R} := P(X > Y) = \frac{\alpha_1/\alpha_2}{(\alpha_1/\alpha_2 - 1)^2} \left[ -\ln\left(\frac{\alpha_1}{\alpha_2}\right) + \frac{\alpha_1}{\alpha_2} - 1 \right]. \quad (4)$$

According to (4) we can conclude the following points.

- Gupta et al. (2010) proved, in this case,  $X$  and  $Y$  are i.i.d., using L'Hospital rule, if  $\alpha_1 \rightarrow \alpha_2$  then  $\mathcal{R} \rightarrow \frac{1}{2}$ . Thus, equal chance that  $X$  is bigger than  $Y$  and conversely  $Y$  is bigger than  $X$ .

- Let  $a = \frac{\alpha_1}{\alpha_2}$ , then

$$\mathcal{R} = \frac{a}{(a-1)^2} \left[ -\ln(a) + a - 1 \right],$$

therefore

$$\frac{d\mathcal{R}}{da} = \frac{2(1-a) + (1+a)\ln(a)}{(1-a)^3}.$$

Thus when  $0 < a \leq 1$ ,  $\frac{d\mathcal{R}}{da} < 0$  as a result  $\mathcal{R}$  is a decreasing function of  $a = \frac{\alpha_1}{\alpha_2}$ . That is, for a fixed  $\alpha_2$ ,  $\mathcal{R}$  is a decreasing function of  $\alpha_1$ .

- The  $\mathcal{R}$  does not depend on the base distribution of  $X$  and  $Y$  and only depends on the parameters  $\alpha_1, \alpha_2$ .

We want to estimate the  $\mathcal{R}$  which is a function of  $\alpha_1$  and  $\alpha_2$  only, so focus our attention in estimating only the parameter  $\alpha_1$  and  $\alpha_2$  assuming that  $\lambda_1$  and  $\lambda_2$  are known.

Here in this paper, we assume that  $\mathbf{X} = (X_{1:m_1:n_1}, \dots, X_{m_1:m_1:n_1})'$  and  $\mathbf{Y} = (Y_{1:m_2:n_2}, \dots, Y_{m_2:m_2:n_2})'$ , two progressively Type- II censored samples from MOEE distribution with unknown parameters  $\alpha_1, \alpha_2$  and known parameters  $\lambda_1, \lambda_2$ .

Now, we are interested in estimating  $\mathcal{R}$  based on  $\mathbf{X}$  and  $\mathbf{Y}$ . Then, the log-likelihood function based on the two independent mentioned samples is given by

$$\begin{aligned} \ell = l(\alpha_1, \alpha_2; \mathbf{x}, \mathbf{y}) &= \sum_{i=1}^{m_1} \ln[g(x_i; \alpha_1)] + \sum_{j=1}^{m_2} \ln[g(y_j; \alpha_2)] \\ &= \sum_{i=1}^{m_1} (R_{1i} + 1) \ln(\alpha_1) + \ln(\lambda_1) + \sum_{i=1}^{m_1} \lambda_1 x_i (1 - R_{1i}) \\ &\quad - 2 \sum_{i=1}^{m_1} \ln(e^{\lambda_1 x_i} - \bar{\alpha}_1) - \sum_{i=1}^{m_1} R_i \ln(1 - \bar{\alpha}_1 e^{-\lambda_1 x_i}) \\ &\quad + \sum_{j=1}^{m_2} (R_{2j} + 1) \ln(\alpha_2) + \ln(\lambda_2) + \sum_{j=1}^{m_2} \lambda_2 y_j (1 - R_{2j}) \\ &\quad - 2 \sum_{j=1}^{m_2} \ln(e^{\lambda_2 y_j} - \bar{\alpha}_2) - \sum_{j=1}^{m_2} R_{2j} \ln(1 - \bar{\alpha}_2 e^{-\lambda_2 y_j}), \end{aligned}$$

where  $\mathbf{x} = (x_1, \dots, x_{m_1})'$  and  $\mathbf{y} = (y_1, \dots, y_{m_2})'$  are the observations of  $\mathbf{X}$  and  $\mathbf{Y}$ , respectively.

To this end, we start off with obtaining the MLE's of  $\mathcal{R}$ . In this regard, the MLE's of the parameters  $\alpha_1$  and  $\alpha_2$ , denoted by  $\hat{\alpha}_1, \hat{\alpha}_2$ , are the solutions of the following nonlinear equations.

$$\begin{aligned}\frac{\partial l}{\partial \alpha_1} &= \frac{\sum_{i=1}^{m_1} (R_{1i} + 1)}{\alpha_1} - 2 \sum_{i=1}^{m_1} \frac{1}{e^{\lambda_1 x_i} - \bar{\alpha}_1} - \sum_{i=1}^{m_1} \frac{R_{1i} e^{-\lambda_1 x_i}}{1 - \bar{\alpha}_1 e^{-\lambda_1 x_i}} = 0, \\ \frac{\partial l}{\partial \alpha_2} &= \frac{\sum_{j=1}^{m_2} (R_{2j} + 1)}{\alpha_2} - 2 \sum_{j=1}^{m_2} \frac{1}{e^{\lambda_2 y_j} - \bar{\alpha}_2} - \sum_{j=1}^{m_2} \frac{R_{2j} e^{-\lambda_2 y_j}}{1 - \bar{\alpha}_2 e^{-\lambda_2 y_j}} = 0.\end{aligned}\quad (5)$$

Since the above equations do not have the theoretical solution, we use numerical methods to solve them. From (4) and using the invariance property of MLE, we have the MLE of  $\mathcal{R}$  as

$$\hat{\mathcal{R}} = \frac{\hat{\alpha}_1/\hat{\alpha}_2}{(\hat{\alpha}_1/\hat{\alpha}_2 - 1)^2} \left[ -\ln \left( \frac{\hat{\alpha}_1}{\hat{\alpha}_2} \right) + \frac{\hat{\alpha}_1}{\hat{\alpha}_2} - 1 \right]. \quad (6)$$

### 3 Interval estimation of $\mathcal{R}$

In the section, we derive two confidence interval of  $\mathcal{R}$  given by (4). Hence, we limit ourselves in obtaining the approximate CI's such as asymptotic and bootstrap ones. It seems that there is no closed form for the distribution.

#### 3.1 Asymptotic confidence interval

Suppose that the convergence in distribution denote  $\xrightarrow{d}$ . We know that as  $m_1, m_2 \rightarrow \infty$ , then  $(\hat{\alpha}_1 - \alpha) \xrightarrow{d} N(0, Var(\hat{\alpha}_1))$  and  $(\hat{\alpha}_2 - \alpha) \xrightarrow{d} N(0, Var(\hat{\alpha}_2))$  where  $Var(\hat{\alpha}_1) = I_{11}^{-1}$ ,  $Var(\hat{\alpha}_2) = I_{22}^{-1}$ . By performing algebraic calculations, we have

$$\begin{aligned}I_{11} = E \left( -\frac{\partial^2 l}{\partial \alpha_1^2} \right) &= \sum_{i=1}^{m_1} \frac{(R_{1i} + 1)}{\alpha_1^2} - 2m_1 E \left( \frac{1}{(e^{\lambda_1 X} - \bar{\alpha}_1)^2} \right) \\ &\quad - \sum_{i=1}^{m_1} R_{1i} E \left( \frac{e^{-2\lambda_1 X}}{(1 - \bar{\alpha}_1 e^{-\lambda_1 X})^2} \right),\end{aligned}\quad (7)$$

$$\begin{aligned}I_{22} = E \left( -\frac{\partial^2 l}{\partial \alpha_2^2} \right) &= \sum_{j=1}^{m_2} \frac{(R_{2j} + 1)}{\alpha_2^2} - 2m_2 E \left( \frac{1}{(e^{\lambda_2 Y} - \bar{\alpha}_2)^2} \right) \\ &\quad - \sum_{j=1}^{m_2} R_{2j} E \left( \frac{e^{-2\lambda_2 Y}}{(1 - \bar{\alpha}_2 e^{-\lambda_2 Y})^2} \right),\end{aligned}\quad (8)$$

$$I_{12} = I_{21} = E \left( -\frac{\partial^2 l}{\partial \alpha_1 \partial \alpha_2} \right) = 0.$$

Now, suppose that  $m_1, m_2$  are sufficiently large in such a way that  $\frac{m_2}{m_1} \rightarrow p$ , where  $p \in (0, 1)$ , so we readily conclude that the asymptotic property of the MLE, as  $m_1 \rightarrow \infty$  and  $m_2 \rightarrow \infty$ ,

$$\sqrt{m_1} \begin{pmatrix} \hat{\alpha}_1 - \alpha_1 \\ \hat{\alpha}_2 - \alpha_2 \end{pmatrix} \xrightarrow{d} N_2 \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{a_{11}} & 0 \\ 0 & \frac{1}{p \times a_{22}} \end{pmatrix} \right),$$

where

$$\begin{aligned} a_{11} &= \lim_{m_1, m_2 \rightarrow \infty} \frac{1}{m_1} E \left( -\frac{\partial^2 l}{\partial \alpha_1^2} \right), \\ a_{22} &= \lim_{m_1, m_2 \rightarrow \infty} \frac{1}{m_2} E \left( -\frac{\partial^2 l}{\partial \alpha_2^2} \right). \end{aligned} \quad (9)$$

Taking

$$g(t_1, t_2) = \frac{t_1}{t_2(\frac{t_1}{t_2} - 1)^2} \left[ -\ln \left( \frac{t_1}{t_2} \right) + \frac{t_1}{t_2} - 1 \right].$$

Then using delta method yields (see, e.g. Lehman and cassela, 1998; page: 240-241)

$$(\hat{\mathcal{R}} - \mathcal{R}) = \left( g(\hat{\alpha}_1, \hat{\alpha}_2) - g(\alpha_1, \alpha_2) \right) \xrightarrow{d} N(0, Var(\hat{\mathcal{R}})), \quad (10)$$

so that

$$Var(\hat{\mathcal{R}}) = Var(\hat{\alpha}_1) \left( \frac{\partial \hat{\mathcal{R}}}{\partial \hat{\alpha}_1} \right)^2 + Var(\hat{\alpha}_2) \left( \frac{\partial \hat{\mathcal{R}}}{\partial \hat{\alpha}_2} \right)^2, \quad (11)$$

where

$$\frac{\partial \hat{\mathcal{R}}}{\partial \hat{\alpha}_1} = \frac{\hat{\alpha}_2}{(\hat{\alpha}_1 - \hat{\alpha}_2)^3} \left[ -2(\hat{\alpha}_1 - \hat{\alpha}_2) + (\hat{\alpha}_1 + \hat{\alpha}_2) \ln \left( \frac{\hat{\alpha}_1}{\hat{\alpha}_2} \right) \right],$$

and

$$\begin{aligned} \frac{\partial \hat{\mathcal{R}}}{\partial \hat{\alpha}_2} &= \frac{\hat{\alpha}_1}{(\hat{\alpha}_1 - \hat{\alpha}_2)^3} \left[ 2(\hat{\alpha}_1 - \hat{\alpha}_2) - (\hat{\alpha}_1 + \hat{\alpha}_2) \ln \left( \frac{\hat{\alpha}_1}{\hat{\alpha}_2} \right) \right], \\ &= -\frac{\hat{\alpha}_1}{\hat{\alpha}_2} \frac{\partial \hat{\mathcal{R}}}{\partial \hat{\alpha}_1}. \end{aligned}$$

By using the property in (10) and applying the Slutsky's theorem, a  $100(1 - \gamma)\%$  asymptotic CI for  $\mathcal{R}$  is derived as follows

$$\left( \hat{\mathcal{R}} - z_{1-\frac{\gamma}{2}} \sqrt{\widehat{Var}(\hat{\mathcal{R}})}, \hat{\mathcal{R}} + z_{1-\frac{\gamma}{2}} \sqrt{\widehat{Var}(\hat{\mathcal{R}})} \right), \quad (12)$$

where  $z_\nu$  is  $\nu$ th quantile of the standard normal distribution and  $\widehat{Var}(\hat{\mathcal{R}})$  is obtained by replacing the plug-in estimators  $\hat{\alpha}_1$  and  $\hat{\alpha}_2$  instead of  $\alpha_1$  and  $\alpha_2$ , respectively, in (11).

### 3.2 Bootstrap confidence interval

The bootstrap method is used when we can not obtain the exact distribution of the interesting statistic. There are several ways to construct bootstrap CI. The following algorithm is used to construct the bootstrap CI for  $\mathcal{R}$  in this paper.

#### Algorithm 3.2. Percentile (Boot<sub>1</sub> CI)

- Step 1. Based on the independent observed progressively Type-II censored samples from MOEE distribution  $\mathbf{x} = (x_1, \dots, x_{m_1})'$  and  $\mathbf{y} = (y_1, \dots, y_{m_2})'$ , calculate  $\hat{\alpha}_1$ ,  $\hat{\alpha}_2$  and  $\widehat{\mathcal{R}}$  from (4) and (3), respectively.
- Step 2. Generate independent bootstrap samples  $X^* \sim MOEE(\hat{\alpha}_1)$ ,  $Y^* \sim MOEE(\hat{\alpha}_2)$  and  $\widehat{\mathcal{R}}^*$ .
- Step 3. Repeat Step 2 for  $b = 1, \dots, B$ , to derive  $\widehat{\mathcal{R}}_j^*$ ,  $j = 1, \dots, B$ .
- Step 4. Let  $\widehat{Q}(\nu)$  be the  $\nu$  percentile of  $\{\widehat{\mathcal{R}}_j^*, j = 1, 2, \dots, B\}$ , i.e.  $\widehat{Q}(\nu)$  is such that

$$\frac{1}{B} \sum_{j=1}^B I(\widehat{\mathcal{R}}_j^* < \widehat{Q}(\nu)) = \nu.$$

Then the  $100(1 - \gamma)\%$  percentile CI of  $\mathcal{R}$  is

$$\left( \widehat{Q}^{-1}\left(\frac{\gamma}{2}\right), \widehat{Q}^{-1}\left(1 - \frac{\gamma}{2}\right) \right).$$

## 4 Simulation study

In this section, we study the performance of the point estimator and CIs for  $\mathcal{R}$ . We generated progressively Type-II censored samples from the MOEE distribution using the algorithm presented in Balakrishnan and Sandhu (1995). Also, for given  $(n_1, m_1)$  and progressive censoring scheme  $(R_1, \dots, R_{m_1})$  for the first population, and  $(n_2, m_2)$  and progressive censoring scheme  $(R'_1, \dots, R'_{m_2})$  for the second population. Here, in order to easy calculation we suppose that  $n_1 = 2m_1$  and  $n_2 = 2m_2$ . All combination of  $m_1 = 10, 15$ ,  $m_2 = 10, 15$ ,  $\alpha_1 = 0.5$ ,  $\lambda_1 = 0.1, 10$ ,  $\lambda_2 = 0.2, 20$ ,  $R = 0.1, 0.3, 0.5, 0.9$  and  $\alpha = 0.95$  are used. In each combination  $10^5$  samples of  $\mathbf{x}$  and  $\mathbf{y}$  from  $MOEE(\alpha_1, \lambda_1)$  and  $MOEE(\alpha_2, \lambda_2)$  are simulated, respectively. Note that, we fix  $B = 300$  and  $B' = 30$ . For a given  $m_1$  and  $m_2$ , three different censoring schemes  $A$ ,  $B$  and  $C$  are used to generate the progressively censored samples:

Censoring scheme	$R = (R_1, \dots, R_m)$
$A$	Type-II censoring, $R = (0, 0, \dots, n - m)$
$B$	left censoring, $R = (n - m, 0, \dots, 0, 0)$
$C$	all $R_i$ , taken the same number

The results of the ACI, Boot<sub>1</sub> confidence intervals are presented in Tables 1 and 2. As mentioned earlier, parameter  $\mathcal{R}$  does not depend on the initial distribution  $F_0$ . Here, the entries of Tables 1 and 2 also confirm this fact. In all CIs, when the sample size increase, we see that the expected length is almost decreasing. It seems that the maximum of the

Table 1: Average bias, average  $MSE$ , Coverage Probability (C.P.) and Expected Length (E.L.) of  $\mathcal{R}$  for  $\lambda_1 = 0.1$ ,  $\lambda_2 = 0.2$  and  $\alpha = 0.95$

$\mathcal{R}$	$m_1$	$m_2$	R	$R'$	MLE		ACI		Boot <sub>1</sub>	
					Bias	MSE	C.P.	E.L.	C.P.	E.L.
0.1	10	10	A	A	0.0065	0.0017	0.8924	0.2070	0.9223	0.1545
0.1	10	10	B	B	0.0098	0.0029	0.8904	0.2090	0.9117	0.1535
0.1	10	10	C	C	0.0073	0.0020	0.8889	0.2106	0.9105	0.1673
0.1	10	15	A	A	0.0044	0.0013	0.8936	0.2050	0.9257	0.1387
0.1	10	15	B	B	0.0076	0.0024	0.8910	0.2055	0.9217	0.1790
0.1	10	15	C	C	0.0051	0.0016	0.8890	0.2103	0.9211	0.1501
0.1	15	10	A	A	0.0058	0.0014	0.9057	0.2030	0.9297	0.1376
0.1	15	10	B	B	0.0093	0.0025	0.8957	0.2040	0.9287	0.1819
0.1	15	10	C	C	0.0067	0.0017	0.8895	0.2059	0.9267	0.1498
0.3	10	10	A	A	0.0057	0.0067	0.8994	0.2137	0.9603	0.2007
0.3	10	10	B	B	0.0092	0.0107	0.8987	0.2159	0.9580	0.2157
0.3	10	10	C	C	0.0066	0.0077	0.8967	0.2168	0.9600	0.2166
0.3	10	15	A	A	0.0031	0.0058	0.8998	0.2126	0.9556	0.1819
0.3	10	15	B	B	0.0059	0.0093	0.8989	0.2152	0.9617	0.2132
0.3	10	15	C	C	0.039	0.0066	0.8975	0.2167	0.9576	0.2145
0.3	15	10	A	A	0.0055	0.0059	0.9057	0.2117	0.9553	0.1709
0.3	15	10	B	B	0.0082	0.0094	0.9015	0.2150	0.9463	0.2122
0.3	15	10	C	C	0.0062	0.0068	0.9010	0.2164	0.9536	0.2114
0.5	10	10	A	A	0.0011	0.0091	0.9225	0.3812	0.9493	0.3532
0.5	10	10	B	B	0.0012	0.0141	0.9217	0.3882	0.9476	0.3665
0.5	10	10	C	C	0.0011	0.0105	0.9205	0.3902	0.9490	0.3740
0.5	10	15	A	A	0.0001	0.0077	0.9236	0.3805	0.9433	0.3258
0.5	10	15	B	B	0.0009	0.0012	0.9229	0.3885	0.9510	0.3657
0.5	10	15	C	C	0.0027	0.0088	0.9202	0.3896	0.9453	0.3464
0.5	15	10	A	A	0.0017	0.0076	0.9247	0.3789	0.9550	0.3248
0.5	15	10	B	B	0.0013	0.0119	0.9238	0.3895	0.9526	0.3647
0.5	15	10	C	C	0.0015	0.0087	0.9200	0.3884	0.9516	0.3462
0.9	10	10	A	A	-0.0070	0.0017	0.9313	0.1678	0.9426	0.1599
0.9	10	10	B	B	-0.0109	0.0030	0.9310	0.1885	0.9446	0.1878
0.9	10	10	C	C	-0.0081	0.0021	0.9297	0.1902	0.9447	0.1889
0.9	10	15	A	A	-0.0049	0.0013	0.9325	0.1667	0.9420	0.1586
0.9	10	15	B	B	-0.0092	0.0024	0.9315	0.1882	0.9483	0.1889
0.9	10	15	C	C	-0.0069	0.0016	0.9309	0.1897	0.9413	0.1876
0.9	15	10	A	A	-0.0050	0.0014	0.9337	0.1652	0.9400	0.1521
0.9	15	10	B	B	-0.0053	0.0022	0.9329	0.1879	0.9427	0.1860
0.9	15	10	C	C	-0.0057	0.0016	0.9319	0.1893	0.9406	0.1866

expected length occurs at  $\mathcal{R} = 0.5$  and the expected lengths are very small for the extreme values of  $\mathcal{R}$  near to zero and one. Note that because the sample size in this simulation is maximum 15, the asymptotic CI is less accurate than the bootstrap CI.

Table 2: Average bias, average  $MSE$ , Coverage Probability (C.P.) and Expected Length (E.L.) of  $\mathcal{R}$  for  $\lambda_1 = 10$ ,  $\lambda_2 = 20$  and  $\alpha = 0.95$

$\mathcal{R}$	$m_1$	$m_2$	R	$R'$	MLE		ACI		Boot <sub>1</sub>	
					Bias	MSE	C.P.	E.L.	C.P.	E.L.
0.1	10	10	A	A	0.0066	0.0018	0.8923	0.2071	0.9225	0.1542
0.1	10	10	B	B	0.0097	0.0027	0.8902	0.2092	0.9115	0.1533
0.1	10	10	C	C	0.0072	0.0022	0.8887	0.2103	0.9102	0.1670
0.1	10	15	A	A	0.0042	0.0012	0.8932	0.2051	0.9255	0.1385
0.1	10	15	B	B	0.0073	0.0023	0.8913	0.2052	0.9212	0.1791
0.1	10	15	C	C	0.0050	0.0014	0.8892	0.2104	0.9212	0.1505
0.1	15	10	A	A	0.0055	0.0016	0.9051	0.2032	0.9294	0.1372
0.1	15	10	B	B	0.0091	0.0027	0.8953	0.2041	0.9281	0.1814
0.1	15	10	C	C	0.0065	0.0016	0.8891	0.2057	0.9266	0.1491
0.3	10	10	A	A	0.0055	0.0065	0.8994	0.2139	0.9603	0.2017
0.3	10	10	B	B	0.0093	0.0105	0.8985	0.2154	0.9580	0.2159
0.3	10	10	C	C	0.0064	0.0074	0.8966	0.2167	0.9600	0.2164
0.3	10	15	A	A	0.0030	0.0058	0.8992	0.2125	0.9556	0.1819
0.3	10	15	B	B	0.0058	0.0093	0.8982	0.2151	0.9617	0.2134
0.3	10	15	C	C	0.0037	0.0065	0.8975	0.2165	0.9578	0.2147
0.3	15	10	A	A	0.0054	0.0057	0.9056	0.2113	0.9555	0.1709
0.3	15	10	B	B	0.0082	0.0092	0.9013	0.2151	0.9463	0.2123
0.3	15	10	C	C	0.0061	0.0065	0.9013	0.2162	0.9538	0.2116
0.5	10	10	A	A	0.0013	0.0091	0.9226	0.3815	0.9493	0.3532
0.5	10	10	B	B	0.0011	0.0141	0.9217	0.3882	0.9477	0.3665
0.5	10	10	C	C	0.0015	0.0105	0.9208	0.3902	0.9490	0.3740
0.5	10	15	A	A	0.0003	0.0078	0.9236	0.3807	0.9435	0.3259
0.5	10	15	B	B	0.0008	0.0012	0.9229	0.3885	0.9510	0.3657
0.5	10	15	C	C	0.0027	0.0089	0.9205	0.3898	0.9454	0.3464
0.5	15	10	A	A	0.0017	0.0076	0.9247	0.3789	0.9550	0.3249
0.5	15	10	B	B	0.0015	0.0119	0.9239	0.3898	0.9526	0.3647
0.5	15	10	C	C	0.0012	0.0087	0.9200	0.3884	0.9516	0.3462
0.9	10	10	A	A	-0.0071	0.0017	0.9313	0.1678	0.9426	0.1599
0.9	10	10	B	B	-0.0107	0.0030	0.9311	0.1887	0.9446	0.1878
0.9	10	10	C	C	-0.0085	0.0023	0.9297	0.1902	0.9448	0.1889
0.9	10	15	A	A	-0.0049	0.0014	0.9326	0.1667	0.9420	0.1588
0.9	10	15	B	B	-0.0093	0.0025	0.9317	0.1882	0.9485	0.1889
0.9	10	15	C	C	-0.0069	0.0016	0.9309	0.1897	0.9413	0.1876
0.9	15	10	A	A	-0.0055	0.0014	0.9338	0.1653	0.9405	0.1524
0.9	15	10	B	B	-0.0054	0.0022	0.9329	0.1879	0.9428	0.1865
0.9	15	10	C	C	-0.0057	0.0016	0.9319	0.1893	0.9406	0.1866

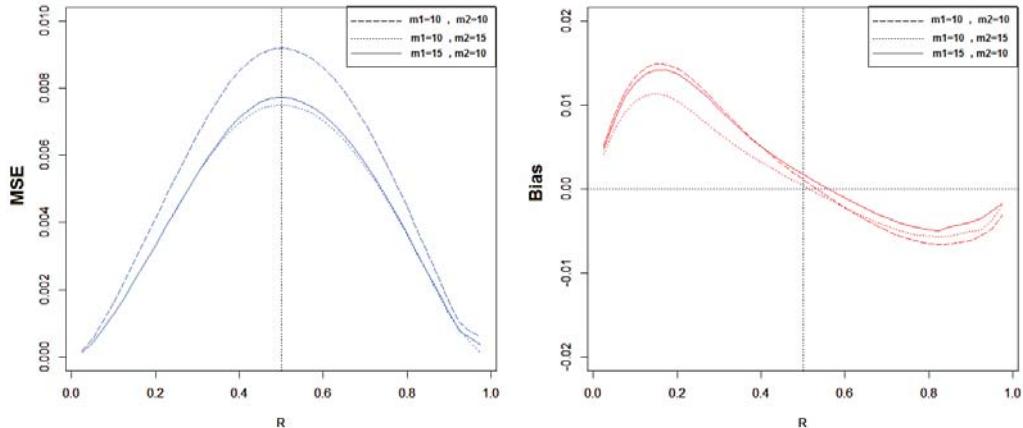


Figure 1: plot of  $MSE(\hat{\mathcal{R}}, \mathcal{R})$  ( left) and  $Bias(\hat{\mathcal{R}}, \mathcal{R})$ (right) versus  $\mathcal{R}$

As regards, there is no difference between different censorship schemes (A, B and C) and the choice of basic distribution parameters. We have plotted MSE and Bias of the MLE estimator of  $\mathcal{R}$ , also used of the A Censoring scheme and base distribution parameters  $\lambda_1 = 0.1, \lambda_2 = 0.2$ .

From Figure 1 we observe the following points:

- When  $\mathcal{R} < 0.5$ , the bias is positive and when  $\mathcal{R} > 0.5$ , the bias is negative.
- As it is expected, the MSE and the bias are reduced by increasing the sample sizes.
- The  $MSE(\hat{\mathcal{R}}, \mathcal{R})$  is symmetric, about  $\mathcal{R} = 0.5$ , when  $m_1 = m_2$ , and departures from symmetry when  $m_1 < m_2$  and  $m_1 > m_2$ , respectively.

## Acknowledgement

The authors are grateful to the referees to provide useful comments in promoting the presentation.

## References

- [1] Aggarwala R. (1996), Advances in life testing: progressive censoring and generalized distribution [PhD thesis]. Hamilton, ON: McMaster University.
- [2] Balakrishnan N, Aggarwala R. (2000), Progressive censoring, theory, methods and applications. Boston, MA: Birkhäuser.
- [3] Balakrishnan N. (2007), Progressive censoring methodology: *an appraisal Test* , **16**, 290–296.
- [4] Balakrishnan N, Sandhu RA. (1995), A simple simulative algorithm for generating progressive type-II censored samples. *Am Stat* , **49**, 229–230.
- [5] Banerjee S, Dey DK. (2005), Semi parametric proportional odds models for spatially correlated survival data. *Lifetime Data Anal* , **11**, 175–191.

- [6] Basirat M, Baratpour S, Ahmadi J. (2015), Statistical inferences for stress–strength in the proportional hazard models based on progressive Type-II censored samples. *J Stat Comput Simul*, **85**, 431–449.
- [7] Bennett S. (1983), Analysis of survival data by the proportional odds model. *Stat Med*, **2**, 273–277.
- [8] Birnbaum ZW. (1956), On a use of Mann–Whitney statistics. In: Neyman, J., editor. Proceedings of the Third Berkeley Symposium in Mathematics, Statistics and Probability, Vol. 1. Berkeley, CA: University of California Press, 13–17.
- [9] Casella G, Berger RL. (1990), Statistical inference. Pacific Grove, CA: Wadsworth & Brooks/Cole.
- [10] Erylmaz S. (2010), On system reliability in stress–strength setup. *Stat Probab Lett*, **80**, 834–839.
- [11] Gupta RC, Ghitany ME, Al-Mutairi DK. (2010), Estimation of reliability from Marshall–Olkin extended Lomax distributions. *J Stat Comput Simul*, **80**, 937–947.
- [12] Hunter DR, Lange K. (2002), Comparing estimates in the proportional odds model. *Ann Inst Stat Math*, **54**, 155–168.
- [13] Kotz S, LumelskiiY, Pensky M. (2003), The stress–strength model and its generalizations: theory and applications. Singapore: World Scientific.
- [14] Kundu D, Gupta RD. (2005), Estimating of  $P(Y < X)$  for the generalized exponential distribution. *Metrika*, **61**, 291–308.
- [15] Kundu D, Raqab MZ. (2009), Estimation of  $P(Y < X)$  for three-parameter Weibull distribution. *Stat Probab Lett*, **79**, 1839–1846.
- [16] Marshall AW, Olkin I. (1997), A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, **84**, 641–652.
- [17] Murphy SA, Rossini AJ, Van derVarrt AW. (1997), Maximum likelihood estimation in the proportional odds model. *J Am Stat Assoc*, **92**, 968–976.
- [18] Salehi M, Ahmadi Jafar. (2015), Estimation of Stress–Strength reliability using record ranked set sampling scheme from the exponential distribution. *FILOMAT*, **29**, 1149–1162.
- [19] Zeng D, Lin DY, Yin G. (2005), Maximum likelihood estimation for the proportional odds model with random effects. *J Am Stat Assoc*, **100**, 470–483.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## Comparison of equivalency and optimality between constant-stress and step-stress tests under type-I censoring for Rayleigh distribution

Hakamipour, N.<sup>1</sup>

Department of Mathematics, Buein Zahra Technical University, Buein Zahra, Qazvin

### Abstract

By running life tests at higher stress levels than normal operating conditions, accelerated life testing quickly yields information on the lifetime distribution of a test unit. The lifetime at the design stress is then estimated through extrapolation using a regression model. In constant-stress testing, a unit is tested at a fixed stress level until failure or the termination time point of test, whereas step-stress testing allows the experimenter to gradually increase the stress levels at some pre-fixed time points during the test. In this work, the optimal  $k$ -level constant-stress and step-stress ALTs are compared for the Rayleigh failure data under complete sampling and Type-I censoring. The objective is to quantify the advantage of using the step-stress testing relative to the constant-stress one. A log-linear relationship between the scale parameter and stress level is assumed and the Khamis-Higgins model holds for the effect of changing stress in step-stress testing. The optimal design point is then determined under C/D/A-optimality criteria. The efficiency of step-stress testing to constant-stress one is then discussed in terms of the ratio of optimal objective functions based on the information matrix.

**Keywords:** Accelerated Life Testing, Constant-Stress Testing, Optimal Regression Design, Step-Stress Testing, Type-I Censoring.

---

<sup>1</sup>nooshin.hakami@aut.ac.ir

# 1 Introduction

With increasing reliability and substantially long life-spans of products, it is often difficult for standard life testing methods under normal operating conditions to obtain sufficient information about the failure time distribution of the products. This difficulty is overcome by accelerated life test (ALT) where the test units are subjected to higher stress levels than normal for rapid failures. By applying more severe stresses, ALT collects information on the parameters of lifetime distributions more quickly. A key references in the area of ALT is Nelson [6]. There are two special classes of ALT: constant-stress testing and step-stress testing. In constant-stress testing, a unit is tested at a fixed stress level until failure occurs or the life test is terminated, whichever comes first. On the other hand, step-stress testing allows the experimenter to gradually increase the stress levels at some prefixed time points during the test.

The optimal ALT design has attracted great attention in the reliability literature.

Ng et al. [3] discussed the problem of determining the optimal sample size allocation for a general  $k$ -level model with extreme value regression, whereas Gouno et al. [2]. Balakrishnan and Han [1] discussed the problem of determining the optimal stress duration under progressive Type-I censoring. Under complete sampling, Hu et al. [4] studied the statistical equivalency of a simple step-stress ALT to other stress loading designs.

The main focus of this article is to investigate the advantage of using step-stress ALT relative to constant-stress ALT. Assuming a log-linear relationship between the mode lifetime parameter and stress level, with the Khamis-Higgins (KH) model for the effect of changing stress in step-stress ALT. The KH model is based on a time transformation of the cumulative exposure model. Khamis and Higgins [5] have proposed this model for multiple step-stress testing. The optimal design point is determined under various optimality criteria. In particular, the cases of Type-I censoring is considered under Rayleigh lifetime distribution for units subjected to stress. Using the ratio of optimal objective functions as a measure of relative efficiency, comparison of  $k$ -level step-stress testing to  $k$ -level constant-stress testing is discussed through a numerical study.

The rest of the paper is organized as follows Section 2 presents the model description and formulation, derives the MLEs of the model parameters and the associated Fisher information for  $k$ -level constant-stress ALT and step-stress ALT. Section 3 then defines the three optimality criteria based on the Fisher information (viz., variance, determinant and trace). Section 4 provides the results of a numerical study and discusses the relative efficiency of these two classes of ALT. Finally, Section 5 is devoted to some conclusions.

## 2 Model description, MLEs and Fisher information

Let  $s$  be the given stress loading for ALT. Also, let  $s_k$  be the upper bound of stress level and  $s_0$  be the normal use-stress level that  $s_k < s_0$ . The standardized stress loading is then defined as

$$x_i = \frac{s_i - s_0}{s_k - s_0}, \quad \text{for } i = 1, 2, \dots, k;$$

so that the range of  $x_i$  is  $[0, 1]$ .

Let us define  $0 \leq x_0 \leq x_1 < x_2 < \dots < x_k \leq 1$  to be the ordered standardized stress levels to be used in the test. As the basis of constructing an ALT model, it is further assumed that under any specific stress level  $x_i$ , the Rayleigh distribution describes the failure mechanism of a test unit. That is, the probability density function (PDF) and the corresponding cumulative distribution function (CDF) of the lifetime of a test unit at stress level  $x_i$  are

$$f_i(y) = \frac{t}{\theta_i} e^{-\frac{t^2}{2\theta_i^2}}, \quad 0 \leq y, \quad (1)$$

$$F_i(y) = 1 - e^{-\frac{t^2}{2\theta_i^2}}, \quad 0 \leq y. \quad (2)$$

respectively. Also, it is assumed that under any stress level  $x_i$ , the mode time to failure (MTTF) of a test unit,  $\theta_i$ , is a log-linear function of stress given by

$$\log \theta_i = \beta_0 + \beta_1 x_i, \quad (3)$$

where the regression parameters  $\beta_0$  and  $\beta_1$  are unknown and need to be estimated. The log-linear relationship is a commonly used and well studied model in ALT.

Here we consider two popular classes of ALT: constant-stress and step-stress. In constant-stress testing, a unit is tested at a fixed stress level until failure occurs or the life test is terminated, whichever comes first. On the other hand, step-stress testing allows the experimenter to gradually increase the stress levels at some pre-fixed time points during the test. The following subsections present the MLEs of  $\beta_0$  and  $\beta_1$  and the associated Fisher information for general  $k$ -level constant-stress ALT and step-stress ALT.

## 2.1 $k$ -level constant-stress test under time constraint

A constant-stress test under time constraint proceeds as follows. For  $i = 1, 2, \dots, k$ ,  $N_i \equiv n\pi_i$  units are allocated on test at stress level  $x_i$  such that  $\sum_{i=1}^k N_i = n$ . or equivalently,  $\sum_{i=1}^k \pi_i = 1$ .  $\pi_i = N_i/n$  is the allocation proportion of units (out of total  $n$  units under the test) assigned to stress level  $x_i$ . The allocated units are then tested until time  $\Delta_i$  at which point all the surviving items are withdrawn, thereby terminating the life test. Let  $n_i$  denote the number of units failed at stress level  $x_i$  in time interval  $[0, \Delta_i]$  and  $y_{i,l}$  denote the  $l$ th ordered failure time of  $n_i$  units at  $x_i$ ,  $l = 1, 2, \dots, n_i$  while  $N_i - n_i$  denotes the number of units censored at time  $\Delta_i$ . Obviously, when there is no right censoring (viz.,  $\Delta_i = \infty$  and  $n_i = N_i$ ), this situation corresponds to the  $k$ -level constant-stress testing under complete sampling as a special case.

Now, using (1) and (2), the likelihood function of  $\mathbf{n} = (n_1, n_2, \dots, n_k)$  and  $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k)$  with  $\mathbf{y}_i = (y_{i,1}, y_{i,2}, \dots, y_{i,n_i})$  is obtained as

$$f_J(\mathbf{y}, \mathbf{n}) = \left[ \prod_{i=1}^k \frac{N_i!}{(N_i - n_i)!} \right] \left[ \prod_{i=1}^k \prod_{l=1}^{n_i} \theta_i^2 \right] \left[ \prod_{i=1}^k \prod_{l=1}^{n_i} y_{il} \right] e^{-\sum_{i=1}^k \frac{U_i}{2\theta_i^2}}, \quad (4)$$

where

$$U_i = \sum_{l=1}^{n_i} y_{i,l}^2 + (N_i - n_i)\Delta_i^2, \quad i = 1, 2, \dots, k.$$

Using (4) and the log-linear link given in (3), the log-likelihood function of  $(\beta_0, \beta_1)$  can be written as

$$\ell(\beta_0, \beta_1) \propto -2\beta_0 \sum_{i=1}^k n_i - 2\beta_1 \sum_{i=1}^k x_i n_i - \sum_{i=1}^k n_i \left( \frac{1}{2} e^{-2\beta_0 - 2\beta_1 x_i} U_i \right). \quad (5)$$

Upon differentiating (5) with respect to  $\beta_0$  and  $\beta_1$ , the MLEs  $\hat{\beta}_0$  and  $\hat{\beta}_1$  are obtained as simultaneous solutions to the following two equations:

$$\left[ \sum_{i=1}^k n_i \right] \left[ \sum_{i=1}^k e^{-2\hat{\beta}_1 x_i} U_i x_i \right] = \left[ \sum_{i=1}^k x_i n_i \right] \left[ \sum_{i=1}^k e^{-2\hat{\beta}_1 x_i} U_i \right], \quad (6)$$

$$\hat{\beta}_0 = \frac{1}{2} \log \left( \frac{\sum_{i=1}^k e^{-2\hat{\beta}_1 x_i} U_i}{2 \sum_{i=1}^k n_i} \right). \quad (7)$$

From (6), for the existence of  $\hat{\beta}_1$  and also of  $\hat{\beta}_0$ , at least one failure has to be observed from at least two different stress levels. Otherwise, the parameters are not estimable. Both sides of (6) are also monotone decreasing functions of  $\hat{\beta}_1$ . Thus, their intersection guarantees the unique existence of  $\hat{\beta}_1$  and also of  $\hat{\beta}_0$  from (7).

As shown above,  $\hat{\beta}_0$  and  $\hat{\beta}_1$  are nonlinear functions of random quantities and thus, statistical inference with these MLEs can be based on the asymptotic distributional result that the vector  $(\hat{\beta}_0, \hat{\beta}_1)$  is approximately distributed as a bivariate normal with mean vector  $(\beta_0, \beta_1)$  and variance-covariance matrix  $\mathbf{I}_n^{-1}(\beta_0, \beta_1)$ , where  $\mathbf{I}_n(\beta_0, \beta_1)$  is the Fisher information matrix. By using the proposition 2.1,  $\mathbf{I}_n(\beta_0, \beta_1)$  is obtained in the Theorem 2.1.

**Proposition 2.1.** *For  $i = 1, 2, \dots, k$ , the random variable  $n_i$  has a binomial distribution with parameters  $(N_i, F_i(\Delta_i))$ .*

**Theorem 2.1.** *Under this setup of the constant-stress test with time constraint, the Fisher information matrix is*

$$\mathbf{I}_n(\beta_0, \beta_1) = 2n \begin{pmatrix} \sum_{i=1}^k A_i & \sum_{i=1}^k A_i x_i \\ \sum_{i=1}^k A_i x_i & \sum_{i=1}^k A_i x_i^2 \end{pmatrix} \quad (8)$$

where

$$A_i = \pi_i F_i(\Delta_i), \quad i = 1, 2, \dots, k. \quad (9)$$

## 2.2 $k$ -level step-stress test under type-I censoring

For  $i = 1, 2, \dots, k$ , let us first define  $n_i$  to be the number of units failed at stress level  $x_i$  in time interval  $[\tau_i - 1, \tau_i)$  and  $y_{i,l}$  to be the  $l$ -th ordered failure time of  $n_i$  units at  $x_i$ ,  $l = 1, 2, \dots, n_i$ . Furthermore, let  $N_i$  denote the number of units operating and remaining on test at the start of stress level  $x_i$  (viz.,  $N_i = n - \sum_{j=1}^{i-1} n_j = N_{i-1} - n_{i-1}$ ). Then, a

step-stress test under time constraint proceeds as follows. A total of  $N_1 \equiv n$  test units is initially placed at stress level  $x_1$  and tested until time  $\tau_1$  at which point the stress is changed to  $x_2$ . The test is continued on remaining  $N_2 = n - n_1$  units until time  $\tau_2$  at which the stress is changed to  $x_3$ , and so on. Finally, at time  $\tau_k$ , all the surviving items are withdrawn, thereby terminating the life test. Note that the number of surviving items at time  $\tau_k$  is  $n - \sum_{i=1}^k n_i = N_k - n_k$ . Obviously, when there is no right censoring (viz.,  $\tau_k = \infty$  and  $n_k = N_k$ ), this situation corresponds to the  $k$ -level step-stress testing under complete sampling as a special case.

Now, under the K-H model and using (1) and (2) the pdf and cdf of a test unit are

$$f(y) = \frac{y}{\theta_i^2} e^{-\frac{t^2 - \tau_{i-1}^2}{2\theta_i^2}} \left[ \prod_{j=1}^{i-1} e^{-\frac{\tau_j^2 - \tau_{j-1}^2}{2\theta_j^2}} \right], \quad \tau_{i-1} \leq y \leq \tau_i, \quad (10)$$

$$F(y) = 1 - e^{-\frac{y^2 - \tau_{i-1}^2}{2\theta_i^2}} \left[ \prod_{j=1}^{i-1} e^{-\frac{\tau_j^2 - \tau_{j-1}^2}{2\theta_j^2}} \right], \quad \tau_{i-1} \leq y \leq \tau_i \quad (11)$$

Then, using (10) and (11), the likelihood function of  $\mathbf{n} = (n_1, n_2, \dots, n_k)$  and  $\mathbf{y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_k)$  with  $\mathbf{y}_i = (y_{i,1}, y_{i,2}, \dots, y_{i,n_i})$  is obtained as in (4) where

$$U_i = \sum_{l=1}^{n_i} (y_{i,l}^2 - \tau_{i-1}^2) + (N_i - n_i)(\tau_i^2 - \tau_{i-1}^2) \quad (12)$$

Using (4) and the log-linear link in (3), the log-likelihood function of  $(\beta_0, \beta_1)$  can be written as in (5) and as a result, we obtain the MLEs  $\hat{\beta}_0$  and  $\hat{\beta}_1$  as simultaneous solutions to (6) and (7) with  $U_i$  given in (12).

Just like in the case of constant-stress testing,  $\hat{\beta}_0$  and  $\hat{\beta}_1$  are nonlinear functions of random quantities and hence, inference using these MLEs are based on the asymptotic distributional result that the vector  $(\hat{\beta}_0, \hat{\beta}_1)$  is approximately distributed as a bivariate normal with mean  $(\beta_0, \beta_1)$  and variance-covariance matrix  $\mathbf{I}_n^{-1}(\beta_0, \beta_1)$ . Again, by using the proposition 2.2, we can derive  $\mathbf{I}_n(\beta_0, \beta_1)$  in Theorem 2.2.

**Proposition 2.2.** *The random variable  $n_1$  has a binomial distribution with parameters  $(n, F_1(\tau_1))$ . For  $i = 2, 3, \dots, k$ , given  $n_1, n_2, \dots, n_{i-1}$ , the random variable  $n_i$  has a binomial distribution with parameters  $(N_i, p_i)$  where  $p_i = \frac{F_i(\tau_i) - F_{i-1}(\tau_{i-1})}{1 - F(\tau_{i-1})}$ .*

**Theorem 2.2.** *Under this setup of the step-stress test with time constraint, the Fisher information matrix is as in (8) where*

$$A_i = \left[ 1 - e^{-\frac{\tau_i^2 - \tau_{i-1}^2}{2\theta_i^2}} \right] \left[ \prod_{j=1}^{i-1} e^{-\frac{\tau_j^2 - \tau_{j-1}^2}{2\theta_j^2}} \right], \quad i = 1, 2, \dots, k. \quad (13)$$

### 3 Optimality criteria and existence of optimal design points

In this section, we define different optimality criteria for determining the optimal design points, which then can be used to compare between the multi-level constant-stress test

and step-stress test. For the  $k$ -level constant-stress testing, the focus is to determine the optimal allocation proportions  $\pi^* = (\pi_1^*, \pi_2^*, \dots, \pi_k^*)$  with  $\pi_k^* = 1 - \sum_{i=1}^{k-1} \pi_i^*$  while it is to determine the optimal stress durations  $\Delta^* = (\Delta_1^*, \Delta_2^*, \dots, \Delta_k^*)$  for the  $k$ -level step-stress testing. These objective functions are purely based on the Fisher information matrix  $\mathbf{I}_n(\beta_0, \beta_1)$  derived in the preceding section. Since  $A_i$ s in (9) and (13) are always positive, it eliminates any disconcerting anomalies and ensures a positive determinant of  $\mathbf{I}_n(\beta_0, \beta_1)$  as well as a positive variance function. Therefore, there is no particular restriction on the search region for the global optimal design points in these cases.

### 3.1 C-optimality

In an ALT experiment, researchers often wish to estimate the parameters of interest with maximum precision and minimum variability possible. In both the constant-stress and stepstress settings under consideration here, the parameter of interest is  $\theta_0$  which is essential to calculate the MTTF at the use-condition. For this purpose, we consider an objective function given by

$$\phi = nAV(\log \hat{\theta}_0) = nAV(\hat{\beta}_0 + \hat{\beta}_1 x_0) = nAV(\hat{\beta}_0) = 2 \left( \sum_{i=1}^k A_i x_i^2 \right) \left( \sum_{i=1}^k \sum_{j=1}^k A_i A_j (x_i - x_j)^2 \right). \quad (14)$$

where AV stands for asymptotic variance. The C-optimal design points are the ones that minimize  $\phi$  in (14).

### 3.2 D-optimality

Another optimality criterion often used in planning ALT is based on the determinant of the Fisher information matrix, which equals to the reciprocal of the determinant of the asymptotic variance-covariance matrix. Note that the overall volume of the Wald-type joint confidence region of  $(\beta_0, \beta_1)$  is proportional to  $|\mathbf{I}_n^{-1}(\beta_0, \beta_1)|^{1/2}$  at a fixed level of confidence. Consequently, a larger value of  $\mathbf{I}_n(\beta_0, \beta_1)$  would correspond to a smaller asymptotic joint confidence ellipsoid of  $(\beta_0, \beta_1)$  and thus a higher joint precision of the estimators of  $\beta_0$  and  $\beta_1$ . Motivated by this, our second objective function is simply given by

$$\delta = n^{-2} \mathbf{I}_n(\beta_0, \beta_1) = \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k A_i A_j (x_i - x_j)^2. \quad (15)$$

The D-optimal design points are obtained by maximizing (15) for the maximal joint precision of  $(\hat{\beta}_0, \hat{\beta}_1)$ .

### 3.3 A-optimality

Another optimality criterion considered in this study is based on the trace of the first-order approximation of the variance-covariance matrix of the MLEs. It is identical to

the sum of the diagonal elements of  $\mathbf{I}_n^{-1}(\beta_0, \beta_1)$ . The A-optimality criterion provides an overall measure of the average variance of the parameter estimates and gives the sum of the eigenvalues of the inverse of the Fisher information matrix. The A-optimal design points minimize the objective function defined by

$$a = n \operatorname{tr}(\mathbf{I}_n^{-1}(\beta_0, \beta_1)) = 2 \left( \sum_{i=1}^k A_i (1 + x_i^2) \right) \left( \sum_{i=1}^k \sum_{j=1}^k A_i A_j (x_i - x_j)^2 \right)^{-1}. \quad (16)$$

## 4 Numerical results

A numerical study was conducted in order to investigate the relative efficiency of stepstress testing compared to constant-stress testing and to evaluate it as a function of varying parameters. Although a general  $k$ -level step-stress test under Type-I censoring, for the purpose of optimization, an equal step duration  $\Delta$ , i.e.

$\Delta = \tau_{i+1} - \tau_i$ , for  $i = 1, 2, \dots, k-1$ , was considered in this study. For this proposed, considered  $k = 4$ ,  $x = (0.2, 0.4, 0.7, 1)$  and  $\theta = (0.1, 0.3, 0.5, 0.7)$ . Table 1 presents the values of the optimal step duration  $\Delta_C^*$ ,  $\Delta_D^*$  and  $\Delta_A^*$ . It is observed that  $\Delta_C^* > \Delta_A^* > \Delta_D^*$ .

Using the optimal step durations calculated from Table 1 as the censoring time points at each stress level, the allocation proportions  $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$  were then optimized for a four-level sequential constant-stress test under time constraint. Table 2 presents the values of these optimal allocation proportions  $\pi_C^*$ ,  $\pi_D^*$  and  $\pi_A^*$  along with the corresponding optima of each objective function described in the section 3. From Table 2, it is observed that  $\pi_C^* > \pi_A^* > \pi_D^*$ . It is also interesting to note that except for the first and last stress levels, the C-optimality and the A-optimality do not allocate any test units in the intermediate stress levels.

To formally assess the efficiency of the constant-stress ALT and the step-stress ALT optimal values discussed in this area, pairwise ratios of the optima under each criterion were computed using the results obtained in Tables 1 and 2. For example, the efficiencies of test constant-stress compared to test step-stress under each criterion were computed by  $eff_C = \phi_{SS}^*/\phi_{CS}^*$ ,  $eff_D = \delta_{SS}^*/\delta_{CS}^*$  and  $eff_A = a_{SS}^*/a_{CS}^*$ . The values are tabulated in Table 3 and the number greater than 1 indicates higher efficiency of test constant-stress compared to test step-stress since the objective of the D-optimality is to maximize (15) while the objectives of the C-optimality and the A-optimality are to minimize (14) and (16), respectively. It is observed from Table 3 that comparing between the constant-stress and step-stress tests, the highest efficiency is achieved by the D-optimality. Overall from Table 3, the step-stress test is empirically shown to be more efficient compared to the corresponding constant-stress.

## 5 Conclusions

In this article, the optimal  $k$ -level constant-stress ALT and step-stress ALT were compared for the Rayleigh failure data under Type-I censoring. One of the objectives of this article was to quantify the advantage of using the step-stress ALT relative to the constant-stress ALT. A log-linear relationship was assumed between the scale parameter and stress level,

Table 1: Optimal baseline durations and objective optima for the four-level step-stress test under Type-I censoring.

	C-Optimality	D-Optimality	A-Optimality
Baseline Duration	$\Delta_C^*$ 354.19	$\Delta_D^*$ 2.313	$\Delta_A^*$ 3.243
Optima	$\phi_{ss}^*$ 1.000	$\delta_{ss}^*$ 0.081	$a_{ss}^*$ 16.564

Table 2: Optimal allocation proportions and objective optima for the four-level constant-stress test under Type-I censoring.

	C-Optimality	D-Optimality	A-Optimality
Allocation Proportion	$\pi_C^*$ (0.999,0,0,0.001)	$\pi_D^*$ (0.200,0.312,0,0.488)	$\pi_A^*$ (0.741,0,0,0.259)
Optima	$\phi_{cs}^*$ 4.518	$\delta_{cs}^*$ 0.034	$a_{cs}^*$ 18.886

Table 3: Efficiency of the four-level constant-stress and step-stress tests under Type-I censoring.

Efficiency	C-Optimality	D-Optimality	A-Optimality
Step-Stress vs. Constant-Stress	1.01	2.40	1.12

and the KH model was assumed for the effect of changing stress levels in the step-stress ALT. Based on the information matrix, the optimal design points were determined under C-optimality, D-optimality, and A-optimality criteria. The relative efficiency of the  $k$ -level step-stress ALT compared to the  $k$ -level constant-stress ALT was then measured via pairwise ratios of the optima under each criterion, and some interesting observations were made from the results of the numerical study under a particular setup. It was demonstrated that the step-stress ALT is overall more efficient than the corresponding constant-stress ALT.

## References

- [1] Balakrishnan, N. and Han, D. (2009), Optimal step-stress testing for progressively Type-I censored data from exponential distribution, *Journal of Statistical Planning and Inference*, **139**, 1782–1798.
- [2] Gouno, E., Sen, A. and Balakrishnan, N. (2004), Optimal step-stress test under progressive Type-I censoring, *IEEE Transactions on Reliability*, **53**, 383–393.
- [3] Ng, H.K.T., Balakrishnan, N. and Chan, P.S. (2007), Optimal sample size allocation for tests with multiple levels of stress with extreme value regression, *Naval Research Logistics*, **54**, 237–249.

- [4] Hu, C.H., Plante, R.D. and Tang, J. (2013), Statistical equivalency and optimality of simple step-stress accelerated test plans for the exponential distribution, *Naval Research Logistics*, **60**, 19–30.
- [5] Khamis, I. H. and Higgins, J. J. (1998), A new model for step-stress testing. *IEEE Transactions on Reliability*, **47**, 131–134.
- [6] Nelson, W. (1990), *Accelerated Testing-Statistical Models*, Test Plans and Data Analysis, John Wiley, New York.



## Cumulative residual entropy in sequential order statistics and some characterizations

Hashempour, M. <sup>1</sup>      Doostparast, M. <sup>2</sup>

<sup>1</sup> Department of Statistics, School of Basic Sciences, University of Hormozgan, Bandar Abbas  
<sup>2</sup> Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad

### Abstract

This article considers the problem of characterizing the parent distributions based on the cumulative residual entropy of sequential order statistics under a conditional proportional hazard rates model. It is shown that the equality of the cumulative residual entropy in first sequential order statistics determine uniquely the parent distribution. Subsequently, we characterized the Weibull distribution based on the ratio of the cumulative residual entropy of first sequential order statistics to the mean of the first sequential order statistics. Also, we considered characterizations based on the dynamic cumulative residual entropy and derived some bound for the cumulative residual entropy of residual lifetime of the first sequential order statistics.

**Keywords:** Cumulative residual entropy, Sequential order statistics, Residual lifetime.

## 1 Introduction

Kamps [7] introduced the concept of sequential order statistics (SOS) as an extension of order statistics and used for modelling lifetimes of sequential  $r$ -out-of- $n$  systems. Specifically, consider a given  $r$ -out-of- $n$  system consisting of  $n$  components and  $X_1, \dots, X_n$

<sup>1</sup>ma.hashempour@hormozgan.ac.ir

<sup>2</sup>doustparast@um.ac.ir

denote component lifetimes. The system lifetime coincides to the  $r$ -th order statistics among  $X_1, \dots, X_n$ , denoted by  $X_{r:n}$ . In (usual)  $r$ -out-of- $n$  systems, it is assumed that  $X_1, \dots, X_n$  are independent and identically distributed (i.i.d.) with a common cumulative distribution function (CDF), say  $F$ . Notice that failing a component does not change lifetimes of surviving components; For more details, see, David and Nagaraja [8], Hashempour and Doostparast [5] and references therein. Motivated by Cramer and Kamps [2, 3], in practice, the failure of a component may result in a higher load on remaining components and hence causes the distribution of the surviving components change. In these cases, system lifetimes may be modelled by SOSs. To see this, suppose that  $F_j$ , for  $j = 1, \dots, n$ , denotes the common CDF of the lifetime components when  $n - j + 1$  components are working. Components begin to work at time  $t = 0$  independently with the common CDF  $F_1$ . When at time  $x_1$ , the first failure occurs, the remaining  $n - 1$  components work independently with the common CDF  $F_2$ . This process continues to  $n - r + 1$  components independently with the common CDF  $F_r$  work until the  $r$ -th failure occurs at time  $x_r$  and hence the whole system fails. The mentioned system is called *sequential  $r$ -out-of- $n$  system* and the system lifetime coincides to the  $r$ -th component failure time, denoted by  $X_{(r)}^*$ . In the literature,  $(X_{(1)}^*, \dots, X_{(n)}^*)$  is called SOSs. The problem of estimating parameters on the basis of SOS has been considered in the literature. For example, Cramer and Kamps [2] considered the problem of estimating the parameters on the basis of  $s$  independent SOSs samples under a conditional proportional hazard rates (CPHR) model, defined by  $\bar{F}_j(t) = \bar{F}_0^{\alpha_j}(t)$  for  $j = 1, \dots, r$ . In this case, the hazard rate function of the CDF  $F_j$ , defined by  $h_j(t) = f_j(t)/\bar{F}_j(t)$  for  $t > 0$  and  $j = 1, \dots, n$ , is proportional to the hazard rate function of the baseline CDF  $F_0$ , i.e.  $h_j(t) = \alpha_j h_0(t)$  where  $\alpha_j > 0$ ; See, also Cramer and Kamps [2, 3], and Hashempour and Doostparast [5, 6].

Shannon [5] introduced the concept of entropy which is widely used in the fields of physics, probability, statistics, communication theory, information theory, economics, and so forth. In information theory, entropy is a measure of the uncertainty associated with a random variable. Shannon entropy represents an absolute limit on the best possible lossless compression of any communication. Shannon entropy of a continuous random variable  $X$  with the probability density function (PDF)  $f(x)$  is defined as

$$H(X) = -E(\log f(X)) = - \int_{-\infty}^{+\infty} f(x) \log f(x) dx,$$

where “ $\log$ ” means the natural logarithm. In this paper, we suppose  $X$  is a positive and continuous random variable. We focus our attention on the cumulative residual entropy (CRE) introduced by Rao et al. [9]. It is defined by

$$\begin{aligned} \text{CRE}(X) &= - \int_0^{+\infty} \bar{F}(x) \log \bar{F}(x) dx \\ &= - \int_0^1 \frac{u \log u}{f(F^{-1}(1-u))} du. \end{aligned} \tag{1}$$

where  $\bar{F}(x) = 1 - F(x)$  is the survival function of  $X$ . Rao et al. [9] showed that it is more general than Shannon entropy and possesses more general mathematical properties

than Shannon entropy. The CRE can be easily estimated from data [6] and its estimation asymptotically converges to the true value. CRE has applications in reliability engineering and computer vision; For more details, see Rao [6]. Asadi and Zohrevand [4] defined the dynamic measure of CRE and obtained some of its properties. The dynamic CRE for the residual lifetime distribution of a system is

$$\text{DCRE}(X_t) = - \int_t^\infty \bar{F}_t(x) \log \bar{F}_t(x) dx, \quad (2)$$

where  $\bar{F}_t(x) = P(X - t > x | X > t) = \bar{F}(x+t)/\bar{F}(t)$  and  $\bar{F}(x) = 1 - F(x)$  is the survival function (SF). The DCRE is a measure of the information in the residual life distribution. The authors show that the CRE and the DCRE is connected with some well-known reliability measures such as the mean residual lifetime and the hazard rate. Also, they prove that if the DCRE is an non-decreasing function on  $t$  then it characterizes the underlying distribution function uniquely. We also tackled DCRE of SOS.

## 2 Main results

Let  $X_{1:n}^*$  be the first sequential order statistic under a CPHR model of size  $n$  from a positive and continuous random variable  $X$  with the CDF  $F$  and the PDF  $f$ . Then, the CDF of  $X_{1:n}^*$  is

$$F_{X_{1:n}^*}(x) = 1 - \bar{F}^{n\alpha_1}(x), \quad \forall x \in \mathbb{R}.$$

Thus,

$$\text{CRE}(X_{1:n}^*) = -n\alpha_1 \int_0^1 \frac{u^{n\alpha_1} \log u}{f(F^{-1}(1-u))} du. \quad (3)$$

Now we consider three cases:

### 2.1 Exponential model

A random variable  $X$  follows the exponential distribution, if

$$F(x) = 1 - \exp\{-\lambda x\}, \quad x > 0, \quad \lambda > 0, \quad (4)$$

where  $\lambda$  is the scale parameter. Notice that  $E(X) = \lambda^{-1}$  and  $\text{CRE}(X) = \lambda^{-1}$ . Thus,

$$\Lambda_1 = \frac{\text{CRE}(X)}{E(X)} = 1. \quad (5)$$

Substituting Equation (4) into Equation (3), we have

$$\begin{aligned} \text{CRE}(X_{1:n}^*) &= \frac{n\alpha_1}{\lambda} \int_0^1 u^{n\alpha_1-1} (-\log u) du \\ &= (n\alpha_1 \lambda)^{-1}. \end{aligned} \quad (6)$$

since  $E(X_{1:n}^*) = (n\alpha_1\lambda)^{-1}$ , Then

$$\Lambda_1^* = \frac{\text{CRE}(X_{1:n}^*)}{E(X_{1:n}^*)} = 1. \quad (7)$$

For all  $n$ ,

- the ratio (7) is constant Under the CPHR model with a baseline exponential distribution.
- $\Phi_1 = \text{CRE}(X) - \text{CRE}(X_{1:n}^*) > 0$ , that means uncertainty of  $X$  is more than  $X_{1:n}^*$ .
- Notice that

$$\frac{\text{CRE}(X)}{\text{CRE}(X_{1:n}^*)} = n\alpha_1, \quad (8)$$

does not depend on  $\lambda$ . A plot of  $\text{CRE}(X)/\text{CRE}(X_{1:n}^*)$  is show in Figure red...

## 2.2 Pareto model

Here, suppose that  $X$  has the Pareto distribution, with shape parameter  $\gamma > 0$  and scale parameter  $\beta > 0$ ; that is,  $F(x) = 1 - (\beta/x)^\gamma$  for  $x \geq \beta$ , and denoted by  $P(\gamma, \beta)$ . From Equations (1) and (3), the  $\text{CRE}(X)$  and the  $\text{CRE}(X_{1:n}^*)$  are derived as

$$\text{CRE}(X) = \frac{\gamma\beta}{(\gamma - 1)^2}, \quad \gamma > 1,$$

and

$$\text{CRE}(X_{1:n}^*) = \frac{n\alpha_1\gamma\beta}{(n\alpha_1\gamma - 1)^2}, \quad \gamma > \frac{1}{n\alpha_1},$$

and otherwise CREs are infinity. For  $\gamma > 1$ ,  $\Phi_2 = \text{CRE}(X) - \text{CRE}(X_{1:n}^*) \geq 0$ , then uncertainty of  $X$  is more than  $X_{1:n}^*$ .  $\Phi_2$  is an increasing function of  $n$  for  $n > (\gamma\alpha_1)^{-1}$ . Notice that

$$\frac{\text{CRE}(X)}{\text{CRE}(X_{1:n}^*)} = \frac{(n\alpha_1\gamma - 1)^2}{n\alpha_1(\gamma - 1)^2}, \quad (9)$$

does not depend on  $\beta$ . The following lemma is used in the rest of this section. For any increasing sequence of positive integers  $\{m_i, i \geq 1\}$ ,  $\sum_{i=1}^{\infty} m_i^{-1}$  is infinite, if and only if the sequence of polynomials  $\{x^{m_i}\}$  is complete on  $L(0, 1)$ .

### 2.3 Weibull model

A random variable  $X$  follows the Weibull distributed, if its CDF is

$$F(x) = 1 - \exp\{-(x\lambda)^q\}, \quad x > 0, \quad q > 0, \quad \lambda > 0,$$

where  $q$  and  $\lambda$  are shape and scale parameters, respectively. One can show that  $E(X) = \lambda^{-1}\Gamma(1+q^{-1})$  and  $\text{CRE}(X) = (\lambda q)^{-1}\Gamma(1+q^{-1})$ , where  $\Gamma(x) = \int_0^\infty t^{x-1} \exp\{-t\} dt$  stands for the complete gamma function. Thus,  $\text{CRE}(X)/E(X) = q^{-1}$ .

**Theorem 2.1.** *Suppose that  $X_1, \dots, X_n$  are independent and identically distributed observations from an absolutely continuous CDF  $F(x)$  and PDF  $f(x)$ . Then  $F$  belongs to the Weibull family, if and only if  $\text{CRE}(X_{1:n}^*)/E(X_{1:n}^*) = k$ , for all  $n = n_j$ ,  $j \geq 1$ , such that  $\sum_{j=1}^{+\infty} n_j^{-1}$  is infinite.*

For all  $n$ ,

$\Phi_3 = \text{CRE}(X) - \text{CRE}(X_{1:n}^*) = (1 - (n\alpha_1)^{\frac{1}{q}})(\lambda q)^{-1}\Gamma(1 + q^{-1}) \geq 0$ , then uncertainty of  $X$  is more than  $X_{1:n}^*$  and is increasing in  $n$ . Notice that

$$\frac{\text{CRE}(X)}{\text{CRE}(X_{1:n}^*)} = (n\alpha_1)^{\frac{1}{q}}. \quad (10)$$

does not depend on  $\lambda$ .

**Theorem 2.2.** *Suppose  $Y$  and  $Z$  be two positive random variables with PDFs  $f(x)$  and  $g(x)$  and absolutely continuous CDFs  $F(x)$  and  $G(x)$ , respectively. Then  $\text{CRE}(Y_{1:n}^*) = \text{CRE}(Z_{1:n}^*)$ , for  $j \geq 1$ ,  $n = n_j$ , such that  $\sum_{j=1}^{+\infty} n_j^{-1} = \infty$ , if and only if  $F$  and  $G$  belong to the same family of distributions, but for a possible location shift.*

The CRE for the residual lifetime distribution of  $X_{1:n}^*$ , that is  $\text{DCRE}(X_{1:n,t}^*)$  is

$$\begin{aligned} \text{CRE}(X_{1:n,t}^*) &= -(\bar{F}(t))^{-n\alpha_1} \int_t^\infty (\bar{F}(x))^{n\alpha_1} \log(\bar{F}(x))^{n\alpha_1} dx \\ &\quad + n\alpha_1 M_{X_{1:n}^*}(t) \log \bar{F}(t), \end{aligned} \quad (11)$$

where  $M_{X_{1:n}^*}(t) = E(X_{1:n}^* - t | X_{1:n}^* > t)$  is the mean residual lifetime of system.

For  $t = 0$ ,  $\text{CRE}(X_{1:n,0}^*) = \text{CRE}(X_{1:n}^*)$ . In sequel, we provide a lower bounds for the  $\text{DCRE}(X_{1:n,t}^*)$ .

**Proposition 2.1.** *For all  $t$ ,*

$$n\alpha_1 M_{X_{1:n}^*}(t) |\log \bar{F}(t)| (\bar{F}(t))^{n\alpha_1} \leq \text{CRE}(X_{1:n}^*). \quad (12)$$

At the end of this section,

**Proposition 2.2.** *Suppose that  $Y$  and  $Z$  be two positive random variables with PDFs  $g(x)$  and  $h(x)$  and absolutely continuous CDFs  $G(x)$  and  $H(x)$ , respectively. Then  $G$  and  $H$  belong to the same family of distributions, but for a change in location and scale, if and only if for  $t > 0$ ,  $\text{CRE}(Y_{1:n,t}) = \text{CRE}(Z_{1:n,t})$ , for  $n = n_j$ ,  $j \geq 1$  such that  $\sum_{j=1}^{+\infty} n_j^{-1}$  is infinite.*

## References

- [1] Asadi, M. and Zohrevand, Y. (2007), On the dynamic cumulative residual entropy, *Journal of Statistical Planning and Inference*, **137**, 931-1941.
- [2] Cramer, E. and Kamps, U. (1996), Sequential order statistics and k-out-of-n systems with sequentially adjusted failure rates, *Annals of the Institute of Statistical Mathematics*, **48**(3), 535-549.
- [3] Cramer, E. and Kamps, U. (2001a), Estimation with Sequential Order Statistics from exponential distributions, *Annals of the Institute of Statistical Mathematics*, **53**(2), 307-324.
- [4] David, H. A. and Nagaraja, H. N. (2003), *Order Statistics*, 3rd edition. John Wiley, New York.
- [5] Hashempour, M. and Doostparast, M. (2016), Bayesian inference on multiply sequential order statistics from heterogeneous exponential populations with GLR test for homogeneity, *Communications in Statistics-Theory and Methods*, **46**(16), 8086-8100.
- [6] Hashempour, M. and Doostparast, M. (2016), Statistical evidences in sequential order statistics arising from a general family of lifetime distributions, *Journal of the Turkish Statistical Association*, **9**(1), 29-41.
- [7] Kamps, U. (1995b), A concept of generalized order statistics, *Journal of Statistical Planning and Inference*, **48**, 1-23.
- [8] Rao, M. (2005), More on a new concept of entropy and information. *Journal Theoretical Probability*, **18**, 967-981.
- [9] Rao, M., Chen, Y., Vemuri, B. C. and Wang, F. (2004), Cumulative residual entropy: a new measure of information, *IEEE Transactions on Information Theory*, **6**, 1220-1228.
- [10] Shannon, C. E. (1948), A mathematical theory of communication, *Bell Syst. Tec. J*, **27**, 379-432.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## Some results on proportional mean past lifetime frailty model

Hooti, F.<sup>1</sup>      Ahmadi, J.<sup>2</sup>

<sup>1,2</sup> Department of Statistics, Ferdowsi University of Mashhad

### **Abstract**

The general proportional mean past lifetime frailty model is introduced and some of its properties are studied. The unconditional cumulative distribution and density functions of the lifetime variable are derived. Also, some dependency concepts between the two variables are investigated.

**Keywords:** Frailty model, PMPL model, Stochastically increasing (decreasing), Totally positive of order 2, Reverse regular of order 2.

## 1 Introduction

In the literature, several models have been introduced for modeling and analyzing failure time data. Vaupel et al. (1979) introduced the term frailty, it can be attributed to a person who is more vulnerable than others. They used it in order to capture the difference between individuals at risks, even when they appear to have height, weight, age, etc. to be the same. Frailty models are extensively used in survival analysis to account for unobserved heterogeneity in individual risks to disease and death. This model is a random effect model for time to event data and takes into account that the population is not homogeneous. Heterogeneity is usually explained by covariates, but when important covariates have not been observed, this leads to unobserved heterogeneity. For more details, we refer the readers to the books by, Duchateau and Janssen (2008), Hanagal

---

<sup>1</sup>hooti.fatemeh@gmail.com

<sup>2</sup>ahmadi-j@um.ac.ir

(2011) and Wienke (2011) and the references therein. In survival analysis, Cox proportional hazard (PH) model is one of the most popular statistical models which is used to investigate the relation between the survival time and covariates. The PH frailty model is given by  $r(t|v) = vr_0(t)$ ,  $t > 0$ , where  $r_0(t)$  is the baseline hazard rate and does not depend on  $v$ , with  $v$  being the unobserved heterogeneity, known as frailty. Gupta and Gupta (2009) proposed a general frailty model and obtained some results for stochastic comparisons, and these results have been used to obtain some properties of the PH frailty model and the additive frailty model. Rezaei and Gholizadeh (2015) introduced mixture proportional mean residual life model and investigated dependence structure and various stochastic comparisons for this model. Kayid et al. (2016) considered general proportional mean residual life model and established many characterizations and preservation properties of this model under different stochastic orders and aging classes.

The rest of this paper is organized as follows. The proportional mean past lifetime (PMPL) model is reminded in Section 2. In Section 3, the general proportional mean past lifetime (GPMPL) frailty model is introduced then some distributional properties of this introduced model and also dependency between lifetime variable and frailty random variable are discussed. In Section 3, some concluding comments are made.

## 2 PMPL model and some dependency concepts

The mean past lifetime (MPL) measures the expected value of the elapsed lifetime of the system, given that it has failed at or by time  $t$ ,  $t > 0$ . Note that the mean past lifetime is also called the reversed mean residual life or the mean inactivity time in the literature. The MPL function of a non-negative random variable  $X$  with cumulative distribution function (cdf)  $F$  is given by

$$\begin{aligned} m^*(x) &= E(x - X | X \leq x) \\ &= \frac{\int_0^x F(u)du}{F(x)}, \quad x > 0. \end{aligned}$$

The definition and some properties of MPL have been given by Nanda et al. (2003). Izadkhah et al. (2013) obtained some properties of MPL order for weighted distributions.

Suppose  $T_1$  and  $T_2$  are two non-negative random variables with MPL functions  $m_1^*(.)$  and  $m_2^*(.)$ , respectively. Asadi and Berred (2012) proposed the proportional mean past lifetime (PMPL) model as

$$m_2^*(t) = cm_1^*(t). \quad (1)$$

Rezaei (2016) studied some reliability properties and implications of stochastic orders of random variables possessing the PMPL model. He then investigated the mixture model of PMPL and the dependency structure of these random variables. Let  $F_1(.)$  and  $F_2(.)$  be the cdf's,  $f_1(.)$  and  $f_2(.)$  be probability distribution functions (pdf), and  $r_1^*(.)$  and  $r_2^*(.)$  be reversed hazard rate functions of random variables  $T_1$  and  $T_2$ , respectively. Further, let  $S_{T_i} = (a_i, b_i)$  be the support of  $T_i$ , where  $a_i = \inf\{t : F_i(t) > 0\}$  and  $b_i = \sup\{t : F_i(t) < 1\}$ , for  $i = 1, 2$ . Then, Rezaei (2016) showed (see his Theorem 2.1) that if (1) holds and  $b_1 = b_2 = b$ , then

$$(i) \quad F_2(t) = F_1(t) \left( \frac{\int_0^t F_1(x)dx}{\int_0^b F_1(x)dx} \right)^{\frac{1}{c}-1};$$

$$(ii) \quad f_2(t) = \left( \frac{\int_0^t F_1(x)dx}{\int_0^b F_1(x)dx} \right)^{\frac{1}{c}-2} \left[ f_1(t) \frac{\int_0^t F_1(x)dx}{\int_0^b F_1(x)dx} + \frac{F_1^2(t)}{\int_0^b F_1(x)dx} \left( \frac{1}{c} - 1 \right) \right];$$

$$(iii) \quad r_2^*(t) = \frac{1}{c} \left( c r_1^*(t) + \frac{1-c}{m_1^*(t)} \right).$$

The purpose of this paper is to study the GPMPL frailty model and discuss various dependency concepts from this model. We first present some definitions and notions of dependency that will be used throughout this paper.

**Definition 2.1.** Let  $g(.,.)$  be a non-negative function defined on  $[0, \infty) \times [0, \infty)$ . Then,  $g$  is said to be totally positive of order 2 (TP2) (reverse regular of order 2 (RR2)) if for all  $0 \leq x_1 \leq x_2$  and  $0 \leq y_1 \leq y_2$ , we have

$$g(x_1, y_1) g(x_2, y_2) \geq (\leq) g(x_1, y_2) g(x_2, y_1).$$

This is equivalent to

(i)  $\frac{g(x, y_2)}{g(x, y_1)}$  is increasing (decreasing) in  $x \geq 0$  for all  $0 \leq y_1 \leq y_2$ ;

(ii)  $\frac{g(x|y_2)}{g(x|y_1)}$  is increasing (decreasing) in  $x \geq 0$  for all  $0 \leq y_1 \leq y_2$ ;

(iii)  $\frac{\partial^2 \ln g(x, y)}{\partial x \partial y} \geq (\leq) 0$ .

**Definition 2.2.** Assume that  $F(.,.)$ ,  $\bar{F}(.,.)$  and  $f(.,.)$  are the joint cdf, joint survival function and joint pdf of  $(X, Y)$ , respectively. Then:

1. The random variable  $Y$  is said to be stochastically increasing (decreasing) in  $X$ , denoted by  $SI(Y|X)$  ( $SD(Y|X)$ ), if  $P(Y \leq y|X = x)$  is non-increasing (non-decreasing) in  $x$ ;
2. The random variable  $Y$  is said to be left-tail decreasing (increasing) in  $X$ , denoted by  $LTD(Y|X)$  ( $LTI(Y|X)$ ), if  $P(Y < y|X < x) = \frac{F(x, y)}{F(x)}$  is non-increasing (non-decreasing) in  $x$ ;
3. Two random variables  $X$  and  $Y$  are said to be left-corner set decreasing (increasing), denoted by  $LCSD(X, Y)$  ( $LCSI(X, Y)$ ), iff for all  $x, x', y$  and  $y'$ ,  

$$P(X \leq x, Y \leq y|X \leq x', Y \leq y')$$
  
is decreasing (increasing) in  $x', y'$ .

The following relationships are known:

$$LCSI(X, Y) \implies LTI(Y|X) \text{ and } LTI(X|Y);$$

$$LCSD(X, Y) \implies LTD(Y|X) \text{ and } LTD(X|Y);$$

$$LCSD(X, Y) \iff F(x, y) \text{ is TP2};$$

$$LCSI(X, Y) \iff F(x, y) \text{ is RR2}.$$

For more details, we refer the readers to Nelsen (2006).

### 3 GPMPL frailty model and some distributional properties

Let  $T$  be the lifetime of an individual and  $V$  be the frailty variable. The conditional mean past lifetime function for a given frailty variable,  $V = v$ , at time  $t \geq 0$ , is given by

$$m^*(t|v) = \alpha(v)m_0^*(t), \quad (2)$$

where  $m_0^*(t)$  is the mean past lifetime function of the baseline random variable  $T_0$  at time  $t \geq 0$ , and  $\alpha(\cdot)$  is a positive function of the frailty random variable. Let  $F_0(\cdot)$  and  $S_{T_0} = (a, b)$  be the cdf and the support of the baseline random variable  $T_0$ . Then, the conditional cdf, for given frailty at time  $t \geq 0$ ,  $F(t|v)$ , is given by

$$F(t|v) = F_0(t) \left( \frac{\int_0^t F_0(x)dx}{\int_0^b F_0(x)dx} \right)^{\frac{1}{\alpha(v)} - 1}. \quad (3)$$

The corresponding conditional pdf is

$$f(t|v) = F_0(t) \left( \frac{\int_0^t F_0(x)dx}{\int_0^b F_0(x)dx} \right)^{\frac{1}{\alpha(v)} - 1} \left[ r_0^*(t) + \frac{1}{m_0^*(t)} \left( \frac{1}{\alpha(v)} - 1 \right) \right], \quad (4)$$

where  $r_0^*(t)$  is the baseline reversed hazard rate function. Consequently, we obtain the following expression for the conditional reversed hazard rate:

$$r^*(t|v) = r_0^*(t) + \left( \frac{1}{\alpha(v)} - 1 \right) \frac{1}{m_0^*(t)}. \quad (5)$$

The unconditional cdf is simply

$$F(t) = \frac{\int_0^b F_0(x)dx}{m_0^*(t)} E \left[ \left( \frac{\int_0^t F_0(x)dx}{\int_0^b F_0(x)dx} \right)^{\frac{1}{\alpha(V)}} \right], \quad (6)$$

where  $h(\cdot)$  is the pdf of the random variable  $V$ . Similarly, the unconditional pdf, in terms of  $m_0^*(t)$ , is

$$\begin{aligned} f(t) &= \frac{\int_0^b F_0(x)dx}{m_0^*(t)} \left\{ r_0^*(t) E \left[ \left( \frac{\int_0^t F_0(x)dx}{\int_0^b F_0(x)dx} \right)^{\frac{1}{\alpha(V)}} \right] \right. \\ &\quad \left. + \frac{1}{m_0^*(t)} E \left[ \left( \frac{1}{\alpha(V)} - 1 \right) \left( \frac{\int_0^t F_0(x)dx}{\int_0^b F_0(x)dx} \right)^{\frac{1}{\alpha(V)}} \right] \right\}. \end{aligned} \quad (7)$$

From the above equations, the population level reversed hazard rate function can be expressed as

$$r^*(t) = r_0^*(t) + \frac{1}{m_0^*(t)} \left[ E \left( \frac{1}{\alpha(V)} \middle| T \leq t \right) - 1 \right]. \quad (8)$$

Note that

$$\begin{aligned} h(v|T \leq t) &= \frac{F(t|v) h(v)}{F(t)} \\ &= \frac{\left( \frac{\int_0^t F_0(x)dx}{\int_0^b F_0(x)dx} \right)^{\frac{1}{\alpha(v)}} h(v)}{\int_0^\infty \left( \frac{\int_0^t F_0(x)dx}{\int_0^b F_0(x)dx} \right)^{\frac{1}{\alpha(u)}} h(u)du}. \end{aligned} \quad (9)$$

Now, with these conditional and unconditional distributions and reversed hazard rate functions, and by the use of (3), we obtain the following theorem.

**Theorem 3.1.** *Assume that the GPMPL frailty model (2) holds. If  $\alpha(v)$  is increasing (decreasing) in  $v$ , then we have:*

- (i)  $SD(T|V)$  ( $SI(T|V)$ );
- (ii)  $LCSI(T|V)$  ( $LCSD(T|V)$ ).

*Proof.* (i) From (3), for all  $0 \leq v_1 < v_2$ , we have

$$F(t|v_2) - F(t|v_1) = \frac{\int_0^b F_0(x)dx}{m_0^*(t)} \left[ \left( \frac{\int_0^t F_0(x)dx}{\int_0^b F_0(x)dx} \right)^{\frac{1}{\alpha(v_2)}} - \left( \frac{\int_0^t F_0(x)dx}{\int_0^b F_0(x)dx} \right)^{\frac{1}{\alpha(v_1)}} \right]. \quad (10)$$

The expression within the brackets in (10) is positive (negative) if  $\alpha(v)$  is increasing (decreasing).

(ii) It is suffices to show that the joint distribution function of  $T$  and  $V$ ,  $F(t, v)$ , is  $RR2(TP2)$ . To this end, we can write

$$\begin{aligned} F(t, v) &= \int_0^v P(T \leq t | V = u) h(u) du \\ &= \frac{\int_0^b F_0(x)dx}{m_0^*(t)} \int_0^\infty \phi(t, u) \varphi(u, v) du, \end{aligned} \quad (11)$$

where  $\phi(t, u) = \left( \frac{\int_0^t F_0(x)dx}{\int_0^b F_0(x)dx} \right)^{\frac{1}{\alpha(u)}}$  and  $\varphi(u, v) = I_{(0,v)}(u)h(u)$ .

It is easy to show that if  $\alpha(u)$  is increasing (decreasing) in  $u \in S_V$ , then  $\phi(t, u)$  is RR2 (TP2) in  $(t, u) \in \mathbb{R}^+ \times \mathbb{R}^+$ . In addition,  $\varphi(u, v)$  is TP2 in  $(u, v) \in S_V \times S_V$ . Therefore, by using Lemma 1.1 of Karlin (1968, page 99) in (11), the proof gets completed.  $\square$

From Theorem 1, we readily obtain the following corollary. Under the assumption of Theorem 1, if  $\alpha(v)$  is increasing (decreasing) in  $v$ , then:

- (i)  $F(t|V \leq v)$  is increasing (decreasing) in  $v > 0$  and  $H(v|T \leq t)$  is increasing (decreasing) in  $t$ , for  $t > 0$ ;
- (ii)  $E(T|V \leq v)$  is a decreasing (increasing) function of  $v$  and  $E(V|T \leq t)$  is a decreasing (increasing) function of  $t$ .

## 4 Conclusions

In the literature, several models have been introduced for modeling and analyzing failure time data. Such as the proportional hazard (PH) rate model, the proportional reversed hazard (PRH) rate model, the additive hazard model, the additive reversed hazard rate model, proportional mean residual life model and frailty model are all good examples. Frailty models are extensively used in survival analysis to account for unobserved heterogeneity in individual risks to disease and death. This model is a random effect model for time to event data and takes into account that the population is not homogeneous. In this paper, we introduced the GPMPL frailty model. Some distributional properties of the proposed model are studied. Based on this model, dependency between the two variables are studied.

## References

- [1] Asadi, M. and Berred, A. (2012). Properties and estimation of the mean past lifetime. *Statistics*, **46**, 405–417.
- [2] Duchateau, L. and Janssen, P. (2008). *The Frailty Model*. Springer, New York.
- [3] Gupta, R. C. and Gupta, R. D. (2009). General frailty model and stochastic orderings. *Journal of Statistical Planning and Inference*, **139**, 3277–3287.
- [4] Hanagal, D. D. (2011). *Modeling Survival Data Using Frailty Models*. Chapman and Hall/ CRC Press, New York.
- [5] Izadkhah, S., Rezaei, R. A. H. and Mohtashami B. G. R. (2013). On properties of reversed mean residual life order for weighted distributions. *Communications in Statistics-Theory and Methods*, **42**, 838–851.
- [6] Kayid, M., Izadkhah, S. and Almufarrej, D. (2016). General proportional mean residual life model. *Applications of Mathematics*, **61**, 607–622.

- [7] Nanda, A. K., Singh, H., Misra, N. and Paul, P. (2003). Reliability properties of reversed residual lifetime. *Communications in Statistics-Theory and Methods*, **32**, 2031–2042.
- [8] Nelsen, R. B. (2006). *An Introduction to Copulas*. Springer, New York.
- [9] Rezaei, M. (2016). On proportional mean past lifetimes model. *Communications in Statistics-Theory and Methods*, **45**, 4035–4047.
- [10] Rezaei, M. and Gholizadeh, B. (2015). On mixture proportional mean residual life model. *Communications in Statistics-Theory and Methods*, **44**, 4263–4277.
- [11] Vaupel, J. W., Manton, K. G. and Stallard, E. (1979). The impact of heterogeneity on individual frailty on the dynamic of mortality. *Demography*, **16**, 439–454.
- [12] Wienke, A. (2011). *Frailty Models in Survival Analysis*. Chapman and Hall/ CRC Press, New York.



*Forth Seminar on*  
**Reliability Theory and its Applications**

*April 25-26, 2018*



## On mean time to failure in age replacement

Izadi, M. <sup>1</sup>      Khaledi, B. <sup>2</sup>

<sup>1</sup> Department of Statistics, Razi University, Kermanshah

### Abstract

In-service, failure of a unit can sometimes lead to costly or hazardous consequences. Age replacement policy is the most common maintenance policy to prevent a unit from the failure during operation. In age replacement policy, we replace the item either at failure time or at the pre-specified time  $t$  if it is active at time  $t$ . The mean of the first in-service failure time of an item under the age replacement policy, known as mean time to failure (MTTF) in age replacement, is used to evaluate the performance and effectiveness of the age replacement policy. Based on the MTTF function, we introduce two new non-parametric classes of lifetime distributions with non-monotonic mean time to failure in age replacement; increasing then decreasing MTTF (IDMTTF) and decreasing then increasing MTTF (DIMTTF). We characterize the IDMTTF and DIMTTF classes of distributions in terms of the scaled TTT-transform function. The implications between these classes of distributions and some existing classes of non-monotonic aging classes are studied.

**Keywords:** Age replacement, bathtub failure rate, non-monotonic aging class, total time on test transform.

### Extended abstract

In-service failure of a unit can sometimes lead to costly or hazardous consequences. Age replacement policy is the most common maintenance policy to prevent a unit from the failure during operation. In age replacement policy, we replace the item either at failure

---

<sup>1</sup>m.izadi@razi.ac.ir

<sup>2</sup>bkhaledi@hotmail.com

time or at time  $t$  if it is active at time  $t$  (known as planned replacement age). Let  $X_{[t]}$  stand for the first in-service failure time of an item under the age replacement policy with the planned replacement time  $t$ . The mean of  $X_{[t]}$ , known as the mean time to failure (MTTF) in age replacement, has been offered by Barlow and Proschan [1] to evaluate the performance and usefulness of the age replacement policy. Assuming that  $F$  is the lifetime distribution of a new item, the mean of  $X_{[t]}$  (denoted by  $M_F(t)$ ) is given by

$$M_F(t) = \frac{\int_0^t \bar{F}(x)dx}{F(t)}, \quad t > 0 \quad (1)$$

where  $\bar{F} = 1 - F$ . Based on the MTTF function, we introduce two new non-parametric classes of lifetime distributions with non-monotonic mean time to failure in age replacement; increasing then decreasing MTTF (IDMTTF) and decreasing then increasing MTTF (DIMTTF). We characterized these classes of distribution in terms of the scaled TTT-transform function. The bathtub (upside-down bathtub) failure rate (BTF (UBTF)) and new worse (better) then better (worse) than used in the expectation (NWBUE (NBWUE)) are well known non-monotonic aging classes which have been extensively studied in the literature (see [2]). The implications between the IDMTTF and DIMTTF classes of distributions and the BTF, UBTF, NWBUE and NBWUE are studied. It is shown that

$$\text{BFR} \Rightarrow \text{IDMTTF} \Rightarrow \text{NWBUE} \text{ and } \text{UBFR} \Rightarrow \text{DIMTTF} \Rightarrow \text{NBWUE}.$$

The trend change point of the MTTF function in the IDMTTF class of distributions is a possible optimal age replacement time. Thus, we consider the estimation problem of the change point which is of great interest. We estimate the change point and study the asymptotic behaviour of the estimator.

## References

- [1] Barlow, R.E., Proschan, F. (1965). *Mathematical theory of reliability*. Wiley, New York
- [2] Lai, C.D., Xie, M. (2006). *Stochastic ageing and dependence for reliability*. Springer, New York.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## Reliability estimation for the two-parameter exponential distribution based on records

Jafari, A. A.<sup>1</sup>      Bafekri Fadafan, S.<sup>2</sup>

<sup>1,2</sup> Department of Statistics, Yazd University

### Abstract

In this paper, we estimate the reliability function of two-parameter exponential distribution based on record data. For this purpose, we consider different estimators including the maximum likelihood estimator, a bootstrap estimator, a Bayesian estimator and an estimation based on the generalized pivot approach. These estimators are compared by simulation studies, and finally, they are illustrated by using a real data set.

**Keywords:** Maximum likelihood estimate, Bootstrap estimate, Bayesian estimate, generalized pivot variables, record data.

## 1 Introduction

Recently, estimating the reliability function,  $R(t)$ , based on records data is interesting. This type data were first introduced and studied by Chandler [5]. Let  $\{X_n\}_{n=1}^{\infty}$  is a sequence of iid random variables. An observation  $X_j$  will be upper record that  $X_j > X_i$  for every  $i < j$ . Therefore, the  $\{R_n\}_{n=1}^{\infty}$  is a sequence of upper records that defined by

$$R_0 = X_1, \quad R_n = X_{\min j: X_j > R_{n-1}}.$$

One of the reliability topics is estimation of  $R(t)$  that investigated by researchers based on record and censored data. Gui [8] estimated the reliability function for two parameter

---

<sup>1</sup>aa.jafari@yazd.ac.ir

<sup>2</sup>s.bafekri@gmail.com

exponential family under progressively type-II censored data. Baklizi [4] obtained confidence intervals for the reliability function in the two-parameter exponential distribution with Type-II censored data. Let  $X$  has the two-parameter exponential  $\text{Exp}(\mu, \sigma)$  with cumulative distribution function as follows:

$$F(x; \mu, \sigma) = \begin{cases} 1 - \exp\left\{-\frac{(x-\mu)}{\sigma}\right\} & x \geq \mu \\ 0 & x < \mu, \end{cases}$$

where  $\mu$  and  $\sigma$  are location and scale parameters, respectively. So, the reliability function is given by

$$R(t) = \begin{cases} \exp\left\{-\frac{(t-\mu)}{\sigma}\right\} & t \geq \mu \\ 1 & t < \mu. \end{cases} \quad (1)$$

## 2 Estimations methods

Suppose  $R_0, \dots, R_n$  are the first  $(n+1)$  upper record values observed from  $\text{Exp}(\mu, \sigma)$ . To estimate  $R(t)$ , we use the following methods.

The likelihood function is

$$L(\mu, \sigma) = \sigma^{-(n+1)} \exp\left\{-\frac{r_n - \mu}{\sigma}\right\}, \quad \mu < r_0 < \dots < r_n.$$

By maximizing the likelihood function, the maximum likelihood estimators (MLEs) of the parameters  $\mu$  and  $\sigma$  are as follows

$$\hat{\mu} = R_0, \quad \hat{\sigma} = \frac{R_n - R_0}{n+1}, \quad (2)$$

respectively. See Baklizi [2].

Therefore, the maximum likelihood estimator of reliability function is

$$\hat{R}(t) = \begin{cases} \exp(-\frac{(t-\hat{\mu})}{\hat{\sigma}}) & t \geq 0 \\ 1 & t < 0. \end{cases}$$

Efron and Tibshirani [6], obtained the bootstrap estimator and confidence intervals for the parameters  $\mu$  and  $\sigma$ . To obtain the bootstrap estimator (Boot.E), we use the below algorithm.

- (i) For given record data sets, comput the MLEs  $\mu$  and  $\sigma$ .
- (ii) Generat  $(n+1)$  upper record values observed from  $\text{Exp}(\hat{\mu}, \hat{\sigma})$ .
- (iii) Calculate the reliability function estimate, using the record data.
- (iv) Repeat steps (ii) and (iii) for  $i = 1, \dots, B$ .
- (v) Bootstrap estimate for  $R(t)$  is  $\hat{R}^*(t) = \frac{1}{B} \sum_{i=1}^B \hat{R}_i^*(t)$ .

The generalized variable approach (GV) for obtain confidence interval of stress-strength reliability in exponential distribution based on records data proposed by Baklizi [2]. Let  $\hat{\mu}_0$  and  $\hat{\sigma}_0$  be the observed values of  $\hat{\mu}$  and  $\hat{\sigma}$ , respectively. The generalized pivot variable  $G_\mu$  for  $\mu$  should satisfy two conditions:

- (1) The value of  $G_\mu$  at  $(\hat{\mu}, \hat{\sigma}) = (\hat{\mu}_0, \hat{\sigma}_0)$  should be  $\mu$ .
- (2) For a given  $\hat{\mu}_0$ , the distribution of  $G_\mu$  should be independent of any unknown parameters.

Noting that  $\frac{2(\hat{\mu} - \mu)}{\sigma} \sim \chi_2^2$  and  $\frac{2(n+1)\hat{\sigma}}{\sigma} \sim \chi_{2n}^2$ , we define the generalized pivot variable for  $\mu$  and  $\sigma$  as below:

$$\begin{aligned} G_\mu &= \hat{\mu}_0 - \frac{\chi_2^2}{\chi_{2n}^2}(R_n - R_0), \\ G_\sigma &= \frac{2(R_n - R_0)}{\chi_{2n}^2}. \end{aligned}$$

Therefore, the generalized pivot variable for  $R(t)$  is defined as:

$$G_{R(t)} = \begin{cases} \exp(-\frac{(t-G_\mu)}{G_\sigma}) & G_\mu \geq 0 \\ 1 & G_\mu < 0. \end{cases}$$

We use the following algorithm for estimating  $R(t)$  based on the generalized variable approach. This algorithm includes the following steps:

- (i) For given record data sets, comput the MLEs  $\mu$  and  $\sigma$ .
- (ii) Generate  $Q \sim \chi_2^2$  and  $W \sim \chi_{2n}^2$ .
- (iii) Compute  $G_\mu$ ,  $G_\sigma$  and  $G_{R(t)}$ .
- (iv) Repeat (ii) and (iii) for  $m$  times and compute  $G_{R(t)}$ .
- (v) Compute the mean of  $G_{R(t)}$ s.

Let  $R_0, \dots, R_n$  be record values observed from the distribution  $\text{Exp}(\mu, \sigma)$ . To obtain Bayesian estimator (BE) of  $R(t)$ , we consider non-information prior distribution of  $\mu$  and the conjugate inverted gamma prior for the corresponding  $\sigma$  parameter

$$\begin{aligned} \pi(\mu) &= 1, \quad -\infty < \mu < \infty, \\ \pi(\sigma) &= \frac{\beta^\alpha}{\Gamma(\alpha)\sigma^{(\alpha+1)}} \exp\left\{-\frac{\beta}{\sigma}\right\}, \quad \sigma > 0, \end{aligned}$$

where  $\alpha$  and  $\beta$  are shape and rate parameters, respectively. Therefore, joint posterior distribution of  $\mu$  and  $\sigma$  given  $\mathbf{r}$  is

$$\pi^*(\mu, \sigma \mid \mathbf{r}) = \frac{(\beta + r_n - r_0)^{n+\alpha}}{\Gamma(n+\alpha)\sigma^{n+\alpha+2}} \exp\left\{-\frac{(\beta + r_n - \mu)}{\sigma}\right\}.$$

See Baklizi [2]. So, the Bayes estimator of  $R(t)$  is given by if  $t < R_0$  is

$$\hat{R}_B(t) = E(R(t) | \mathbf{r}) = 1 - \frac{(\beta + r_n - r_0)^{n+\alpha}}{2(\beta + r_n - t)^{n+\alpha}},$$

and if  $t > R_0$

$$\hat{R}_B(t) = E(R(t) | \mathbf{r}) = \frac{(\beta + r_n - r_0)^{n+\alpha}}{2(t + \beta + r_n - 2r_0)^{n+\alpha}}.$$

### 3 Simulation studies

In this section, we compare the performance of mentioned estimators using Monte Carlo simulation. We consider different values for  $t$  and  $\mu$  and different sample sizes include  $n = 10$ ,  $n = 15$  and  $n = 20$ . Also, we have taken  $(\sigma, \alpha, \beta) = (2, 2, 4)$ . Average estimates (AE) and the mean of square errors MSE of reliability function estimators are represented in Table 1. These studies investigated based on  $N = 1000$  replication. In bootstrap method, we consider  $B = 1000$  and the generalized variable approach  $m = 1000$ .

Table 1: Comparing of different estimators.

$n$	$\mu$	Method	$t = 2$		$t = 5$		$t = 8$	
			AE	MSE	AE	MSE	AE	MSE
10	0	MLE	0.7083	0.1931	0.2446	0.1011	0.0905	0.0382
		Boot.E	0.8823	0.2852	0.4434	0.2087	0.2012	0.0848
		GV	0.4559	0.0696	0.1603	0.0413	0.0660	0.0170
		BE	0.4556	0.0686	0.1598	0.0409	0.0640	0.0167
		MLE	0.8176	0.1516	0.3302	0.1470	0.1103	0.0445
		Boot.E	0.9424	0.2072	0.5396	0.2660	0.2412	0.1033
	0.6	GV	0.5343	0.0552	0.2148	0.0606	0.0785	0.0194
		BE	0.5353	0.0552	0.2147	0.0601	0.0764	0.0190
		MLE	0.7158	0.1908	0.2730	0.1214	0.0866	0.0333
	15	Boot.E	0.8948	0.2941	0.4857	0.2374	0.8876	0.2873
		GV	0.4567	0.0696	0.1742	0.0523	0.4436	0.0670
		BE	0.4570	0.0691	0.1736	0.0519	0.4442	0.0667
		MLE	0.8304	0.1540	0.3314	0.1348	0.1074	0.0432
		Boot.E	0.9526	0.2136	0.5678	0.2146	0.2441	0.1001
		GV	0.5467	0.0591	0.2065	0.0540	0.0728	0.0205
20	0	BE	0.5664	0.0584	0.2059	0.0535	0.0719	0.0204
		MLE	0.7131	0.1899	0.2743	0.1161	0.0833	0.0318
		Boot.E	0.8952	0.2935	0.9492	0.2333	0.2057	0.0774
		GV	0.4546	0.0712	0.5270	0.0562	0.0541	0.0136
		BE	0.4550	0.0707	0.5272	0.0558	0.0534	0.0135
		MLE	0.8425	0.1580	0.3428	0.1445	0.1144	0.0424
	0.6	Boot.E	0.9595	0.2185	0.5836	0.2883	0.2614	0.1091
		GV	0.5577	0.0591	0.2128	0.0598	0.0749	0.0226
		BE	0.5579	0.0586	0.2124	0.0595	0.0744	0.0226

### Discussion and results

From Table 1, it is observed that the GV and BE methods perform better than ML and bootstrap methods. Also, when  $t$  increases, MSE decreases.

## A real-data example

Engelhardt and Bain [7] considered data on the mileages at which 19 military carriers failed in services. The data are given as: 162, 200, 271, 320, 393, 508, 539, 629, 706, 777, 884, 1008, 1101, 1082, 1463, 1603, 1984, 2355, 2880. The different estimations for some values of  $t$ , are represented in Table 2.

Table 2: Different estimations.

Method	$t = 200$	$t = 500$	$t = 800$
MLE	0.7667	0.0941	0.0115
Boot.E	0.9683	0.3008	0.0606
GV	0.3920	0.0598	0.0112
BE	0.3789	0.0481	0.0074

## References

- [1] Baklizi, A. (2008). Estimation of  $Pr(X < Y)$  using record values in the one and two -parameter exponential distributions, *Communications in Statistics-Theory and Methods*, **84**, 37-41.
- [2] Baklizi, A.(2013). Interval estimation of the stress-strength reliability in the two-parameter exponential distribution based on records, *Journal of Statistical Computation and Simulation*, **84**, 37-41.
- [3] Baklizi, A. (2014). Bayesian inference for  $Pr(X < Y)$  in the exponential distribution based on records, *Applied Mathematical Modelling*, **38**, 1698-1709.
- [4] Baklizi, A. (2016). Confidence intervals for the two-parameter exponential reliability with Type-II censored data, *American Journal of Mathematical and Management Sciences*, **35(4)**, 297-308.
- [5] Chandler, K. N. (1952). The distribution and frequency of record values, *Journal of the Royal Statistical Society. Series B (Methodological)*, 220-228.
- [6] Efron, B., Tibshirani, R. (1993). *An introduction to the bootstrap*, London, UK: Chapman and Hall.
- [7] Engelhardt, M., and Bain, L. J. (1978). Tolerance limits and confidence limits on reliability for the two-parameter exponential distribution, *Technometrics*, **20(1)**, 37-39.
- [8] Gui, W. (2017). Reliability estimation for the two-parameter exponential family based on progressively Type-II censored data, *International Journal of Reliability, Quality and Safety Engineering*, 1850005.



## Multi-objective redundancy allocation problem with entropy constraint using two approaches

Kamandi, R.<sup>1</sup>      Zarezadeh, S.<sup>2</sup>      Mirdehghan, M.<sup>3</sup>

<sup>1,3</sup> Department of Mathematics, Shiraz University

<sup>2</sup> Department of Statistics, Shiraz University

### Abstract

A redundancy allocation problem consists of the selection of a number of redundancies to be allocated in each subsystem of a series-parallel configuration to maximize and minimize reliability and overall system cost, respectively. In this study, two methods based on weighted sum and goal programming approaches are proposed to solve the multi-objective redundancy allocation problem with the entropy constraint. A standard example is presented to show the effectiveness and advantages of the two proposed methods comparing to the existing approaches.

**Keywords:** Reliability, series-parallel system, optimization, goal programming method, weighted-sum method.

## 1 Introduction

Redundancy allocation problem (RAP) includes the use of additional components in a parallel, series-parallel, or parallel-series system to construct a system with the most reliability while the overall system cost is minimized. In the other word, the RAP can be considered as an optimization problem with some criteria and objectives, includes maximizing the system reliability and minimizing the total cost of the system. Chern [3] has

<sup>1</sup>r.kamandi@shirazu.ac.ir

<sup>2</sup>s.zarezadeh@shirazu.ac.ir

<sup>3</sup>mirdehghan@shirazu.ac.ir

proved that the RAP is an NP-hard problem, i.e., it does not solve in a polynomial time. Many authors have studied the RAP using some optimization approaches under different point of view. Fyffe et al. [4], explored the RAP by considering only system reliability using dynamic programming, and integer programming optimization. Salmasnia et al. [8] investigated the RAP including the objectives of the system reliability and its variance as well as the total cost and its variance using loss function approach of multi-objective optimization.

The concept of entropy is a measure of uncertainty in information theory. Since the components are usually affected by external or internal shocks which fall down the system until it does not function, considering the entropy is an essential tool for RAP to have a safe system. In recent years, the entropy of system is used to analyze the system reliability and importance of its components. Rocchi [6] applied the entropy to study the reliability and reparability of system. Mahapatra and Mahapatra [5] used intuitive fuzzy optimization technique for RAP in a system based on entropy.

Roy et al. [7] defined variable  $x_i$  as the number of components in the  $i$ th subsystem of a series-parallel system and determined the optimal values of  $x_i$ 's by maximizing system reliability, minimizing the overall system cost and a constraint on the entropy using genetic algorithm. In the aforementioned work, it is assumed that the components in all subsystems have one type whereas the systems are often designed based on more than one type. This motivates us to define variable  $x_{ij}$  as the number of components of type  $j$  in subsystem  $i$ ,  $i = 1, \dots, s$ ,  $j = 1, \dots, m$ . We explore the RAP based on maximizing the system reliability, minimizing the overall system cost which is the sum of operational cost and maintenance cost and minimizing the cost variance. We also consider some constraints on the total weight (or volume) of system and entropy measure.

Multi-objective optimization is a branch of optimization in which all problems have at least two objective functions restricting to some linear or nonlinear constraints. In multi-objective optimization, usually, there does not exist any solution which optimizes all of the objective functions. Hence, multi-objective optimization consists of the trade-off between objective functions. To obtain Pareto and weak efficient solutions of a multi-objective optimization problem, usually, the problem is converted to a scalar optimization problem. There are some scalarizing approaches to transform a multi-objective optimization problem to a scalar optimization problem. Weighted sum method is one of the simplest and popular method to convert a multi-objective optimization problem to a scalar optimization problem. In weighted sum method some nonnegative constants (weights) are multiplied to the objective functions, and then the summation of the weighted objective functions is considered as the objective function of the scalar optimization. Goal programming (GP) is one of the oldest techniques in multi-objective optimization which was introduced in [2]. GP has widely been applied to many different areas.

One of the applications of GP is transformation of a multi-objective optimization problem to a single-objective optimization problem such that we added objective reflection variable to the objective functions and turn them to constraints. In GP a set of targets for the objective functions are clarified by decision maker and the Unwanted deviations from this set of target values are minimized by an achievement function. This function can be a vector or a weighted sum regarding to the GP variant used.

In this paper, we give a model to explore the RAP using the objective functions

as maximizing the system reliability and minimizing of both total cost of system and the variance of the total cost with the constraints on weight (or volume), and entropy of the system. Minimizing the variance of the cost of system is considered as one of the objective functions for the robustness in system. Then we use the WS method to minimize the weighted sum of the objective functions including the system reliability, the system cost and the variance of the system cost. Furthermore, we propose GP to minimize the goal deviation variables which are defined for the system reliability, the total cost and the variance of the system cost. In these methods, the multi-objective RAP is transformed to a scalar optimization problem.

This paper is organized as follows. In section 2, we introduce the RAP model and proposed approaches based on WS and GP for solving RAP model. An experimental example is given in Section 3 to illustrate the proposed methods and determine the optimal configuration of each approaches.

## 2 RAP Model and multi-objective approaches based on WS and GP

In this section, we explore the RAP model in which maximizing the reliability and minimizing the cost and its variance are considered as the objective functions. Moreover, this model has constraints on the total weight, number of components and entropy. The used notations and variables are shown in Table 1.

The measure of allocation is defined as the expected information of the message which transforms in the system based on the components number of its subsystems. Then the Shannon entropy is written as

$$H(\mathbf{p}) = - \sum_{i=1}^s p_i \log p_i,$$

where

$$p_i = \frac{\sum_{j=1}^m x_{ij}}{\sum_{i=1}^s \sum_{j=1}^m x_{ij}}, \quad i = 1, \dots, s.$$

From concept of entropy, it is concluded that whatever the amount of entropy measure increases, uncertainty increases. Hence, indicating an upper bound for entropy or uncertainty is necessary in the RAP. Therefore, the model is stated as follows:

Table 1: The parameters and variables which are used in the proposed method

Notation	Description	Notation	Description
$i$	Index for a subsystem	$j$	Index for a component type
$s$	Number of subsystems	$m$	Number of available component types (equally functioning) for subsystem $i$
$r_j$	Reliability of component type $j$	$R(\mathbf{x})$	Overall system reliability
$x_{ij}$	A decision variables denoting the number of component type $j$ in subsystem $i$	$\mathbf{x}$	A vector specifying the redundancy levels and component choice of the multi-objective RAP
$C(\mathbf{x})$	Overall system cost	$C_o(\mathbf{x})$	Operational cost
$C_{cm}(\mathbf{x})$	Corrective maintenance cost	$c_j^o$	Cost per unit time of operating the $j$ th component type
$c_j^{cm}$	Corrective maintenance cost of the $j$ th component type	$t_j$	Useful life of $j$ th component type
$n_j$	The number of repairs that $j$ th component type takes during the test	$sw_j^{(v)}$	Volume for the $j$ th type of component
$sw_j^{(m)}$	Weight for $j$ th type of component	$SW_m$	Upper bound on the total system weight
$SW_v$	Upper bound on the total system volume	$Var(\cdot)$	The variance of the quantity in the parenthesis
$\sigma_{t_j}^2$	Variance of the useful life of component $j$ th type	$\sigma_{n_j}^2$	Variance of the number of repairs that $j$ th component type takes during the test
$n_{min,i}$	Minimum number of components required for subsystem $i$ to function	$n_{max,i}$	Maximum number of components used in parallel in subsystem $i$ (user defined)
$E(\cdot)$	The expected value of the quantity in the parenthesis	$E_s$	The maximum of entropy in the model
$d_R^-$	A positive variable denoting the negative deviation from the reliability's goal	$d_R^+$	A positive variable denoting the positive deviation from the reliability's goal
$d_C^-$	A positive variable denoting the negative deviation from the cost's goal	$d_C^+$	A positive variable denoting the positive deviation from the cost's goal
$d_{VC}^-$	A positive variable denoting to negative deviation from the variance of cost's goal	$d_{VC}^+$	A positive variable denoting the positive deviation from the variance of cost's goal

## Model:

$$\text{Maximize } \hat{R}(\mathbf{x}), \quad (1)$$

$$\text{Minimize } \hat{C}(\mathbf{x}), \quad (2)$$

$$\text{Minimize } Var(\hat{C}(\mathbf{x})) \quad (3)$$

$$\text{Subject to } \sum_{i=1}^s \sum_{j=1}^m \mathbf{s}w_j x_{ij} \leq \mathbf{SW}, \quad (4)$$

$$n_{min,i} \leq \sum_{j=1}^m x_{ij} \leq n_{max,i}, \quad (5)$$

$$-\sum_{i=1}^s \frac{\sum_{j=1}^m x_{ij}}{\sum_{i=1}^s \sum_{j=1}^m x_{ij}} \log \left( \frac{\sum_{j=1}^m x_{ij}}{\sum_{i=1}^s \sum_{j=1}^m x_{ij}} \right) \leq E_s, \quad (6)$$

$$x_{ij} \in \{0, 1, 2, \dots\}, \quad i = 1, 2, \dots, s, \quad j = 1, 2, \dots, m.$$

The objectives are represented in (1)-(3) which express maximizing the system reli-

bility, minimizing the system cost and its variance, respectively. Relations (4)-(6) are our model's constraints. In (4),  $\mathbf{sw}_j$  ( $j = 1, \dots, m$ ), and  $\mathbf{SW}$  denote the vectors of weights and volumes for the  $j$ th type of components and upper bound on the total system weight and volume, respectively. In other words,

$$\mathbf{SW} = \begin{pmatrix} SW_m \\ SW_v \end{pmatrix}, \quad \mathbf{sw}_j = \begin{pmatrix} sw_j^{(m)} \\ sw_j^{(v)} \end{pmatrix}, \quad j = 1, \dots, m. \quad (7)$$

For a series-parallel system, the system reliability is estimated as follows [1]:

$$\hat{R}(\mathbf{x}) = \prod_{i=1}^s (1 - \prod_{j=1}^m \hat{q}_j^{x_{ij}}),$$

where  $\hat{q}_j = 1 - \hat{r}_j$  and  $\hat{r}_j$  is the reliability of component  $j$ th type. The overall cost consists of operational cost,  $C_o(\mathbf{x})$ , and corrective maintenance cost,  $C_{cm}(\mathbf{x})$ . Operational cost is attained by multiplication of the operation cost per unit time and its associated time. Corrective maintenance (CM) cost can be attained by the product of the corrective maintenance cost of each component and the number of repairs of component in the mission time [10, 9]. Therefore, the estimation of overall system cost can be declared as

$$\hat{C}(\mathbf{x}) = \hat{C}_o(\mathbf{x}) + \hat{C}_{cm}(\mathbf{x}) = \sum_{i=1}^s \sum_{j=1}^m C_j^o \hat{t}_j x_{ij} + \sum_{i=1}^s \sum_{j=1}^m C_j^{cm} \hat{n}_j x_{ij},$$

where  $\hat{t}_j$  is the useful life estimate of component  $j$  having normal distribution with mean  $\mu_{t_j}$  and standard deviation  $\sigma_{t_j}$ , and  $\hat{n}_j$  is the estimation of the number of repairs that the component type  $j$  carries out during the test according to a Poisson distribution with the mean  $\mu_j$  and hence the standard deviation  $\sigma_j = \sqrt{\mu_j}$ ,  $j = 1, \dots, m$ . Therefore, the variance of the total cost of system is obtained as follows:

$$\begin{aligned} Var(\hat{C}(\mathbf{x})) &= \sum_{i=1}^s \sum_{j=1}^m (c_j^o x_{ij})^2 var(t_j) + \sum_{i=1}^s \sum_{j=1}^m (c_j^{cm} x_{ij})^2 var(n_j) \\ &= \sum_{i=1}^s \sum_{j=1}^m (c_j^o x_{ij})^2 \sigma_{t_j}^2 + \sum_{i=1}^s \sum_{j=1}^m (c_j^{cm} x_{ij})^2 \sigma_{n_j}^2. \end{aligned} \quad (8)$$

Now, we investigate the multi-objective redundancy allocation problem (MORAP) with weighted sum optimization, denoted by  $WS(\lambda_1, \lambda_2, \lambda_3)$ , as the following model.

## WS model:

$$\text{Minimize} \quad -\lambda_1 \hat{R}(\mathbf{x}) + \lambda_2 \hat{C}(\mathbf{x}) + \lambda_3 \text{Var}(\hat{C}(\mathbf{x})) \quad (9)$$

$$\text{subject to} \quad \sum_{i=1}^S \sum_{j=1}^m \mathbf{s}\mathbf{w}_j x_{ij} \leq \mathbf{SW}, \quad (10)$$

$$n_{min,i} \leq \sum_{j=1}^m x_{ij} \leq n_{max,i}, \quad (11)$$

$$-\sum_{i=1}^s \frac{\sum_{j=1}^m x_{ij}}{\sum_{i=1}^s \sum_{j=1}^m x_{ij}} \log \left( \frac{\sum_{j=1}^m x_{ij}}{\sum_{i=1}^s \sum_{j=1}^m x_{ij}} \right) \leq E_s, \quad (12)$$

$$x_{ij} \in \{0, 1, 2, \dots\}, \quad i = 1, 2, \dots, s, \quad j = 1, 2, \dots, m, \quad (13)$$

where  $\sum_{k=1}^3 \lambda_k = 1$ ,  $\lambda_k \geq 0$ ,  $k = 1, 2, 3$ . Based on decision maker, we can choose parameter  $(\lambda_1, \lambda_2, \lambda_3)$  such that  $\sum_{k=1}^3 \lambda_k = 1$ . In the numerical example we choose some parameters and then compare the corresponding results.

In the GP model, the objective function (14) is given to minimize the deviations of the reliability, the cost and the cost variance. Since the reliability is more important than other objective functions, the coefficient 100 is used for that. The objective function of the model find a feasible solution of RAP model in which the values of the reliability, the cost and its variance corresponding to that is the nearest values to  $g_R$ ,  $g_C$  and  $g_{VC}$ , respectively. Relations (15)-(17) indicate the constraints on the reliability, cost and variance of cost using the deviation variables and the expected goals of them. Goals  $g_R$ ,  $g_C$ , and  $g_{VC}$  are identified corresponding to the decision maker preferences.

## GP model:

$$\text{Minimize} \quad 100d_R^- + d_C^- + d_{VC}^+, \quad (14)$$

$$\text{subject to} \quad \hat{R}(\mathbf{x}) + d_R^- - d_R^+ \geq g_R, \quad (15)$$

$$\hat{C}(\mathbf{x}) + d_C^- - d_C^+ \leq g_C, \quad (16)$$

$$\text{Var}(\hat{C}(\mathbf{x})) + d_{VC}^- - d_{VC}^+ \leq g_{VC}, \quad (17)$$

$$\sum_{i=1}^S \sum_{j=1}^m \mathbf{s}\mathbf{w}_j x_{ij} \leq \mathbf{SW}, \quad (18)$$

$$n_{min,i} \leq \sum_{j=1}^m x_{ij} \leq n_{max,i}, \quad (19)$$

$$-\sum_{i=1}^s \frac{\sum_{j=1}^m x_{ij}}{\sum_{i=1}^s \sum_{j=1}^m x_{ij}} \log \left( \frac{\sum_{j=1}^m x_{ij}}{\sum_{i=1}^s \sum_{j=1}^m x_{ij}} \right) \leq E_s, \quad (20)$$

$$x_{ij} \in \{0, 1, 2, \dots\}, \quad i = 1, 2, \dots, s, \quad j = 1, 2, \dots, m,$$

$$d_R^-, d_R^+, d_C^-, d_C^+, d_{VC}^-, d_{VC}^+ \geq 0$$

The objective of reliability has conflict with the objective of the total cost and its variance. This means that if the system reliability increases, then the system cost and its variance also increase. The behavior of these three objective functions is considered as the trade-off between the reliability, the total cost and its variance. As mentioned earlier, the goals  $g_R$ ,  $g_C$  and  $g_{VC}$  are determined corresponding to the viewpoint of decision maker. This implies that it is possible that there does not exist any feasible solution of the model such that the values of reliability, the total cost and its variance are greater than  $g_R$ , less than  $g_C$  and less than  $g_{VC}$ , respectively.

### 3 Experimental results

Consider a series-parallel system consists of three subsystems and four components in each subsystem where the components are independent and may have different types. In this example we consider only system weight and ignore the system volume. The acceptable total weight of system is assumed to be at most 45 ( $SW_m = 45$ ). Similarly, we can use the system volume and then solve the problem. The minimum and maximum number of components which are selected in each subsystem are one and seven, respectively,  $n_{min,i} = 1$  and  $n_{max,i} = 7$ ,  $i = 1, \dots, 4$ . Table 2 shows all component data which are required in the proposed model. To solve MORAP for the described system, the WS and GP methods are implemented in GAMS running on a system Intel CORE<sup>TM</sup> i5-2530, CPU @ 2.30 GHz, Ram 4 GB.

Table 2: Reliability, cost and cost variance data

$j$	$X_j$	$N_j$	$\hat{r}_j$	$sw_j$	$C_j^o$	$C_j^{cm}$	$\hat{t}_j$	$\hat{n}_j$	$\sigma_{t_j}^2$	$\sigma_{n_j}^2$
1	4	20	0.8	5	0.2	0.305	0.25	0.15	0.133	0.093
2	3	20	0.85	6	0.15	0.3	0.375	0.135	0.2	0.106
3	4	25	0.86	7	0.3	0.2	0.15	0.31	0.146	0.146
4	9	50	0.82	8	0.175	0.165	0.188	0.2	0.106	0.14

Tables 3 and 4 show the results including of decision maker values, the expectation value, and the variance of reliability and cost by employing the WS method with some parameters.

Table 3: Results of employing WS approach with different parameters

Approach	$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$	$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$	$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
$WS(0.5, 0.4, 0.1)$	0	1	0	0	0	0	0	2	0	1	0	0
$WS(0.6, 0.3, 0.1)$	0	0	0	2	0	1	0	0	0	1	0	1
$WS(0.7, 0.2, 0.1)$	0	0	0	2	0	1	0	1	0	1	0	0

Table 4: Expectation value and variance of reliability and cost by employing WS approach with different parameters

Approach	$E[\hat{R}(\mathbf{x})]$	$E[\hat{C}(\mathbf{x})]$	$E[Var(\hat{C}(\mathbf{x}))]$	$E[WS]$
$WS(0.5, 0.4, 0.1)$	0.6991	0.3253	0.0563	0.2138
$WS(0.6, 0.3, 0.1)$	0.8003	0.3912	0.0634	0.3565
$WS(0.7, 0.2, 0.1)$	0.8003	0.3912	0.0634	0.3565

The optimal configuration of WS method with different parameters is depicted in Figure 1.

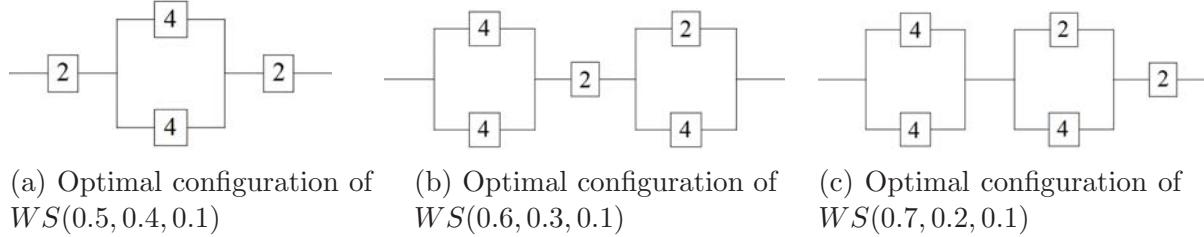


Figure 1: Configuration of WS method

Now we solve the problem with GP approach and give the results including of decision maker value, the expectation value, and the variance of reliability and cost in Tables 5 and 6. The configuration of GP method with different parameters is shown in Figure 2.

Table 5: Results of employing GP approach with different parameters

Approach	$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$	$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$	$x_{31}$	$x_{32}$	$x_{33}$	$x_{34}$
$GP(0.94, 0.2, 3)$	3	0	0	0	0	0	0	2	0	1	0	1
$GP(0.85, 0.1, 3)$	0	0	1	1	0	0	0	2	0	0	0	1
$GP(0.9, 0.5, 5)$	0	0	2	1	0	0	1	0	0	0	1	0

Table 6: Expectation value and variance of reliability and cost by employing GP approach with different parameters

Approach	$E[\hat{R}(\mathbf{x})]$	$E[\hat{C}(\mathbf{x})]$	$E[Var(\hat{C}(\mathbf{x}))]$	$E[WS]$
$GP(0.94, 0.2, 3)$	0.9340	0.5820	3.0000	96.9820
$GP(0.85, 0.1, 3)$	0.7730	0.2810	3.0000	80.5810
$GP(0.9, 0.5, 5)$	0.7370	0.5000	5.0000	79.2000

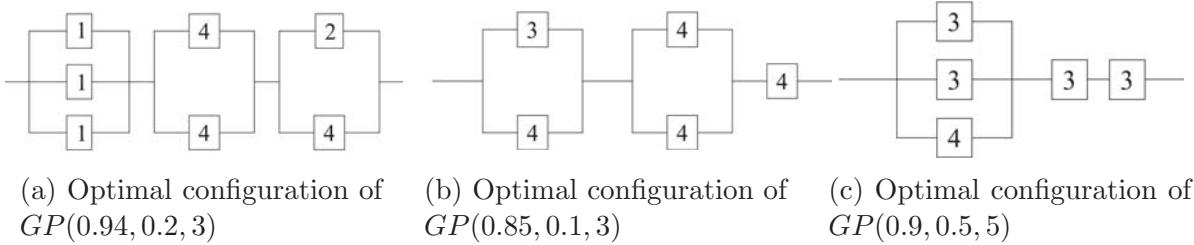


Figure 2: Configuration of GP method

## References

- [1] Barlow, R.E., Proschan, F. (1975), *Statistical theory of reliability and life testing: probability models*. Florida State Univ Tallahassee.
- [2] Charnes, A., Cooper, W.W., Ferguson, R. (1955), *Optimal estimation of executive compensation by linear programming*, Management Science, 1, 138-151.

- [3] Chern, M.S. (1992), *On the computational complexity of reliability redundancy allocation in a series system*. Oper. Res. Lett. 11(1) 309-315.
- [4] Fyffe, D.E., Hines, W.W., Lee, N.K. (1968), *System reliability allocation problem and a computational algorithm*, IEEE Trans. Reliab. 17 (1) 64-69.
- [5] Mahapatra, G.S., & Mahapatra, B.S. (2011), *Intuitionistic fuzzy optimization technique on reliability redundancy optimization with entropy function*. International Journal of Performability Engineering, (7) 155-164.
- [6] Rocchi, P. (2002), *Boltzmann-like entropy in reliability theory*. Entropy, (4) 142-150.
- [7] Roy, P., Mahapatra, B.S., Mahapatra, G.S., Roy, P.K. (2014), *Entropy based region reducing genetic algorithm for reliability redundancy allocation in interval environment*. Expert Systems with Applications.
- [8] Salmasnia, A., Ameri, E., Niaki, S.T.K. (2016), *A robust loss function approach for a multi-objective redundancy allocation problem*, Appl. Math. Model. 40 (1) 635-645.
- [9] Wang, K.-H., Kuo, C.-C. (2000), *Cost and probabilistic analysis of series systems with mixed standby components*, Appl. Math. Model. 24 (12) 957-967.
- [10] Wang, L., Chu, J., Wu, J. (2007), *Selection of optimum maintenance strategies based on a fuzzy analytic hierarchy process*, Int. J. Prod. Econ. 107 (1) 151-163.



Forth Seminar on  
**Reliability Theory and its Applications**

April 25-26, 2018



## Statistical inference for a repairable system with Weibull distribution subject to shocks

Kamranfar, H.<sup>1</sup>      Etminan, J.<sup>2</sup>      Chahkandi, M.<sup>3</sup>

<sup>1,2,3</sup> Department of Statistics, Faculty of Mathematics and Statistics, University of Birjand

### Abstract

Consider a repairable system that is replaced at the  $n$ th type I failure, or at the first type II failure whichever comes first. Type I failure occurs with probability  $p$  and is removed by minimal repair, whereas type II failure occurs with probability  $1 - p$  and is removed by replacement. Theoretical results usually consider an interarrival-time distribution  $F$  with known parameters, and obtain an optimal policy in selecting  $n$  to minimize the expected cost. However, this is not generally the case in practice, and the parameters are usually unknown. In this paper, we assume that  $F$  is a Weibull distribution, and then study the problem of estimating the parameters of  $F$  and the repair efficiency parameter  $p$ . The likelihood-ratio test statistic is also obtained for testing the parameters. Finally, a Monte Carlo simulation study is conducted to compute the critical values and power estimates of the proposed test.

**Keywords:** Reliability and Maintainability, Imperfect repair, Weibull Distribution, Maximum likelihood estimation, Hypothesis test.

## 1 Introduction

A lot of models are proposed to study the stochastic behavior of a repairable system. The most common models for repairable systems are minimal repair and perfect repair.

<sup>1</sup>Hoda.kamranfar@birjand.ac.ir

<sup>2</sup>jetminan@birjand.ac.ir

<sup>3</sup>mchahkandi@birjand.ac.ir

More realistic notions of repair fall between these two extremes. This situation is referred to as “imperfect repair”. Brown and Proschan (1983) (BP) proposed one of the first models for imperfect repair that takes into account two types of repairs (perfect and minimal). Kijima (1989) considered general repairs and introduced the concept of virtual age in general repair models for a repairable system. Last and Szekli (1998) introduced a general model for repairable systems that contains most of other models. They considered a repairable system with random lifetime  $T_1$  and random degree of repair  $Z_1$ ;  $0 \leq Z_1 \leq 1$ ; where the case of 0 corresponds to the minimal repair and the case of 1 to the perfect repair. If  $\lambda_1(t) = \lambda(t)$  denotes the failure rate of  $T_1$ , then the distribution of the time until the next failure has the failure rate  $\lambda_2(t) = \lambda(t - T_1 + V_1)$ . Now assume that  $T_n$  and  $Z_n$  denote the time of  $n$ th ( $n \geq 1$ ) failure and the degree of repair at that time, respectively. Therefore, the failure rate of the  $(n+1)$ th waiting time until the next failure can be obtained by  $\lambda_{(n+1)}(t) = \lambda(t - T_n + V_n)$ , where  $T_n \leq t < T_{(n+1)}$ ,  $n \geq 0$  and  $V_n = (1 - Z_n)(V_{(n-1)} + T_n - T_{(n-1)})$ ,  $V_0 = 0$ ,  $T_0 = 0$ . They finally introduced the virtual age process as  $V(t) = t - T_n + V_n$ ,  $T_n \leq t < T_{(n+1)}$ ,  $n \geq 1$ . Figure 1 is a simple representation of the model.

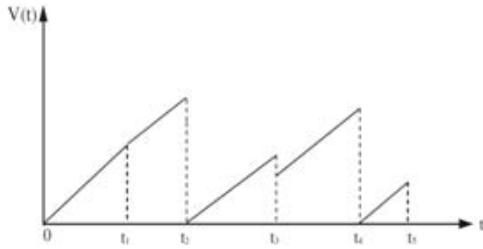


Figure 1: The virtual age process.

In Figure 1  $t_i, i = 1, \dots, 5$  are the realizations of  $T_i$ , at time  $t_1$  we have a minimal repair. At times  $t_2$  and  $t_4$  we have a perfect repair and an imperfect repair is occurred at  $t_3$ . As an especial case of the virtual age process introduced by Last and Szekli (1998), one may obtained the Brown and Proschan (BP) model as follows. Consider the i.i.d. random variables  $Z_k$ ;  $1 \leq k \leq n$  independent of  $(T_1, \dots, T_k)$  such that  $1 - P(Z_k = 0) = P(Z_k = 1) = p$ ;  $p \in (0, 1)$ , then the failure process is given by  $\lambda_{n+1}(t) = \lambda(t - T_n + (1 - Z_n)M_n)$ , where  $M_n$  is the time from  $T_n$  to the last perfect repair.

Sheu and Griffith (1996) studied another expansion of BP model to the case that the system is replaced at the  $n$ th type-I failure, or at the first type-II failure whichever occurs first and based on an optimal replacement policy, determined the optimal number of  $n$ . They also supposed that shocks occur according to a non-homogeneous Poisson process (NHPP). In this policy, the optimization criterion is often a long run average cost rate or availability evaluated from system lifetime distribution and maintenance unit costs. This optimal maintenance policy depends on the parameters of the system lifetime. Most of the authors suppose that the reliability parameters and repair effects are known. However, in practice, these parameters are often unknown. Estimation of reliability and maintainability parameters of a repairable system is essential in determining the maintenance policies. The problem of estimation of parameters in the virtual age process introduced by Last and Szekli (1998) and its especial cases such that BP models are investigated by many authors. The most famous are Gasmi and Kahle (1998), Gasmi (2012) and Gasmi (2013).

In this paper, we discuss the use of likelihood function for characterizing the failure process behavior of the model introduced by Sheu and Griffith (1996). We find the maximum likelihood estimators of the model parameters and develop the likelihood ratio (LR) method for hypothesis tests. The uncertainties and errors in the estimation of the parameters, critical values and power estimates of the proposed tests are determined by Monte Carlo simulation.

## 2 Model description

In this section the failure process with unknown reliability parameters and unknown repair effect will be characterized with the likelihood function in order to find the parameter estimations. First, we introduce some notations which will be used in our derivations. Suppose that  $F$  is the system lifetime before the first failure,  $M$  is the number of shocks until the first type-II failure and  $n$  is the maximum number of minimal repairs before replacement. Consider the system with two types of failures, minor and catastrophic, that the first one is fixed by a minimal repair with probability  $q = 1 - p$ , and the second one is removed by an unplanned replacement with probability  $p$ . The system is replaced at the  $n$ th type-I failure, or at the first type-II failure whichever occurs first. Let  $\mathbf{x} = (x_1, \dots, x_m)$ ;  $1 \leq m \leq n$ , denote the observed failure times of the system before replacement. Then the corresponding likelihood function can be expressed as

$$\begin{aligned} L(p, F | \mathbf{x}, m) &= \left( \prod_{i=1}^{n-1} q^{I(m>i)} p^{I(m=i)} [f(x_i | x_{i-1})]^{I(m \geq i)} \right) [f(x_n | x_{n-1})]^{I(m \geq n)} \\ &= q^{m-1} \left( \prod_{i=1}^m \frac{f(x_i)}{1 - F(x_{i-1})} \right) p^{I(m < n)}, \end{aligned} \quad (1)$$

where  $f(x_i | x_{i-1})$  is the truncated density function of  $x_i$  given  $x_{i-1}$  and  $x_0 = -\infty$ .

So, for a sample of  $k$  i.i.d systems (or  $k$  replacements), the likelihood function can be written as,

$$\begin{aligned} L(p, F | \tilde{\mathbf{x}}, \mathbf{m}) &= \prod_{j=1}^k \prod_{i=1}^{m_j} \left( q^{m_j-1} p^{I(m_j < n)} \frac{f(x_{j,i})}{1 - F(x_{j,i-1})} \right) \\ &= q^{(m^{(k)} - k)} p^{\sum_{j=1}^k I(m_j < n)} \prod_{j=1}^k \prod_{i=1}^{m_j} \frac{f(x_{j,i})}{\bar{F}(x_{j,i-1})}. \end{aligned} \quad (2)$$

where  $\tilde{\mathbf{x}} = (\mathbf{x}_1, \dots, \mathbf{x}_k)$ ,  $m^{(k)} = \sum_{j=1}^k m_j$  and  $\bar{F}(\cdot) = 1 - F(\cdot)$  is the survival function of  $F$ .

Now, we apply the log function to the expression of the likelihood function,

$$\begin{aligned} \log(L(p, F | \tilde{\mathbf{x}}, \mathbf{m})) &= (m^{(k)} - k) \log(q) + \left( \sum_{j=1}^k I(m_j < n) \right) \log(p) \\ &\quad + \sum_{j=1}^k \sum_{i=1}^{m_j} \log \left( \frac{f(x_{j,i})}{\bar{F}(x_{j,i-1})} \right). \end{aligned} \quad (3)$$

To maximize the likelihood function with respect to  $p$  we must calculate the derivative of (3) with respect to  $p$ , and solve the following equation:

$$-\frac{m^{(k)} - k}{1 - p} + \frac{\sum_{j=1}^k I(m_j < n)}{p} = 0.$$

Hence, the maximum likelihood estimate (MLE) of  $p$  is:

$$\hat{p} = \frac{\sum_{j=1}^k I(m_j < n)}{m^{(k)} - k + \sum_{j=1}^k I(m_j < n)},$$

where is free of  $\tilde{\mathbf{x}}$ .

Now, suppose that  $F$  is a Weibull distribution with the density function  $f(x) = \alpha\lambda x^{\alpha-1}e^{-\lambda x^\alpha}$ , where  $x, \lambda, \alpha > 0$ . Then the likelihood function (1) can be rewritten as,

$$\begin{aligned} L(\alpha, \lambda, p | \mathbf{x}, m) &= q^{m-1} \left( \prod_{i=1}^m \frac{\alpha\lambda x_i^{\alpha-1} e^{-\lambda x_i^\alpha}}{e^{-\lambda x_{i-1}^\alpha}} \right) p^{I(m < n)} \\ &= q^{m-1} p^{I(m < n)} (\alpha\lambda)^m e^{-\lambda x_m^\alpha} \left( \prod_{i=1}^m x_i^{\alpha-1} \right). \end{aligned}$$

For a sample of  $k$  i.i.d systems, we also have

$$L(\alpha, \lambda, p | \tilde{\mathbf{x}}, \mathbf{m}) = q^{\sum_{j=1}^k m_j - k} p^{\sum_{j=1}^k I(m_j < n)} (\alpha\lambda)^{\sum_{j=1}^k m_j} e^{-\lambda \sum_{j=1}^k x_{j,m_j}^\alpha} \left( \prod_{j=1}^k \prod_{i=1}^{m_j} x_{j,i}^{\alpha-1} \right).$$

which yields

$$\begin{aligned} \log(L(\alpha, \lambda, p | \tilde{\mathbf{x}}, \mathbf{m})) &= (m^{(k)} - k) \log(q) + \sum_{j=1}^k I(m_j < n) \log(p) \\ &\quad + m^{(k)} \log(\alpha\lambda) - \lambda \sum_{j=1}^k x_{j,m_j}^\alpha + (\alpha - 1) \sum_{j=1}^k \sum_{i=1}^{m_j} \log(x_{j,i}). \end{aligned} \tag{4}$$

## 2.1 Maximum likelihood estimation

There are three parameters  $p$ ,  $\lambda$  and  $\alpha$  that need to be estimated. But, the parameter  $p$  has already been estimated. The derivatives of (4) with respect to  $\alpha$  and  $\lambda$  are

$$\frac{m^{(k)}}{\alpha} - \lambda \sum_{j=1}^k x_{j,m_j}^\alpha \log(x_{j,m_j}) + \sum_{j=1}^k \sum_{i=1}^{m_j} \log(x_{j,i}) = 0, \tag{5}$$

and

$$\frac{m^k}{\lambda} - \sum_{j=1}^k x_{j,m_j}^\alpha = 0, \tag{6}$$

respectively. A numerical algorithm has been developed to solve the system of nonlinear equations in (5) and (6).

## 2.2 Likelihood ratio test

The likelihood ratio test (LRT) is a statistical test of the goodness-of-fit between two hierarchically nested models. In this approach, the likelihood ratio, or equivalently its logarithm, can be used to compute a p-value, or compared to a critical value to decide whether or not to reject the null hypothesis. The likelihood ratio test statistic is defined as

$$\Lambda(\mathbf{x}) = \frac{L(\hat{\boldsymbol{\theta}}_0)}{L(\hat{\boldsymbol{\theta}})} = \frac{\sup_{\boldsymbol{\theta} \in \Theta_0} L(\boldsymbol{\theta} | \mathbf{x})}{\sup_{\boldsymbol{\theta} \in \Theta} L(\boldsymbol{\theta} | \mathbf{x})}.$$

Let  $\boldsymbol{\theta}_0 = (p_0, \lambda_0, \alpha_0)'$ , where  $p_0 \in (0, 1)$ ,  $\lambda_0 > 0$  and  $\alpha_0 > 0$ , are known. In what follows, we are interested in testing the hypothesis

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \quad vs \quad H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0.$$

By inserting representation (4), it is readily seen that the test statistic  $T_{LR} = \log(\Lambda(\mathbf{x}))$  of the LR test is given by

$$\begin{aligned} T_{LR} = & m^{(k)} \left( \log\left(\frac{q_0}{\hat{q}}\right) + \log\left(\frac{\alpha_0 \lambda_0}{\hat{\alpha} \hat{\lambda}}\right) + 1 \right) - k \log\left(\frac{q_0}{\hat{q}}\right) + \sum_{j=1}^k I(m_j < n) \log\left(\frac{p_0}{\hat{p}}\right) \\ & - \lambda_0 \sum_{j=1}^k x_{j,m_j}^{\alpha_0} + (\alpha_0 - \hat{\alpha}) \sum_{j=1}^k \sum_{i=1}^{m_j} \log(x_{j,i}). \end{aligned} \quad (7)$$

## 3 Simulation study

This section presents some simulation results to verify the behavior of the MLEs and the proposed test statistic, for different choices of parameters.

Suppose  $X \sim \text{Weibull}(\lambda, \alpha)$ , then it is well known that  $X^\alpha \sim \text{Exp}(\lambda)$ . As shown earlier, the MLE for the repair parameter  $p$  is free of the failure times, hence we consider two cases for our simulation study.

### 3.1 Case I

$\alpha$  is known:

After choosing the value of  $\alpha$  (where is arbitrary), we used 100,000 replications Monte Carlo simulations for  $k = 10$  and  $n = 5$ . The interarrival-time and the sample sizes are simulated from  $\sim \text{Exp}(\lambda)$ , and a truncated geometric random variable with parameter  $p$ , where only takes the values which are less than or equal to  $n$ , respectively.

Table 1 shows the estimated average biases and mean square errors of the estimators  $\hat{p}$  and  $\hat{\lambda}$  for different values of  $p$  and  $\lambda$ . Two important discoveries are made in examining Table 1. The biases are positive, but quite close to zero (a little overestimated). As  $p$  gets smaller (sample size increases), both biases and MSEs get smaller.

Table 1: Biases and MSEs of  $(\hat{p}, \hat{\lambda})$  for  $k = 10$ ,  $n = 5$  and different values of  $(p, \lambda)$ .

$(\lambda, p)$	Bias		MSE	
	$\hat{p}$	$\hat{\lambda}$	$\hat{p}$	$\hat{\lambda}$
(1,0.2)	0.00873295	0.03213010	0.00619848	0.03595313
(1,0.3)	0.01350030	0.03847862	0.00964458	0.04529289
(1,0.4)	0.01755853	0.04842994	0.01272285	0.05753790
(1,0.5)	0.02086166	0.05573677	0.01507586	0.07068132
(2,0.2)	0.00884359	0.06309478	0.00624778	0.14187352
(2,0.3)	0.01303700	0.07769956	0.00955079	0.18004122
(2,0.4)	0.01722413	0.09156155	0.01275044	0.22903375
(2,0.5)	0.02058338	0.11445848	0.01502247	0.28813282
(3,0.2)	0.00893277	0.09291345	0.00627589	0.31914605
(3,0.3)	0.01273263	0.11743729	0.00963599	0.40399604
(3,0.4)	0.01694013	0.14123302	0.01278636	0.51139489
(3,0.5)	0.01963032	0.16867396	0.01489810	0.63957372

In Table 2 the simulated powers of the test statistic  $T_{LR}$  with 100,000 replications are listed for the test problem (8) at significance level 0.05.

$$H_0 : (p, \lambda) = (0.3, 2) \quad vs \quad H_1 : (p, \lambda) \neq (0.3, 2) \quad (8)$$

We observe the LRT is an unbiased test. Also, Table 2 reveals that the proposed test has a good performance.

Table 2: Power values of the test problem (8) at 5% level of significance for some specified alternatives.

$\lambda =$	1	1.25	1.5	1.75	2	2.25	2.5	2.75	3
$p=0.20$	0.95486	0.75125	0.44511	0.23966	0.18172	0.22195	0.32743	0.46109	0.60457
$p=0.24$	0.93458	0.68018	0.3391	0.13938	0.08971	0.12511	0.21772	0.34749	0.50213
$p=0.28$	0.91602	0.62823	0.28405	0.10105	0.05491	0.08229	0.16322	0.28617	0.43151
$p=0.32$	0.90454	0.60586	0.26879	0.09357	0.05547	0.08006	0.1528	0.26422	0.40669
$p=0.36$	0.89952	0.60174	0.28089	0.11775	0.08177	0.10502	0.17454	0.28068	0.4189
$p=0.40$	0.90096	0.6259	0.32601	0.16771	0.12939	0.15334	0.22332	0.32884	0.45547

### 3.2 Case II

$p = 0.3$ ,  $\lambda$  and  $\alpha$  are unknown:

Here, we are interested in the behavior of the estimation of Weibull distribution parameters. Table 3 shows that the bias of  $\hat{\lambda}$  is quite close to zero, and the bias of  $\hat{\alpha}$  increases as  $\alpha$  gets larger.

In Table 4 the simulated powers of test statistic  $T_{LR}$  are listed for the test problem (9) at significance level 0.05.

$$H_0 : (\alpha, \lambda) = (1, 0.5) \quad vs \quad H_1 : (\alpha, \lambda) \neq (1, 0.5) \quad (9)$$

Note, the sample size  $m^{(k)} = \sum_{j=1}^k m_j$  is a random variable in this model. In fact,  $m^{(k)}$  can be rewritten as  $m^{(k)} = \sum_{i=1}^n iN_i$ , such that  $N_i$  is defined as  $\sum_{j=1}^k I_{(m_j=i)}$ , for  $i = 1, \dots, n$ .

Table 3: Biases and MSEs of  $(\hat{\alpha}, \hat{\lambda})$  for  $k = 10$ ,  $n = 5$  and different values of  $(\alpha, \lambda)$ .

$(\alpha, \lambda)$	Bias		MSE	
	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\lambda}$
(2,0.4)	0.1209741	-0.006862928	0.1515161	0.02602586
(2,0.8)	0.1222854	-0.008594671	0.1544467	0.06418861
(2,1.5)	0.1239195	0.009857756	0.1573602	0.14662210
(4,0.4)	0.2425361	-0.006793185	0.6090908	0.02603825
(4,0.8)	0.2453628	-0.008426908	0.6206782	0.06420258
(4,1.5)	0.2481689	0.009938412	0.6308754	0.14658040
(6,0.4)	0.3659331	-0.006616465	1.385850	0.02606924
(6,0.8)	0.3697353	-0.008248846	1.409813	0.06416231
(6,1.5)	0.3735045	0.010106790	1.430124	0.14641980

Table 4: Power values of the test problem (9) at 5% level of significance for some specified alternatives.

$\alpha=$	1	2	3	4	5	6
$\lambda=0.2$	0.98493	0.99364	0.99814	0.9983	0.99842	0.99856
$\lambda=0.4$	0.17164	0.99797	0.99839	0.99851	0.99878	0.99893
$\lambda=0.6$	0.11516	0.99844	0.99863	0.99889	0.99902	0.99921
$\lambda=0.8$	0.5273	0.99876	0.99898	0.99918	0.99937	0.99946
$\lambda=1.5$	0.87674	0.99899	0.99924	0.99943	0.99952	0.9996
$\lambda=1.5$	0.99817	0.99951	0.99966	0.99973	0.99983	0.9999

Hence, the vector variable  $\mathbf{N} = (N_1, \dots, N_n) \sim MB(k; p_1, \dots, p_n)$ , i.e.

$$f_{\mathbf{N}}(\mathbf{y}) = P(N_1 = y_1, \dots, N_n = y_n) = \binom{k}{y_1, \dots, y_n} p_1^{y_1} \cdots p_n^{y_n}, \quad y_1 + \cdots + y_n = k,$$

where,

$$p_i = \begin{cases} p(1-p)^{i-1} & i = 1, \dots, n-1, \\ (1-p)^{n-1} & i = n. \end{cases}$$

Critical values are determined as the percentage points of the obtained empirical distribution of the test statistic and presented in Table 5. Table 5 shows the sensitivity of critical values to sample sizes for some different patterns that are randomly selected, except two extreme cases (patterns 1 and 12). We observe that the critical values are very close together and to the asymptotic critical value, i.e.  $-\chi^2_{(0.95,2)}/2 = -2.9957$ .

## References

- [1] Brown, M. and Proschan, F. (1983), Imperfect repair. *Journal of Applied probability* **20** (4), 851-859.
- [2] Gasmi, S. and Kahle, W. (1998), Parameter estimation in renewal processes with imperfect repair, *Advances in Stochastic Models for Reliability, Quality and Safety*, Birkhauser Book Series, Statistics for Industry and Technology, 53-65.
- [3] Gasmi, S. (2012), Estimation of the Parameters in an Alternating Repair Model Using a Pareto Intensity of the Second Kind, *Applied Mathematical Sciences* **43** (6), 2111 - 2126.

Table 5: Critical values of the test problem (9) at 5% level of significance for some failure patterns of the systems.

Pattern	<b><i>N</i></b>					$m^{(k)}$	Cr. Value
	1	2	3	4	5		
1	10	0	0	0	0	10	-3.311294
2	5	1	1	2	1	23	-3.161967
3	2	4	2	1	1	25	-3.081804
4	4	1	3	1	2	26	-3.121244
5	4	0	2	1	3	29	-3.095242
6	2	5	0	1	2	30	-3.109490
7	3	1	1	3	2	30	-3.099712
8	2	2	2	1	3	31	-3.121101
9	3	0	2	1	4	33	-3.081196
10	2	2	0	2	4	34	-3.092868
11	2	1	1	0	6	37	-3.102512
12	0	0	0	0	10	50	-3.051216

- [4] Gasmi, S. (2013), Estimating parameters of a log-linear intensity for a repairable system, *Applied Mathematical Modelling* **37** (6), 4325-4336.
- [5] Kijima, M. (1989), Some results for repairable systems with general repair, *Journal of Applied probability* **26** (1), 89-102.
- [6] Last, G. and Szekli, R. (1998), Stochastic comparison of repairable systems by coupling, *Journal of Applied probability* **35**(2) 348-370.
- [7] Sheu, S. H. and Griffith, W. S. (1996), Optimal number of minimal repairs before replacement of a system subject to shocks, *Naval Research Logistics (NRL)* **43**(3), 319-333.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



# On the coherent systems with two different types of dependent components

Kelkinnama, M. <sup>1</sup>

Department of Mathematical sciences, Isfahan University of Technology, Isfahan

## Abstract

Stochastic comparisons on coherent systems composed of heterogeneous and dependent variables are investigated. For model the dependence structure among components lifetime we employ Archimedean copula. Also, we suppose that the component distributions follow the proportional hazard rate model. Specifically we consider the 3-components coherent systems and based on some majorization orders between proportionality coefficients and by some assumptions on generator function of copula, stochastically compare the coherent systems.

**Keywords:** Archimedean copula, majorization, stochastic order, proportional hazard rates.

## 1 Introduction

Comparisons of coherent systems with heterogeneous components is an interesting topic in reliability studies. Many papers in this topic considered the series and parallel systems. For a recent work one can see [5] and for a review study, [7] is an excellent article. Recently, Coolen and CoolenMaturi [1] defined a new metric for the coherent systems and called it survival signature. Using this, they represent the reliability function of coherent systems with independent components of  $r$  different types. Samaniego and Navarro [6] using this expression, compared the coherent systems with heterogeneous independent components. Erylimaz [2] investigated a mixture form for reliability function of system lifetime in the case of two types of dependent components.

---

<sup>1</sup>m.kelkinnama@cc.iut.ac.ir

In this paper, we investigate the stochastic comparisons of coherent systems consisting of dependent components such that the distributions of components follow the proportional hazard rate models with different proportionality coefficients. The random variables  $X_1, X_2, \dots, X_n$  are said to follow the proportional hazard rates (PHR) model if for  $i = 1, 2, \dots, n$ , the reliability function of  $X_i$  can be expressed as,  $\bar{F}_i(x) = [\bar{F}(x)]^{a_i}$ , for  $a_i > 0$ , where  $\bar{F}(x)$  is the baseline reliability function. We denote this model by  $\mathbf{X} \sim PHR(\bar{F}, \mathbf{a})$  where,  $\mathbf{a} = (a_1, \dots, a_n)$  with  $a_i > 0, i = 1, \dots, n$ . If  $r(t)$  denotes the hazard rate corresponding to  $\bar{F}$ , then the hazard rate of  $X_i$  is  $a_i r(t), i = 1, 2, \dots, n$ . Many well-known models are special cases of the PHR model, for example, Weibull, Pareto, Lomax distributions.

Before proceeding the main results, let us first recall some definitions and lemmas will be used in the sequel.

**Definition 1.1.** *The random variable  $X$  is said to be smaller than random variable  $Y$  in the usual stochastic order (denoted by  $X \leq_{st} Y$ ) if  $\bar{F}(x) \leq \bar{G}(x)$  for all  $x$ .*

**Definition 1.2.** *Let  $\mathbf{x} = (x_1, \dots, x_n)$  and  $\mathbf{y} = (y_1, \dots, y_n)$  be two real vectors, denote  $x_{(1)} \leq \dots \leq x_{(n)}$  the increasing arrangement of  $x_1, \dots, x_n$ .*

*$\mathbf{x}$  is said to be weakly super-majorized by  $\mathbf{y}$  (denoted as  $\mathbf{x} \preceq^w \mathbf{y}$ ) if  $\sum_{i=1}^j x_{(i)} \geq \sum_{i=1}^j y_{(i)}$  for all  $j = 1, \dots, n-1$ .*

*$\mathbf{x}$  is said to be weakly sub-majorized by  $\mathbf{y}$  (denoted as  $\mathbf{x} \preceq_w \mathbf{y}$ ) if  $\sum_{i=j}^n x_{(i)} \leq \sum_{i=j}^n y_{(i)}$  for all  $j = 1, \dots, n-1$ .*

A permutation-symmetric continuously differentiable function  $\phi(\mathbf{X})$  is Schur-concave (Schur-convex) if and only if

$$(X_i - X_j) \left( \frac{\partial \phi(\mathbf{X})}{\partial X_i} - \frac{\partial \phi(\mathbf{X})}{\partial X_j} \right) \leq (\geq) 0$$

for all  $i \neq j$ .

For a real function  $g$  on  $A \subset R^n$ ,

(i)  $\mathbf{x} \preceq_w \mathbf{y}$  implies  $g(x) \leq g(y)$  if and only if  $g$  is increasing and Schur-convex on  $A$ , and

(ii)  $\mathbf{x} \preceq^w \mathbf{y}$  implies  $g(x) \leq g(y)$  if and only if  $g$  is decreasing and Schur-convex on  $A$

For model the dependency between random variables, the joint distribution (or survival) function may be expressed by a copula.

**Theorem 1.1.** (*Sklar's theorem*). *Let  $H$  be a joint distribution function with margins  $F$  and  $G$ . Then there exists a copula  $C$  such that for all  $x, y$  in  $R$ ,  $H(x, y) = C(F(x), G(y))$ .*

At this paper we use the family of Archimedean copulas.

**Definition 1.3.** *For a non-increasing and continuous function  $\phi : [0, +\infty) \rightarrow [0, 1]$  such that  $\phi(0) = 1$  and  $\phi(+\infty) = 0$ , let  $\psi = \phi^{-1}$  be the pseudo-inverse.*

$$C_\phi(u_1, \dots, u_n) = \phi(\psi(u_1) + \dots + \psi(u_n)), \quad (u_1, \dots, u_n) \in [0, 1]^n. \quad (1)$$

*is called an Archimedean copula with the generator  $\phi$  if  $(-1)^k \phi^{(k)}(x) \geq 0$  for  $k = 0, \dots, n-2$  and  $(-1)^{n-2} \phi^{(n-2)}(x)$  is non-increasing and convex.*

As is well-known, the Archimedean family contains a plenty of useful copulas, including some well-known ones. For example, the Clayton copula  $C(\mathbf{u}) = (\prod_{i=1}^n u_i^{-\theta} - n + 1)^{-1/\theta}$  for  $\theta \geq 0$ , and the Ali-Mikhail-Haq (AMH) copula  $C(\mathbf{u}) = \frac{(1-\theta)\prod_{i=1}^n u_i}{\prod_{i=1}^n (1-\theta + \theta u_i) - \theta \prod_{i=1}^n u_i}$  for  $\theta \in [0, 1]$ . Note that if  $\phi(t) = e^{-t}$  then we have the independence (product) copula  $C(\mathbf{u}) = \prod_{i=1}^n u_i$ .

The following two theorems give sufficient conditions to compare the parallel and series systems with dependent components with Archimedean copula and which distributed as *PHR* models.

**Theorem 1.2.** [4] Suppose  $(X_1, \dots, X_n) \sim \text{PHR}(\bar{F}, \mathbf{a}, \phi_1)$  and  $(Y_1, \dots, Y_n) \sim \text{PHR}(\bar{F}, \mathbf{b}, \phi_2)$ . Then  $X_{n:n} \geq_{st} Y_{n:n}$  if  $\mathbf{b} \preceq^w \mathbf{a}$  and  $\psi_2 \circ \phi_1$  is super-additive<sup>1</sup>.

**Theorem 1.3.** [3] For  $\mathbf{X} \sim \text{PHR}(\bar{F}, \mathbf{a}, \phi_1)$  and  $\mathbf{Y} \sim \text{PHR}(\bar{F}, \mathbf{b}, \phi_2)$

(i) if  $\phi_1$  or  $\phi_2$  is log-convex and  $\psi_2 \circ \phi_1$  is super-additive., then  $\mathbf{b} \preceq_w \mathbf{a}$  implies  $X_{1:n} \geq_{st} Y_{1:n}$ .

(ii) if  $\phi_1$  or  $\phi_2$  is log-concave and  $\psi_2 \circ \phi_1$  is super-additive., then  $\mathbf{b} \preceq^w \mathbf{a}$  implies  $X_{1:n} \leq_{st} Y_{1:n}$ .

Motivated by these results, we want to compare the coherent systems composed of three dependent components of two different types, say type *A* and type *B*.

## 2 Main results

Consider the all 3-components coherent systems which composed of two components of type *A* and one component of type *B*. These systems represented by theirs lifetimes at Table 1. Note that System 4 is *2-out-of-3* system which works if at least 2 components work and hence its lifetimes is equal to the second order statistics of components lifetimes.

Table 1: The coherent systems with two components of type *A* and one component of type *B*

	$T$
System 1	$\min(X_1^A, X_2^A, X_3^B)$
System 2	$\min(X_1^A, \max(X_2^A, X_3^B))$
System 3	$\min(X_1^B, \max(X_2^A, X_3^A))$
System 4	$X_{2:3}$
System 5	$\max(X_1^A, \min(X_2^A, X_3^B))$
System 6	$\max(X_1^B, \min(X_2^A, X_3^A))$
System 7	$\max(X_1^A, X_2^A, X_3^B)$

Consider the second coherent system given at Table 1 and illustrated at Figure 1, composed of two types of components which may be dependent. The lifetime of the system is  $T = \min(X_1^A, \max(X_2^A, X_3^B))$ . Denote the joint reliability function of  $(X_1^A, X_2^A, X_3^B)$  by

<sup>1</sup>A function  $f$  is said to be super-additive if  $f(x + y) \geq f(x) + f(y)$

$\bar{F}(\cdot, \cdot, \cdot)$ . The minimal path sets of the system are  $P_1 = \{1, 2\}$  and  $P_2 = \{1, 3\}$ . Using the representation of system lifetime based on minimal path sets  $T = \max_{j=1 \dots r} X_{P_j}$ , we have

$$\begin{aligned} Pr(T > t) &= Pr(X_{P_1} > t) + Pr(X_{P_2} > t) - Pr(X_{P_1 \cup P_2} > t) \\ &= \bar{F}(t, t, 0) + \bar{F}(t, 0, t) - \bar{F}(t, t, t), \\ &= \hat{C}(\bar{F}_1(t), \bar{F}_1(t), 1) + \hat{C}(\bar{F}_1(t), 1, \bar{F}_2(t)) - \hat{C}(\bar{F}_1(t), \bar{F}_1(t), \bar{F}_2(t)) \end{aligned}$$

where, the last equality follows from the (Sklar's) Theorem 1.1 and  $\hat{C}$  is the survival copula.

Now suppose that the random variables have the Archimedean survival copula with generator  $\phi$ , then we have

$$Pr(T > t) = \phi(2\psi(\bar{F}_1(t))) + \phi(\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t))) - \phi(2\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t)))$$

This representation for reliability functions of systems are given at Table 2.

Table 2: The reliability function of systems lifetime

	$Pr(T > t)$
System 1	$\phi(2\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t)))$
System 2	$\phi(\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t))) + \phi(2\psi(\bar{F}_1(t))) - \phi(2\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t)))$
System 3	$2\phi(\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t))) - \phi(2\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t)))$
System 4	$\phi(2\psi(\bar{F}_1(t))) + 2\phi(\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t))) - 2\phi(2\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t)))$
System 5	$\phi(\psi(\bar{F}_1(t))) + \phi(\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t))) - \phi(2\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t)))$
System 6	$\phi(\psi(\bar{F}_2(t))) + \phi(2\psi(\bar{F}_1(t))) - \phi(2\psi(\bar{F}_1(t)) + \psi(\bar{F}_2(t)))$
System 7	$1 - \phi(2\psi(1 - \bar{F}_1(t)) + \psi(1 - \bar{F}_2(t)))$

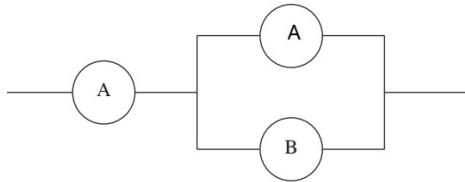


Figure 1: System 2

At the continue, we suppose that the components follows a *PHR* model with baseline reliability function  $\bar{F}$  and proportional hazard rate vector  $\mathbf{a} = (a_1, a_2)$ . For getting the results, let

$$A(\mathbf{a}, s, \phi) = \phi(\psi(s^{a_1}) + \psi(s^{a_2})) + \phi(2\psi(s^{a_1})) - \phi(2\psi(s^{a_1}) + \psi(s^{a_2})). \quad (2)$$

$A(\mathbf{a}, s, \phi)$  is decreasing in  $a_i$ ,  $i = 1, 2$  for all  $0 < s < 1$  and all generating function  $\phi$ .

*Proof.*

$$\begin{aligned} \frac{\partial A(\mathbf{a}, s, \phi)}{\partial a_1} &= \psi'(s^{a_1}) \ln(s) s^{a_1} [\phi'(\psi(s^{a_1}) + \psi(s^{a_2})) + 2\phi'(2\psi(s^{a_1})) - 2\phi'(2\psi(s^{a_1}) + \psi(s^{a_2}))] \\ &= {}^{sgn} \phi'(\psi(s^{a_1}) + \psi(s^{a_2})) + [2\phi'(2\psi(s^{a_1})) - 2\phi'(2\psi(s^{a_1}) + \psi(s^{a_2}))] \\ &\leq 0, \end{aligned}$$

since,  $\phi'$  (and hence  $\psi'$ ) is non-positive and non-decreasing.

Similarly,

$$\frac{\partial A(\mathbf{a}, s, \phi)}{\partial a_2} = \psi'(s^{a_1}) \ln(s) s^{a_1} [\phi'(\psi(s^{a_1}) + \psi(s^{a_2})) - \phi'(2\psi(s^{a_1}) + \psi(s^{a_2}))] \leq 0$$

□

$A(\mathbf{a}, s, \phi)$  is Schur-concave (Schur-convex) in  $\mathbf{a}$  if  $a_1 > a_2$  ( $a_1 < a_2$ ) and  $\phi$  is log-convex (log-concave) for all  $0 < s < 1$  and all generating function  $\phi$ .

*Proof.* For this aim, first note that we have

$$\begin{aligned} \frac{\partial A}{\partial a_1} - \frac{\partial A}{\partial a_2} &= \ln(s) \left( \frac{s^{a_1}}{\phi'(\psi(s^{a_1}))} (\phi'(W_1) + 2\phi'(W_2) - 2\phi'(W_3)) - \frac{s^{a_2}}{\phi'(\psi(s^{a_1}))} (\phi'(W_1) - \phi'(W_3)) \right) \\ &= \ln(s) \left( \frac{\phi(\psi(s^{a_1}))}{\phi'(\psi(s^{a_1}))} (\phi'(W_1) + 2\phi'(W_2) - 2\phi'(W_3)) - \frac{\phi(\psi(s^{a_2}))}{\phi'(\psi(s^{a_1}))} (\phi'(W_1) - \phi'(W_3)) \right) \end{aligned}$$

where,  $W_1 = \psi(s^{a_1}) + \psi(s^{a_2})$ ,  $W_2 = 2\psi(s^{a_1})$ ,  $W_3 = 2\psi(s^{a_1}) + \psi(s^{a_2})$ .

For  $a_1 > a_2$  ( $a_1 < a_2$ ) and  $0 < s < 1$  we have  $\psi(s^{a_1}) \geq (\leq) \psi(s^{a_2})$ . Then log-convexity (log-concavity) of  $\phi$  implies that

$$\frac{\phi(\psi(s^{a_1}))}{\phi'(\psi(s^{a_1}))} \leq \frac{\phi(\psi(s^{a_2}))}{\phi'(\psi(s^{a_2}))}.$$

It is evident that  $W_1 \leq W_3$  and  $W_2 \leq W_3$ , then from the fact that  $\phi'$  is increasing (which implies that  $2\phi'(W_2) - \phi'(W_3)$  is negative) we have

$$\begin{aligned} \frac{\phi(\psi(s^{a_1}))}{\phi'(\psi(s^{a_1}))} (\phi'(W_1) + 2\phi'(W_2) - 2\phi'(W_3)) - \frac{\phi(\psi(s^{a_2}))}{\phi'(\psi(s^{a_1}))} (\phi'(W_1) - \phi'(W_3)) &\geq \\ \frac{\phi(\psi(s^{a_1}))}{\phi'(\psi(s^{a_1}))} (2\phi'(W_2) - \phi'(W_3)) &\geq 0 \end{aligned}$$

This result, implies that  $(a_1 - a_2)(\frac{\partial A(\mathbf{a}, s, \phi)}{\partial a_1} - \frac{\partial A(\mathbf{a}, s, \phi)}{\partial a_2})$  is non-positive (non-negative). Therefore, the desired result follows based on Lemma 1.

□

**Theorem 2.1.** Consider the coherent system illustrated at Figure 1. Suppose that the components distributions follow the model  $(X_1^A, X_2^A, X_3^B) \sim PHR(\bar{F}, \mathbf{a}, \phi)$  and denote the system lifetime with  $T_1$ . Again, consider the system with components following the  $(Y_1^A, Y_2^A, Y_3^B) \sim PHR(\bar{F}, \mathbf{b}, \phi)$  and denote the system lifetime in this case by  $T_2$ . Then

- (i)  $T_1 \leq_{st} T_2$  if  $\mathbf{a} \preceq^w \mathbf{b}$ ,  $a_1 < a_2$ ,  $b_1 < b_2$  and  $\phi$  is log-concave.
- (ii)  $T_1 \geq_{st} T_2$  if  $\mathbf{a} \preceq_w \mathbf{b}$ ,  $a_1 > a_2$ ,  $b_1 > b_2$  and  $\phi$  is log-convex.

*Proof.* (i) From Lemmas 2 and 2 we have that  $A(\mathbf{a}, s, \phi)$  is Schur-convex in  $\mathbf{a}$  and decreasing in  $a_i$ , then from Lemma 1(ii) we have under the assumption of theorem  $A(\mathbf{a}, s, \phi) \leq A(\mathbf{b}, s, \phi)$  for all  $0 < s < 1$  and generator function  $\phi$  which implies that  $T_1 \leq_{st} T_2$ .

(ii) As  $A(\mathbf{a}, s, \phi)$  is Schur-concave then  $-A(\mathbf{a}, s, \phi)$  is Schur-convex. Also,  $-A(\mathbf{a}, s, \phi)$  is increasing in  $a_i$ , then using Lemma 1(i) the assumptions of theorem implies that  $-A(\mathbf{a}, s, \phi) \leq -A(\mathbf{b}, s, \phi)$  and we get the  $T_1 \geq_{st} T_2$ . □

Table 3: The behavior of functions

Generator	
$\phi(t) = (\theta t + 1)^{-1/\theta}$ (Clayton)	log-convex for $\theta > 0$
$\phi(t) = \frac{1-\theta}{\exp(t)-\theta}$ (AMH)	log-convex for $\theta > 0$ and log-concave for $\theta < 0$
$\phi(t) = \exp(-t^{1/\theta})$	log-convex for $\theta \geq 1$
$\phi(t) = 1 - (1 - \exp(-t))^{1/\theta}$	log-convex for $\theta \geq 1$
$\phi(t) = \exp(\frac{1-\exp(t)}{\theta})$	log-concave for $\theta \in (0, 1]$

At Table 3, the behavior of some generating functions based on  $t$  are given.

**Theorem 2.2.** *The results obtained at Theorem 2.1 hold similarly, for System 1, System 4 and System 5.*

Also, we have the following result for System 1.

**Theorem 2.3.** *Under the frame of Theorem 2.1 for System 1, suppose that  $(X_1^A, X_2^A, X_3^B) \sim PHR(\bar{F}, \mathbf{a}, \phi_1)$  and  $(Y_1^A, Y_2^A, Y_3^B) \sim PHR(\bar{F}, \mathbf{b}, \phi_2)$ . Then*

- (i)  $T_1 \leq_{st} T_2$  if  $\mathbf{a} \preceq^w \mathbf{b}$ ,  $a_1 < a_2, b_1 < b_2$ ,  $\phi_1$  or  $\phi_2$  is log-concave, and  $\psi_2 \circ \phi_1$  is super-additive.
- (ii)  $T_1 \geq_{st} T_2$  if  $\mathbf{a} \preceq_w \mathbf{b}$ ,  $a_1 > a_2, b_1 > b_2$  and  $\phi_1$  or  $\phi_2$  is log-convex, and  $\psi_1 \circ \phi_2$  is super-additive.

*Proof.* See [3]. □

The proofs of the following two theorems are similar to Theorem 2.1 and then the details are omitted.

**Theorem 2.4.** *Consider the System 3 at Table 1. Suppose that the components distributions follows the model  $(X_1^A, X_2^A, X_3^B) \sim PHR(\bar{F}, \mathbf{a}, \phi)$  and denote the system lifetime with  $T_1$ . Again, consider the system with components following the  $(Y_1^A, Y_2^A, Y_3^B) \sim PHR(\bar{F}, \mathbf{b}, \phi)$  and denote the system lifetime in this case by  $T_2$ . Then*

- (i)  $T_1 \leq_{st} T_2$  if  $\mathbf{a} \preceq^w \mathbf{b}$ ,  $a_1 > a_2, b_1 > b_2$  and  $\phi$  is log-concave.
- (ii)  $T_1 \geq_{st} T_2$  if  $\mathbf{a} \preceq_w \mathbf{b}$ ,  $a_1 < a_2, b_1 < b_2$  and  $\phi$  is log-convex.

**Theorem 2.5.** *Consider the System 6 at Table 1. Suppose that the components distributions follows the model  $(X_1^A, X_2^A, X_3^B) \sim PHR(\bar{F}, \mathbf{a}, \phi)$  and denote the system lifetime with  $T_1$ . Again, consider the system with components following the  $(Y_1^A, Y_2^A, Y_3^B) \sim PHR(\bar{F}, \mathbf{b}, \phi)$  and denote the system lifetime in this case by  $T_2$ . Then*

- (i)  $T_1 \leq_{st} T_2$  if  $\mathbf{a} \preceq^w \mathbf{b}$ ,  $a_1 > a_2, b_1 > b_2$  and  $\phi$  is log-concave.
- (ii)  $T_1 \geq_{st} T_2$  if  $\mathbf{a} \preceq_w \mathbf{b}$ ,  $a_1 < a_2, b_1 < b_2$  and  $\phi$  is log-concave.

## References

- [1] Coolen, F. P. A. and Coolen-Maturi, T. (2012). On generalizing the signature to systems with multiple types of components. In Complex Systems and Dependability, eds W. Zamojski et al., Springer, Berlin, 115-130.

- [2] Erylimaz, S. (2015). Systems composed of two types of nonidentical and dependent components, *Naval Research Logistics*, **62**(5), 388-394.
- [3] Fang, R., Li, C. and Li, X. (2016). Stochastic comparisons on sample extremes of dependent and heterogenous observations, *Statistics*, DOI 10.1080/02331888.2015.1119151.
- [4] Li, X., Fang, R. (2015). Ordering properties of order statistics from random variables of Archimedean copulas with applications, *Journal of Multivariate Analysis*, **133**, 304320.
- [5] Navarro, J., Torrado, N., Aguila, Y. D., (2017), Comparisons Between Largest Order Statistics from Multiple-outlier Models with Dependence, *Methodol Comput Appl Probab*, DOI 10.1007/s11009-017-9562-7.
- [6] Samaniego F.J. and Navarro J. (2016). On comparing coherent systems with heterogeneous components, *Advances in Applied Probability*, **48**, 124.
- [7] Zhao, P. and Balakrishnan, N. (2013). Ordering properties of order statistics from heterogeneous populations: a review with an emphasis on some recent developments, *Probability in the Engineering and Informational Sciences*, **27**, 403-443



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## Testing exponentiality based on the residual Lin-Wong divergence

Khalili, M.<sup>1</sup>      Habibirad, A.<sup>2</sup>      Yousefzadeh, F.<sup>3</sup>

<sup>1,2</sup> Department of Statistics, Ferdowsi University of Mashhad

<sup>3</sup> Department of Statistics, Faculty of Mathematics and Statistics, University of Birjand

### Abstract

Testing exponentiality has long been an interesting issue in statistical inferences. This article is based on a modified measure of distance between two distributions. New measure is similar Kullback Leibler divergence and it is related to the Lin Wong divergence applied on the residual lifetime data. Modified measure, a consistent test statistic for testing the hypothesis of exponentiality against some alternatives is developed. In order to estimate the divergence, we first consider a method similar to Vasicek's method for estimating the Shannon entropy. Then, critical values of the test are computed by Monte Carlo simulation. Finally, we found that the powers differences between the proposed test and other tests. It is shown that the proposed test is better than other tests of exponentiality, when hazard rate function is increasing. Finally, the use of the proposed test is shown in two illustrative examples.

**Keywords:** Exponentiality test, Goodness of fit testing, Kullback Leibler divergence, Lin Wong divergence, Residual lifetime data.

---

<sup>1</sup>khalilimohadeseh@gmail.com

<sup>2</sup>ahabibi@um.ac.ir

<sup>3</sup>fyousefzadeh@birjand.ac.ir

# 1 Introduction

In the recent past, it is interesting for many researchers that study about the measurement of uncertainty associated with a probability distribution. Of particular interest in probability and statistics is the notion of entropy, introduced by Shannon in 1948. If  $X$  is a random variable having an absolutely continuous distribution function  $G(x)$  with probability density function (pdf)  $g(x)$ , then the entropy of the random variable  $X$  is defined as

$$H(g) = - \int_0^\infty g(x) \log g(x) dx. \quad (1)$$

Hallmark applications are divergence measurement and distance information extraction. A number of divergence measures have been already proposed and extended to approach this task.

Let  $g(x)$  and  $f(x)$  be pdf of variables  $X$  and  $Y$  respectively. The Kullback Leibler (KL) information of  $g(x)$  and  $f(x)$  has been defined by Kullback and Leibler in 1951, based on Shannon entropy as

$$D_{\text{KL}}(g\|f) = \int_0^\infty g(x) \log \frac{g(x)}{f(x)} dx. \quad (2)$$

In this paper, we are focusing the Lin Wong (LW) divergence measures. The LW divergence is a prominent measure of divergence in this realm. The LW divergence measure was defined by Lin and Wong (1990) as

$$D_{\text{LW}}(g\|f) = \int_0^\infty g(x) \log \frac{g(x)}{\frac{1}{2}(g(x) + f(x))} dx. \quad (3)$$

Observing that if a unit has survived up to time  $t$ ,  $H(g)$  is not a useful tool for measuring the uncertainty about the remaining lifetime of the unit. Consider a system  $X$  that has survived upto time  $t$ . In order to calculate the uncertainty about the residual life of such a system, equation (2) is not appropriate. Ebrahimi in 1996, introduced a new measure to ascertain the uncertainty about the residual life of a random variable  $X_t = (X|X \geq t)$  given by

$$H(g_t) = - \int_t^\infty \bar{g}_t(x) \log \bar{g}_t(x) dx, \quad (4)$$

in which,  $X$  denotes the failure time for a system of living organism or a component of it and  $\bar{g}_t(x) = \frac{g(x)}{\bar{G}(t)}$  represents the failure density function, where  $G(t) = P(X \leq t)$  the cumulative distribution function (cdf) and  $\bar{G}(t) = 1 - G(t)$ .

Furthermore, the KL information based on Shannon entropy for the residual lifetime which defined by Ebrahimi and Kirmani in 1996, follows

$$D_{\text{KL}}^{(t,\infty)}(g\|f) = \int_t^\infty \bar{g}_t(x) \log \left( \frac{\bar{g}_t(x)}{\bar{f}_t(x)} \right) dx. \quad (5)$$

In social studies, engineering, medical sciences, reliability studies and management science, it is very important to know whether the underlying data follow a particular distribution. So, many authors were interested in goodness of fit tests. Testing exponentiality still attracts considerable attention and is the topic of a good amount of recent research. Many authors provide test statistics for detecting departures from the hypothesis of exponentiality against specific or general alternatives. For example, see Baratpour and Habibirad (2012) and Alizade and Arghami (2011b).

In the following we illustrate the testing exponentiality based on the LW Information with the residual lifetime data. Before, Abbasnejad et al. (2012) improved the applications of LW measure by proposing a new goodness of fit measure to assess the fitness of exponential distributions with data.

## 2 Test statistics

Based on Khalili et al. (2017), by using the equation (3), the LW divergence on the residual lifetime data is stated as below

$$D_{\text{LW}}^{(t,\infty)}(g\|f) = \int_t^\infty \bar{g}_t(x) \log \frac{\bar{g}_t(x)}{\frac{1}{2}(\bar{g}_t(x) + \bar{f}_t(x))} dx.$$

Many researchers presented the goodness of fit tests based on various entropy estimators. Among these various entropy estimators, Vasicek's sample entropy has been most widely used in goodness of fit tests. Let  $X_1, X_2, \dots, X_n$  be a random sample from a continuous distribution  $G$ . Using  $G(x) = p$ , Vasicek in 1976 expressed equation (2) as

$$H(g) = \int_0^1 \log\left(\frac{d}{dp}G^{-1}(p)\right) dp,$$

and by replacing the distribution function  $G$  by the empirical distribution function  $G_n$  and using a difference operator instead of the differential operator, the derivative of  $G^{-1}(p)$  was estimated by  $\frac{X_{(i+m)} - X_{(i-m)}}{2m}$ . Therefore  $H(g)$  was estimated as

$$V_{mn} = \frac{1}{n} \sum_{i=1}^n \log\left(\frac{n}{2m}(X_{(i+m)} - X_{(i-m)})\right),$$

where,  $m$  is a positive integer with values smaller than  $n/2$ , while  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$  are order statistics satisfying the condition  $X_{(i)} = X_{(1)}$  if  $i < 1$ ,  $X_{(i)} = X_{(n)}$  if  $i > n$ . Vasicek proved that

$$V_{mn} \xrightarrow{pr.} H(g) \text{ as } n \rightarrow \infty, m \rightarrow \infty, \frac{m}{n} \rightarrow 0.$$

Based on Maximum Likelihood Estimating (MLE), estimator the parameter  $\theta$  in exponential distribution on residual lifetime data is given by

$$\hat{\theta} = \frac{1}{\bar{x} - t}. \tag{6}$$

*Proof.* Noting that  $g_t(x) = \frac{\theta e^{-\theta x}}{e^{-\theta t}}$ , the proof is clear.  $\square$

Let  $X_1, \dots, X_n$  be a random sample from a continuous non-negative cdf  $F(x)$  with a pdf  $f(x)$ . Consider the following hypotheses

$$H_0 : f(x; \theta) = f_0(x; \theta),$$

while the alternative hypotheses would be

$$H_1 : f(x; \theta) \neq f_0(x; \theta),$$

where  $f_0(x; \theta) = \theta e^{-\theta x}$ ,  $x > 0$  and  $\theta > 0$  is unknown. It motivates us to use  $D_{\text{LW}}^{(t, \infty)}(f \| f_0)$  as a test statistic for exponentiality. Note that the evaluation of requires the knowledge of  $G$ , which is not operational. Instead, we consider a sample estimate of  $D_{\text{LW}}^{(t, \infty)}(f \| f_0)$ .

The LW divergence on the residual data in favor of  $f(x)$  against  $f_0(x)$  is

$$D_{\text{LW}}^{(t, \infty)}(f \| f_0) = \int_t^\infty \bar{f}_t(x) \log \frac{\bar{f}_t(x)}{\frac{1}{2}(\bar{f}_t(x) + \frac{\theta e^{-\theta x}}{e^{-t\theta}})} dx, \quad (7)$$

under the null hypothesis  $D_{\text{LW}}^{(t, \infty)}(f \| f_0) = 0$  and large values of 0 favor  $H_1$ .

To estimate  $D_{\text{LW}}^{(t, \infty)}(f \| f_0)$ , we use two following methods. In the first method, using  $\bar{F}_t(x) = P$ , similar to Vasicek's method and using difference operator in place of the differential operator, we express equation (7) as

$$\int_0^1 \log \frac{2(-\frac{dF_t^{-1}(p)}{dp})^{-1}}{(-\frac{dF_t^{-1}(p)}{dp})^{-1} + \frac{\theta e^{-\theta x}}{e^{-t\theta}}} dp.$$

There are vast types of estimators for  $\frac{dF_t^{-1}(p)}{dp}$ . Amongst various functions which are proposed for estimating the entropy, Vasicek's sample entropy method has been extensively implemented by scholars when developing statistical procedures based on entropy concept. Based on Vasicek in 1976, we followed estimator for the density quantile function as bellow

$$\frac{2m}{n_0(X_{(i+m)} - X_{(i-m)})} \simeq (\frac{dF_t^{-1}(p)}{dp})^{-1}, \quad (8)$$

where,  $X_{(i)}$  is the ordered statistic and the window size  $m$  is a positive integer with values smaller than  $n/2$ . We get an estimator of  $D_{\text{LW}}^{(t, \infty)}(f \| f_0)$  as

$$\hat{D}_{\text{LW}}^{(t, \infty)}(f \| f_0) = -\frac{1}{n_0} \sum_{i=1}^{n_0} \log \left( \frac{1}{2} + \frac{(X_{(i+m)} - X_{(i-m)})n_0(\hat{\theta}e^{-X_{(i)}\hat{\theta}})}{4m(e^{-t\hat{\theta}})} \right), \quad (9)$$

where,  $m$  is a positive integer with values smaller than  $\frac{n_0}{2}$ , while  $n_0 = \#X_i \geq t$  and  $t \leq X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n_0)}$  are order statistics satisfying the condition  $X_{(i)} = X_{(1)}$  if  $i < 1$ ,  $X_{(i)} = X_{(n_0)}$  if  $i > n_0$ . In equation (9), we used  $\hat{\theta}$  that is MLE of  $\theta$ .

In order to obtain a test statistic, we use the following theorems.

**Theorem 2.1.** Let  $X_1, \dots, X_n$  denote a sample from the continuous distribution with pdf  $g(x)$  concentrated on  $(0, +\infty)$ . Then for any significance level  $\alpha \in (0, 1)$  goodness of fit test based on the test statistic  $\hat{D}_{LW}^{(t, \infty)}(g||f)$  is a consistent test for the hypotheses testing problem.

*Proof.* Suppose  $\exists X_i \in (X_{(i+m)}, X_{(i-m)})$ , in which  $\frac{G(X_{(i+m)}) - G(X_{(i-m)})}{(X_{(i+m)} - X_{(i-m)})} \simeq g(X_i)$ , and  $n \rightarrow \infty, m \rightarrow \infty, \frac{m}{n} \rightarrow 0$ . Then, converging the distribution function  $G$  by the empirical distribution function  $G_n$ , we have  $\frac{2m}{n(X_{(i+m)} - X_{(i-m)})} \xrightarrow{p} g(X_i)$ , where convergence is in probability. Therefore,

$$\frac{1}{n} \sum_{i=1}^n \log \left( 1 + \frac{n f(X_i) (X_{(i+m)} - X_{(i-m)})}{2m} \right) \xrightarrow{p} E_g \left( \log \left( 1 + \frac{f(X_i)}{g(X_i)} \right) \right). \quad (10)$$

Hence, using (10) the test statistic  $\hat{D}_{LW}^{(t, \infty)}(g||f)$  is a consistent test.

The next theorem states that the scale of the random variable  $X$  has no effect on the accuracy of  $\hat{D}_{LW}^{(t, \infty)}(g||f)$ .  $\square$

**Theorem 2.2.** Suppose that  $X_1, \dots, X_n$  be a sequence of iid random variables from a continuous exponential distribution  $G(x)$  with a density function  $g(x)$  on the residual lifetime data and  $t_X$  and  $t_Y$  be the time point for the random variables  $X$  and  $Y$ , respectively. Put  $t_Y = kt_X$ ,  $Y_i = kX_i$ ,  $i = 1, \dots, n$ , where  $k > 0$ . Let  $\hat{D}_{LW_X}^{(t_X, \infty)}(g||f)$  and  $\hat{D}_{LW_Y}^{(t_Y, \infty)}(g||f)$  be the estimators for  $D_{LW_X}^{(t_X, \infty)}(g||f)$  and  $D_{LW_Y}^{(t_Y, \infty)}(g||f)$  on the residual lifetime data, respectively. Then, we can see the test statistic is invariant with respect to scale transformations and the following properties hold

- (i)  $E(\hat{D}_{LW_X}^{(t_X, \infty)}(g||f)) = E(\hat{D}_{LW_Y}^{(t_Y, \infty)}(g||f))$ ,
- (ii)  $Var(\hat{D}_{LW_X}^{(t_X, \infty)}(g||f)) = Var(\hat{D}_{LW_Y}^{(t_Y, \infty)}(g||f))$ ,
- (iii)  $MSE(\hat{D}_{LW_X}^{(t_X, \infty)}(g||f)) = MSE(\hat{D}_{LW_Y}^{(t_Y, \infty)}(g||f))$ .

*Proof.* From the right-hand side of Equation (6), we have

$$\hat{\theta}_Y = \frac{1}{\bar{Y} - t_Y} = \frac{1}{k\bar{X} - kt_X} = \frac{1}{k}\hat{\theta}_X. \quad (11)$$

So, we can get  $\hat{D}_{LW_X}^{(t_X, \infty)}(g||f) = \hat{D}_{LW_Y}^{(t_Y, \infty)}(g||f)$ , and the proof is complete.  $\square$

Similar to  $\hat{D}_{LW}^{(t, \infty)}(g||f)$  in equation (9),  $\hat{D}_{KL}^{(t, \infty)}(g||f)$  for the KL divergence on residual lifetime data is given as

$$\hat{D}_{KL}^{(t, \infty)}(g||f) = \frac{1}{n_0} \sum_{i=1}^{n_0} \log \left( \frac{2m(e^{-t\hat{\theta}})}{(X_{(i+m)} - X_{(i-m)})n_0(\hat{\theta}e^{-X_{(i)}\hat{\theta}})} \right). \quad (12)$$

Also, similar to Theorem 2.1,  $D_{KL}^{(t, \infty)}$  has scale invariant and it is consistent. Furthermore, the Kolmogorov Smirnov (KS) statistic belongs to the supremum class of empirical distribution function (EDF) statistics and this class of statistics is based on the largest vertical difference between the hypothesized and empirical distribution. Given  $n$  ordered

data points,  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ , Conover in 1999 defined the test statistic proposed by Kolmogorov in 1933, as  $T = \sup_x |F^*(x) - F_n(x)|$ , where ‘sup’ stands for supremum which means the greatest.  $F^*(x)$  is the hypothesized distribution function whereas  $F_n(x)$  is the EDF estimated based on the random sample. So, the modifications to the KS statistic for the residual lifetime data, may be written as

$$KS^{(t,\infty)} = \max_{1 \leq i \leq n_0} \left\{ \frac{i}{n_0} - F_t^*(X_i), F_t^*(X_i) - \frac{i-1}{n_0} \right\}, \quad (13)$$

where  $n_0 = \#X_i \geq t$ , and  $F_t^*(X_i) = \frac{F(x_i) - F(t)}{\bar{F}(t)} = \frac{\bar{F}(t) - \bar{F}(X_i)}{\bar{F}(t)} = 1 - \frac{e^{-x_i \hat{\theta}}}{e^{-t \hat{\theta}}}$ .

Anderson and Darling (1954) defined the statistic for this test as

$$W_n^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) \{ \log F^*(X_i) + \log(1 - F^*(X_{n+1-i})) \},$$

where  $F^*(x)$  is the cumulative distribution function of the specified distribution,  $X_{(i)}$ 's are the ordered data and  $n$  is sample size. Therefore, the modified Anderson and Darling (AD) statistic for the residual lifetime data can be rewritten as

$$AD^{(t,\infty)} = -n_0 - \frac{1}{n_0} \sum_{i=1}^{n_0} (2i-1) \{ \log F_t^*(X_i) + \log(1 - F_t^*(X_{n_0+1-i})) \}. \quad (14)$$

The Cramer Von Mises (CVM) test was developed by Cramer (1928), Von Mises (1931). The CVM statistic is as below

$$CVM = \frac{1}{12n} + \sum_{i=1}^n [F^*(x_{(i)}) - \frac{2i-1}{2n}]^2.$$

So, The CVM statistic on the residual lifetime data can be computed as

$$CVM^{(t,\infty)} = \frac{1}{12n_0} + \sum_{i=1}^{n_0} [F_t^*(X_i) - \frac{2i-1}{2n_0}]^2. \quad (15)$$

### 3 Simulation study

A simulation study is performed to analyze the behavior of the proposed test statistic. We determine the critical points using Mont Carlo simulation. For this end, 10000 samples, with the different sample sizes ( $n = 5, 15, \dots, 50$ ), cutting points ( $F(t) = .25, .75$ ) and the significance levels ( $\alpha = .05$ ), were generated for remaining data which followed the exponential distribution with the mean 1. Table 1 gives the critical values of  $\hat{D}_{LW}^{(t,\infty)}(g||f)$  for various sample sizes. Note that the test statistics is considered as the basis for comparison, hence the critical values would not depend the unknown mean parameter. We compute the powers of the test based on  $\hat{D}_{LW}^{(t,\infty)}(g||f)$  statistic by Mont Carlo simulation. For comparison of the power of the proposed test with powers of the tests published, we consider the alternatives according to the type of hazard function as follows:

- a) Increasing Hazard Rate (IHR) including: Gamma:  $G(3,1)$ , Weibull:  $W(3,1)$  and Generalized Exponential:  $GE(3,1)$ .
- b) Decreasing Hazard Rate (DHR) including: Gamma:  $G(0.5,1)$  , Weibull:  $W(0.5,1)$  and Generalized-Exponential:  $GE(.5,1)$ .
- c) Nonmonotone Hazard Rate (NHR) including: Log-Normal:  $LN(0,.5)$ , Log-Laplace:  $LL(0,.5)$  and Generalized-Logistic:  $GL(0.5,1)$ .

The simulation results are summarized in Tables 2 and 3, by taking the proportion of rejections. The values indicated as bold in each row, show that the pertaining statistics has the highest power when compared by other tests on that row.

Table 1: Critical values of  $D_{LW}^{(t,\infty)}(g||f)$ , in  $\alpha = .05$  for various  $\alpha, F(t)$

F(t)	n=5	n=10	n=15	n=20	n=25	n=30	n=35	n=40	n=45	n=50
0.25	0.3993	0.2335	0.1751	0.1376	0.1176	0.0998	0.0908	0.0800	0.0748	0.0681
0.75	0.3907	0.2357	0.1735	0.1382	0.1172	0.1006	0.0902	0.0820	0.0746	0.0683

Table 2: Power comparison for the tests, when  $F(t) = .25, \alpha = .05$

n	Test	IHR			DHR			NHR		
		G(3)	W(3)	GE(3)	G(.5)	W(.5)	GE(.5)	LN(.5)	LL(.5)	GL(.5)
10	LW	<b>0.5206</b>	<b>0.7359</b>	<b>0.3137</b>	0.0276	0.0056	0.0271	0.0365	0.1166	<b>0.1354</b>
	KL	0.4591	0.7019	0.276	0.0341	0.0258	0.0309	0.0428	0.1498	0.1283
	KS	0.3309	0.5407	0.1958	0.063	0.2669	0.0637	0.1008	0.2432	0.0869
	CM	0.3819	0.644	0.2176	0.0639	0.2925	0.064	0.1031	<b>0.248</b>	0.0929
	AD	0.3145	0.5656	0.1658	<b>0.0828</b>	<b>0.356</b>	<b>0.0834</b>	<b>0.1057</b>	0.2231	0.0722
20	LW	<b>0.7711</b>	<b>0.9556</b>	<b>0.4703</b>	0.0191	0.001	0.0219	0.0258	0.0951	<b>0.1876</b>
	KL	0.6645	0.9328	0.3831	0.0309	0.1315	0.0327	0.0528	0.2083	0.1491
	KS	0.595	0.8592	0.337	0.0759	0.4797	0.0693	0.1393	0.3681	0.1231
	CM	0.6906	0.9373	0.3907	0.0752	0.535	0.0712	<b>0.1523</b>	<b>0.3941</b>	0.1403
	AD	0.6511	0.9137	0.3443	<b>0.0891</b>	<b>0.5815</b>	<b>0.0872</b>	0.1491	0.3651	0.1099

Table 3: Power comparison for the tests, when  $F(t) = .75, \alpha = .05$

n	Test	IHR			DHR			NHR		
		G(3)	W(3)	GE(3)	G(.5)	W(.5)	GE(.5)	LN(.5)	LL(.5)	GL(.5)
10	LW	<b>0.1535</b>	<b>0.0875</b>	<b>0.0704</b>	0.0403	0.0109	0.0374	0.0201	0.0095	<b>0.092</b>
	KL	0.1383	0.0795	0.0662	0.0422	0.0188	0.0381	0.0245	0.0676	0.0861
	KS	0.0928	0.0583	0.0496	0.0456	0.1537	0.0473	0.1063	0.2985	0.064
	CM	0.1032	0.0635	0.0523	0.0468	0.1754	0.0467	0.1176	0.325	0.0669
	AD	0.0741	0.0492	0.0456	<b>0.051</b>	<b>0.2097</b>	<b>0.05</b>	<b>0.1369</b>	<b>0.3479</b>	0.0535
20	LW	<b>0.2276</b>	<b>0.1064</b>	<b>0.0849</b>	0.0366	0.0045	0.0395	0.0119	0.0039	<b>0.115</b>
	KL	0.1807	0.0954	0.0724	0.0426	0.0569	0.04	0.0384	0.2591	0.0998
	KS	0.1661	0.0815	0.0668	<b>0.0582</b>	0.2907	0.0534	0.182	0.5522	0.0866
	CM	0.1845	0.0822	0.067	0.0561	0.327	0.0569	0.204	0.5839	0.0919
	AD	0.1477	0.0693	0.0585	0.0568	<b>0.3605</b>	<b>0.0572</b>	<b>0.2217</b>	<b>0.5914</b>	0.075

## 4 Illustrative examples

In this section, we consider the proposed goodness of fit test with real data. We present two examples to illustrate the use of the test  $\hat{D}_{LW}^{(t,\infty)}(g||f)$  for testing the validity of Exponential

distribution.

**Example 4.1.** The data are given below that consist of amount of lifetime for 21 tools from Meaker (1987). A residual truncated reminder sample is obtained in terms of  $t$  value which is determined from the  $F(t) = .25$  based on the real sample. For this task, we obtained the truncated point  $t = 45.6$  equivalent to quantile (.25). So 6 out of 21 observations would have been emitted while 15 data points remained for study (see Table 4). In this way, we can examine the critical value of  $\alpha = .05$  and  $n_0 = 15$  using the sample p-value and the LW statistic value for our sample.

Table 4: Amount of lifetime for 21 tools in Example 4.1.

17.88	28.92	23	41.52	42.42	45.6	48.4	51.84	51.96	54.12	55.56
67.8	68.64	68.88	84.12	93.12	98.64	105.84	127.92	128.04	173.4	

Table 5 shows critical values, test statistics and the p-values. Since the values of  $\hat{D}_{LW}^{(t,\infty)}(g||f)$  are less than the critical values (or p-values are larger than  $\alpha$ ), test accepts the null hypothesis that data follow an exponential distribution at significance levels  $\alpha = 0.05$ .

**Example 4.2.** The normal distribution is considered in this example. Table 6 shows the yearly amounts of rain in New York city between 1910 and 1973 (Parzen, 1979). In this case, we have  $n = 63$ ,  $n_0 = 47$ ,  $t = \text{quantile}(.25) = 61.95$ , equivalently  $F(61.95) = 0.25$ . Result of goodness of fit test for exponential distributions are presented in Table 7.

Table 5: Critical values, test statistics, and the p-values in Example 4.1

Significance level	Critical value	$D_{LW}^{(t,\infty)}(g  f)$	p-value
$\alpha = .05$	0.1751	0.0771	0.4946

Table 6: Yearly rain precipitation in New York city in Example 4.2.

42.5	25.0	39.8	39.9	40.1	46.7	49.1	49.6	51.1	51.6	52.5
54.7	55.5	58.9	58.0	60.3	63.6	65.4	66.1	69.3	70.9	71.4
71.5	71.8	72.9	74.4	76.2	77.8	78.1	78.4	79.0	79.3	79.6
80.7	82.4	82.4	83.0	83.6	83.6	84.8	85.5	87.4	88.7	89.6
89.8	89.9	90.9	97.0	98.3	101.4	102.4	103.9	104.5	105.2	110.0
110.5	110.5	113.7	114.5	115.6	120.7	124.7	126.4			

Table 7: Critical values, test statistics, and the p-values in Example 4.2

Significance level	Critical value	$D_{LW}^{(t,\infty)}(g  f)$	p-value
$\alpha = .05$	0.0717	0.1406	0.0000

## 5 Conclusions

In this article, on the based of LW divergence on the general data, modified measures are introduced when applied on the residual lifetime data. Furthermore, in the present study we construct a consistent goodness of fit test for exponential distribution on LW divergence measure on the residual lifetime data. By a simulation study, the powers of the proposed tests were computed under several alternatives and different sample sizes. It is shown that,  $\hat{D}_{LW}^{(t,\infty)}(g||f)$  test on the residual lifetime data is better than the other test in IHR function. Also,  $\hat{D}_{AD}^{(t,\infty)}(g||f)$  test is more power than other tests in DHR function.

## References

- [1] Abbasnejad, M., Arghami, N.R., Tavakoli, M. (2012). A goodness of fit test for exponentiality based on
- [2] Alizadeh N.H., Arghami, N.R. (2011b). Testing exponentiality based on characterizations of the exponential distribution. *Journal of Statistical Computation and Simulation*. 81(11):1641-1651.
- [3] Anderson, T.W., Darling, D.A. (1954). A test of goodness of fit. *Journal of American Statistical Association*. 49:765-769.
- [4] Baratpour, S., Habibirad, A. (2012). Testing goodness-of-fit for exponential distribution based on cumulative residual entropy. *Communications in Statistics-Theory and Methods*. 41(8):1387-1396.
- [5] Cramer H. (1928). On the composition of elementary errors, *Skandinavisk Aktuarietidskrift*. 11:13-74, 141-180.
- [6] Khalili, M., Habibirad, A., Yousefzadeh, F. (2017). Some Properties of Lin Wong Divergence on the Past Lifetime Data. *Communications in Statistics, Theory and Methods*. 1-13.
- [7] Lin, J., Wong, S.K.M. (1990). A New directed divergence measure and its characterization. *International Journal Of General System*. 17(1):73-81.
- [8] Meaker, M.K. (1987). Clinical Gerontology: A Guide to Assessment and Intervention. *American Journal of Occupational Therapy*. 41(4):269-269.
- [9] Parzen, E. (1979). Nonparametric Statistical Data Modeling. *Journal of the American Statistical Association*, 74(365):105-121.
- [10] Von Mises, R. (1931). Wahrscheinlichkeitsrechnung und Ihre Anwendung in der Statistik und Theoretischen Physik. F. Deuticke, Leipzig. 6(1).



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## Some results on comparison among coherent systems

Khanjari Sadegh, M.<sup>1</sup>      Mirjalili, A.<sup>2</sup>

<sup>1</sup> Department of Statistics, University of Birjand

<sup>2</sup> Department of Statistics, Maybod Branch, Islamic Azad University, Maybod

### Abstract

In this paper we first do an extensive review the literatures on the subject of comparison among systems and then extend some of them. Particularly we obtain some ordering results for comparing of coherent used systems with exchangeable components in terms of the hazard rate and likelihood ratio orders. Also we suggest a simple modified dynamic signature that can be used for ordering coherent used systems sometimes they can not be ordered by using of their usual dynamic signatures. It also makes some more weak conditions for ordering of used systems.

**Keywords:** Coherent systems, stochastic ordering, signature, dynamic signature, used systems.

## 1 Introduction

One of the most important problems in system reliability analysis is comparison among systems. Three most common tools to comparison among coherent systems are structure function, reliability function and the concept of system signature which are defined in the sequel. The comparison of two systems can be done in a fixed point of time(static comparison) or in during of the times(dynamic comparison). We assume that the compared systems are coherent(see Barlow and Proschan for details on coherent systems). In a fixed

---

<sup>1</sup>[mkhanjari@birjand.ac.ir](mailto:mkhanjari@birjand.ac.ir)

<sup>2</sup>[a.mirjalili@maybodiau.ac.ir](mailto:a.mirjalili@maybodiau.ac.ir)

point of time  $t$ , the lifetime and the reliability of components and the system are defined as follows:

$$X_i = \begin{cases} 1 & \text{if } i\text{th component is working} \\ 0 & \text{otherwise} \end{cases}$$

$$p_i = EX_i = P(X_i = 1), i = 1, 2, \dots, n,$$

$$\phi(\mathbf{X}) = \phi(X_1, \dots, X_n) = \begin{cases} 1 & \text{if the system is working} \\ 0 & \text{otherwise} \end{cases}$$

$h(\mathbf{p}) = h(p_1, \dots, p_n) = P(\phi(\mathbf{X}) = 1)$  where  $\phi(\mathbf{X})$  is called the structure function of the system. When  $t > 0$  is not fixed we denote by  $T_1, \dots, T_n$  and  $T = \phi(T_1, \dots, T_n)$  the component lifetimes and the system lifetime, respectively. In this case we have  $\bar{F}_i(t) = P(T_i > t) = 1 - F_i(t)$  and  $h_T(t) = P(T > t) = P(\phi(T_1, \dots, T_n) > t) = h(\bar{F}_1(t), \dots, \bar{F}_n(t))$  as the component and system reliabilities. When  $T_i$ 's are independent and identically distributed(i.i.d.), Samaniego (1985) obtained the following interesting and important result

$$P(T > t) = \sum_{i=1}^n s_i P(T_{i:n} > t) \quad (1)$$

where  $T_{i:n}$  is the  $i$ th ordered lifetime of components and  $s_i = P(T = T_{i:n})$ . The probability vector  $\mathbf{s} = (s_1, \dots, s_n)$  is called the signature of the system.

In iid case Samaniego et al.(2009) showed that

$$P(T = T_{k:n}|T > t, N(t) = r) = \frac{s_k}{\sum_{j=r+1}^n s_j} = w_k, \quad (2)$$

$k = r + 1, \dots, n$  and called the probability vector  $\mathbf{w}_{(n-r)} = (w_{r+1}, \dots, w_n)$  as the dynamic signature of the system at time  $t$ .  $N(t)$  is the number of failed components of the system up to time  $t$ . Under the above conditions the system is called a used system.

About the mentioned concepts, we have done an extensive review on literatures and also given our results which are coming in the next section.

## 2 Main results

We first consider the comparison of two systems based on their structure functions.

**Definition 2.1.** Let  $\phi_1(\mathbf{x})$  and  $\phi_2(\mathbf{x})$  be the structure functions of two systems of order  $n$ . The second system is said to be better than the first one if

$$\phi_1(\mathbf{x}) \leq \phi_2(\mathbf{x}) \quad \forall \mathbf{x} \in \{0, 1\}^n \quad (3)$$

where the inequality is strictly hold at least for one  $\mathbf{x}$ .

**Definition 2.2.** The second system is said to be more reliable than the first one if

$$h_1(\mathbf{p}) = E(\phi_1(\mathbf{X})) \leq h_2(\mathbf{p}) = E(\phi_2(\mathbf{X})) \quad \forall \mathbf{p} \in [0, 1]^n \quad (4)$$

where the inequality is strictly hold at least for one  $\mathbf{p}$ .

Kochar et al.(1999) claimed without proof that the relations (2.1) and (2.2) are equivalent when  $X_i$ 's are independent. Note that in this case

$$h(\mathbf{p}) = E(\phi(\mathbf{X})) = P(\phi(\mathbf{X}) = 1) = \sum_{\mathbf{x}: \phi(\mathbf{x})=1} \prod_{i \in \mathbf{1}_{\mathbf{x}}} p_i \prod_{i \in \mathbf{0}_{\mathbf{x}}} (1 - p_i)$$

where  $\mathbf{1}_{\mathbf{x}} = \{1 \leq i \leq n | x_i = 1\}$  and  $\mathbf{0}_{\mathbf{x}} = \{1 \leq i \leq n | x_i = 0\}$ . Here we give a proof for their claim. The inequality (2.1) simply implies the inequality (2.2). Now suppose (2.2) is hold. If  $\phi_1(\mathbf{x}) = \phi_2(\mathbf{x})$  for all  $\mathbf{x}$  obviously we have  $h_1(\mathbf{p}) = h_2(\mathbf{p})$  for all  $\mathbf{p}$  which is a contradiction. Now suppose there exists a vector  $\mathbf{x}_0$  such that  $\phi_1(\mathbf{x}_0) = 1 > \phi_2(\mathbf{x}_0) = 0$ . If we put  $\mathbf{p}_0 = \mathbf{x}_0$  then  $p_{0i} = 1$  for  $i \in \mathbf{1}_{\mathbf{x}_0}$  and  $1 - p_{0i} = 1$  for  $i \in \mathbf{0}_{\mathbf{x}_0}$ . Also  $P_{\mathbf{p}_0}(\mathbf{X} = \mathbf{x}) = 0$  for  $\mathbf{x} \neq \mathbf{x}_0$ . It is easy to show that  $h_1(\mathbf{p}_0) = 1$  and  $h_2(\mathbf{p}_0) = 0$  which is again a contradiction and this completes the proof.

**Remark 2.1.** Two above inequalities are not necessary equivalent if  $X_i$ 's are i.i.d. that is  $p_i = p$ ,  $i = 1, \dots, n$ . For example if  $\phi_1(\mathbf{x}) = \min\{x_1, \max(x_2, x_3, x_4)\}$  and  $\phi_2(\mathbf{x}) = \max\{\min(x_1, x_2), \min(x_2, x_3), \min(x_3, x_4)\}$  then  $\phi_1(1, 0, 0, 1) = 1 > \phi_2(1, 0, 0, 1) = 0$  and  $\phi_1(0, 0, 1, 1) = 0 < \phi_2(0, 0, 1, 1)$  whereas  $h_1(p) = 3p^2 - 3p^3 + p^4 \leq h_2(p) = 3p^2 - 2p^3$ , for all  $0 \leq p \leq 1$ . We now consider the comparison of coherent systems based on their signatures.

**Theorem 2.1.** Let  $T^1 = \phi_1(T_1, \dots, T_n)$  and  $T^2 = \phi_2(T_1, \dots, T_n)$  denote the lifetimes of two coherent systems where  $T_i$ 's are i.i.d. and continuous and suppose  $\mathbf{s}_1$  and  $\mathbf{s}_2$  are their signatures respectively.

- (a) If  $\mathbf{s}_1 \leq_{st} \mathbf{s}_2$  then  $T^1 \leq_{st} T^2$ .
- (b) If  $\mathbf{s}_1 \leq_{hr} \mathbf{s}_2$  then  $T^1 \leq_{hr} T^2$ .
- (c) If  $\mathbf{s}_1 \leq_{lr} \mathbf{s}_2$  then  $T^1 \leq_{lr} T^2$ .

**Proof.** See Kochar et al.(1999). We also refer to Shaked and Shanthikumar (2007) for details on stochastic orders.

**Example 2.1.** Consider two coherent systems mentioned in Remark 2.1. We have  $\mathbf{s}_1 = (1/4, 1/4, 1/2, 0) \leq_{st} \mathbf{s}_2 = (0, 1/2, 1/2, 0)$  and hence  $T^1 \leq_{st} T^2$ . Note that these two systems could not be ordered by using their structure functions.

The Equation (1.1) is the main tool in comparison among coherent systems by using their signatures. When  $T_i$ 's are i.i.d. and continuous random variables it was first used by Samaniego (1985) and then proved by Kochar et al.(1999). Navarro et al.(2005) claimed that this equation holds true if  $T_i$ 's have an absolutely continuous exchangeable joint distribution.

**Definition 2.3.** The random variables  $T_1, \dots, T_n$  are said to be exchangeable if  $P(T_1 > t_1, \dots, T_n > t_n) = P(T_{\pi(1)} > t_1, \dots, T_{\pi(n)} > t_n)$  for any permutation  $\pi = (\pi(1), \dots, \pi(n))$  of numbers  $\{1, \dots, n\}$ .

They extended Theorem 2.1 for the systems with exchangeable components as follow.

**Theorem 2.2.** (a) Let  $\mathbf{s}_1$  and  $\mathbf{s}_2$  be the signatures of two coherent systems with exchangeable components  $T_1 = \phi_1(X_1, \dots, X_n)$  and  $T_2 = \phi_2(Y_1, \dots, Y_n)$ . If  $(X_1, \dots, X_n) \leq_{st} (Y_1, \dots, Y_n)$  and  $\mathbf{s}_1 \leq_{st} \mathbf{s}_2$  then  $T_1 \leq_{st} T_2$ .

(b) Let  $T_1 = \phi_1(X_1, \dots, X_n)$  and  $T_2 = \phi_2(X_1, \dots, X_n)$ . If  $X_{i:n} \leq_{hr} X_{i+1:n}$ ,  $i = 1, \dots, n-1$  and  $\mathbf{s}_1 \leq_{hr} \mathbf{s}_2$  then  $T_1 \leq_{hr} T_2$ .

(c) If  $X_{i:n} \leq_{lr} X_{i+1:n}$ ,  $i = 1, \dots, n-1$  and  $\mathbf{s}_1 \leq_{lr} \mathbf{s}_2$  then  $T_1 \leq_{lr} T_2$ .

**Proof.** See Navarro et al.(2005)

Khanjari and Tavasolian (2010) extended parts (b) and (c) to the systems with different components. Navarro et al.(2008) showed that the Equation (1.1) holds true if  $T_i$ 's be continuous(or discrete) weakly exchangeable random variables and if  $P(T_i = T_j) = 0$  for all  $i \neq j$ . Marichal et al.(2011) showed that the Equation (1.1) holds true if and only if the following binary random variables are exchangeable

$$X_i(t) = \begin{cases} 1 & T_i > t \\ 0 & T_i \leq t. \end{cases}$$

Navarro et al.(2011) considered the signatures of coherent systems with heterogeneous components. For comparing of two systems, Navarro and Rubio (2010) obtained a necessary and sufficient condition as follow.

**Theorem 2.3.** Let  $T_1 = \phi_1(X_1, \dots, X_n)$  and  $T_2 = \phi_2(X_1, \dots, X_n)$  and  $\mathbf{s}_1$  and  $\mathbf{s}_2$  denote the signatures and let  $\mathbf{F}$  be the joint distribution of  $(X_1, \dots, X_n)$ .

(a)  $\mathbf{s}_1 \leq_{st} \mathbf{s}_2$  iff  $T_1 \leq_{st} T_2$  for any exchangeable joint distribution  $\mathbf{F}$ .

(b)  $\mathbf{s}_1 \leq_{hr} \mathbf{s}_2$  iff  $T_1 \leq_{hr} T_2$  for any exchangeable joint distribution  $\mathbf{F}$  in which  $X_{i:n} \leq_{hr} X_{i+1:n}$ ,  $i = 1, \dots, n-1$ .

(c)  $\mathbf{s}_1 \leq_{lr} \mathbf{s}_2$  iff  $T_1 \leq_{lr} T_2$  for any exchangeable joint distribution  $\mathbf{F}$  in which  $X_{i:n} \leq_{lr} X_{i+1:n}$ ,  $i = 1, \dots, n-1$ .

**Proof.** See Navarro and Rubio (2010).

We now show in the next theorem that if two systems are ordered by using their structure functions their signatures are also ordered.

**Theorem 2.4.** If in Theorem 2.3,  $\mathbf{F}$  is an arbitrary exchangeable joint distribution and  $\phi_1(\mathbf{x}) \leq \phi_2(\mathbf{x})$  for all  $\mathbf{x} \in \{0, 1\}^n$  then  $\mathbf{s}_1 \leq_{st} \mathbf{s}_2$ .

**Proof.** We have  $h_1(\mathbf{p}) = E(\phi_1(\mathbf{X})) \leq E(\phi_2(\mathbf{X})) = h_2(\mathbf{p})$  for all  $\mathbf{p} \in [0, 1]^n$ . On the other hand we note that  $h_i(\mathbf{p}) = P(T_i > t)$ ,  $i = 1, 2$ . Therefore  $T_1 \leq_{st} T_2$ . Now from part (a) of Theorem 2.3, the proof follows.

We now consider the comparison of coherent systems which are of different sizes.

**Definition 2.4.** Let  $\mathbf{p} = (p_1, \dots, p_k)$  be a probability vector and suppose  $T_j = \phi_j(X_1, \dots, X_n)$ ,  $j = 1, \dots, k$  are lifetimes of  $k$  coherent systems. If  $P(T = T_j) = p_j$  then  $T$  is said to be the lifetime of a mixed- $\mathbf{p}$  system. Note that a mixed system is not in general a coherent system and the class of coherent systems is a subset of the mixed systems. If  $\mathbf{s}_j$  is the signature of the  $j$ th coherent system then  $\mathbf{s} = \sum_{j=1}^k p_j \mathbf{s}_j$  is the signature of the mixed- $\mathbf{p}$  system.

For a system of order  $n$ , Samaniego (2006) introduced an equivalent system of order  $n+1$  as follow.

**Theorem 2.5.** Let  $T$  be the lifetime of a coherent or mixed system of order  $n$  with i.i.d. components having common distribution  $F$  and with signature  $\mathbf{s} = (s_1, \dots, s_n)$ . Then  $T =^{st} T^*$  where  $T^*$  is the lifetime of a mixed system of order  $n+1$  with i.i.d. components which have the same distribution  $F$  and its signature is  $\mathbf{s}^* = (s_1^*, \dots, s_{n+1}^*)$  where

$$s_i^* = \frac{(i-1)s_{i-1} + (n-i+1)s_i}{n+1}, \quad i = 1, \dots, n+1.$$

**Proof.** See Samaniego (2006).

For a system of order  $k$  we can define an equivalent system of order  $n(> k)$ . We note that the Theorem 2.5 is also hold true for exchangeable components. We also note that the

equivalent systems are in general mixed systems. We may add that all comparison results among the coherent systems based on signatures are also hold for the mixed systems.

Navarro et al.(2008) extended the Theorem 2.2 for comparing of coherent systems with exchangeable components and with different sizes as follow.

**Theorem 2.6.** Let  $T_1 = \phi_1(Y_1, \dots, Y_{n_1})$  and  $T_2 = \phi_2(Z_1, \dots, Z_{n_2})$  be the lifetimes of two systems where  $Y_i$ 's and  $Z_i$ 's are subsets of the exchangeable random variables  $\{X_1, \dots, X_n\}$  and suppose  $\mathbf{s}_1(n)$  and  $\mathbf{s}_2(n)$  are their equivalent signatures of order  $n$ , respectively.

- (a) If  $\mathbf{s}_1(n) \leq_{st} \mathbf{s}_2(n)$  then  $T_1 \leq_{st} T_2$ .
- (b) If  $\mathbf{s}_1(n) \leq_{hr} \mathbf{s}_2(n)$  and  $X_{i:n} \leq_{hr} X_{i+1:n}$ ,  $i = 1, \dots, n - 1$  then  $T_1 \leq_{hr} T_2$ .
- (c) If  $\mathbf{s}_1(n) \leq_{lr} \mathbf{s}_2(n)$  and  $X_{i:n} \leq_{lr} X_{i+1:n}$ ,  $i = 1, \dots, n - 1$  then  $T_1 \leq_{lr} T_2$ .

**Proof.** See Navarro et al.(2008).

Finally based on dynamic signature defined in previous section we now consider the comparison of coherent used systems.

**Theorem 2.7.** Let  $\mathbf{s}_1$  and  $\mathbf{s}_2$  be the signatures of two coherent systems with i.i.d. components having common continuous distribution  $F$  and let  $T_1 = \phi_1(X_1, \dots, X_n)$  and  $T_2 = \phi_2(X_1, \dots, X_n)$ . Suppose  $T_1 > t$  and  $T_2 > t$  and  $N_1(t) = r_1$  and  $N_2(t) = r_2$ . Let  $\mathbf{w}_1 = (w_{1,1}, \dots, w_{1,n-r_1})$  and  $\mathbf{w}_2 = (w_{2,1}, \dots, w_{2,n-r_2})$  be their dynamic signatures. Also assume that  $\mathbf{w}_1^{(n)}$  and  $\mathbf{w}_2^{(n)}$  are their equivalent dynamic signatures of order  $n$ .

- (a) If  $\mathbf{w}_1^{(n)} \leq_{st} \mathbf{w}_2^{(n)}$  then  $(T_1 - t|T_1 > t, N_1(t) = r_1) \leq_{st} (T_2 - t|T_2 > t, N_2(t) = r_2)$ .
- (b) If  $\mathbf{w}_1^{(n)} \leq_{hr} \mathbf{w}_2^{(n)}$  then  $(T_1 - t|T_1 > t, N_1(t) = r_1) \leq_{hr} (T_2 - t|T_2 > t, N_2(t) = r_2)$ .
- (c) If  $F$  is absolutely continuous and  $\mathbf{w}_1^{(n)} \leq_{lr} \mathbf{w}_2^{(n)}$  then  $(T_1 - t|T_1 > t, N_1(t) = r_1) \leq_{lr} (T_2 - t|T_2 > t, N_2(t) = r_2)$ .

**Proof.** See Samaniego et al.(2009).

They showed that

$$P(T - t > x|T > t, N(t) = r) = \sum_{j=1}^{n-r} w_{r+j} \bar{G}_{j:n-r|t}(x) = \sum_{j=1}^n w_j^{(n)} \bar{G}_{j:n|t}(x)$$

where  $\bar{G}_{j:n-r|t}(x)$  is the reliability function of the  $j$ th order statistic from a random sample of size  $n - r$  from  $\bar{G}(x|t) = \frac{\bar{F}(x+t)}{\bar{F}(t)}$  for  $x > 0$ .

For a used system with exchangeable components let  $T = \phi(T_1, \dots, T_n)$ . Khanjari (2016) obtained the following equation

$$\begin{aligned} P(T - t > x|T > t, N(t) = r) \\ = \sum_{k=r+1}^n w_k P(T_{k:n} - t > x|N(t) = r) = \sum_{k=1}^n w_k^{(n)} P(T_{k:n} - t > x|N(t) = r). \end{aligned}$$

Now based on above equation we extend the Theorem 2.7 not only to the systems with exchangeable components but also to the systems with different components.

**Theorem 2.8.** Let  $T_1 = \phi_1(X_1, \dots, X_n)$  and  $T_2 = \phi_2(Y_1, \dots, Y_n)$  be the lifetimes of two coherent systems with exchangeable components without tie. Suppose  $T_1 > t$  and  $T_2 > t$  and  $N_1(t) = r_1$  and  $N_2(t) = r_2$ .

- (a) If  $\mathbf{w}_1^{(n)} \leq_{st} \mathbf{w}_2^{(n)}$  and  $(X_{i:n} - t|N_1(t) = r_1) \leq_{st} (Y_{i:n} - t|N_2(t) = r_2)$  for  $i = 1, \dots, n$  then  $(T_1 - t|T_1 > t, N_1(t) = r_1) \leq_{st} (T_2 - t|T_2 > t, N_2(t) = r_2)$ .

(b) If  $(X_{i:n} - t|N_1(t) = r_1) \leq_{hr} (Y_{j:n} - t|N_2(t) = r_2)$ , for all  $i, j = 1, \dots, n$  then  $(T_1 - t|T_1 > t, N_1(t) = r_1) \leq_{hr} (T_2 - t|T_2 > t, N_2(t) = r_2)$ .

(c) If  $X_i$ 's and  $Y_j$ 's be absolutely continuous and  $(X_{i:n} - t|N_1(t) = r_1) \leq_{lr} (Y_{j:n} - t|N_2(t) = r_2)$ , for all  $i, j = 1, \dots, n$  then  $(T_1 - t|T_1 > t, N_1(t) = r_1) \leq_{lr} (T_2 - t|T_2 > t, N_2(t) = r_2)$ .

**Proof.** (a) It is known that if  $X \leq_{st} Y$  then  $Eg(X) \leq Eg(Y)$  for any increasing function  $g$ . Now we have

$P(T_1 - t > x|T_1 > t, N_1(t) = r_1) = \sum_{k=1}^n w_{1,k}^{(n)} P(X_{k:n} - t > x|N_1(t) = r_1) \leq \sum_{k=1}^n w_{2,k}^{(n)} P(X_{k:n} - t > x|N_1(t) = r_1) \leq \sum_{k=1}^n w_{2,k}^{(n)} P(Y_{k:n} - t > x|N_2(t) = r_2) = P(T_2 - t > x|T_2 > t, N_2(t) = r_2)$ . In view of Theorem 8 in Navarro et al.(2005) or Lemma 3 in Khanjari and Tavasolian (2010) the proof of part (b) follows. Also the proof of part (c) follows from Khanjari and Tavasolian (2010), page 8.

We note that if the components of the systems are common then the conditions of the Theorem 2.8 become more weak.

Here we propose a simple modified dynamic signature  $\mathbf{w}^*$  which is of order  $n$  and has some useful applications. Its coordinates are given below:

$$w_i^* = \begin{cases} 0 & i \leq r \\ w_i & i > r \end{cases} \quad (5)$$

Obviuosly we have

$$P(T - t > x|T > t, N(t) = r) = \sum_{k=1}^n w_k^* P(T_{k:n} - t > x|N(t) = r) \quad (6)$$

**Theorem 2.9.** Under the conditions of Theorem 2.8 we have

(a) If  $\mathbf{w}_1^* \leq_{st} \mathbf{w}_2^*$  and  $(X_{i:n} - t|N_1(t) = r_1) \leq_{st} (Y_{j:n} - t|N_2(t) = r_2)$  for  $i = r_2 + 1, \dots, n$  then  $(T_1 - t|T_1 > t, N_1(t) = r_1) \leq_{st} (T_2 - t|T_2 > t, N_2(t) = r_2)$ .

(b) If  $(X_{i:n} - t|N_1(t) = r_1) \leq_{hr} (Y_{j:n} - t|N_2(t) = r_2)$ , for  $i = r_1 + 1, \dots, n$  and  $j = r_2 + 1, \dots, n$  then  $(T_1 - t|T_1 > t, N_1(t) = r_1) \leq_{hr} (T_2 - t|T_2 > t, N_2(t) = r_2)$ .

(c) If  $X_i$ 's and  $Y_j$ 's be absolutely continuous and  $(X_{i:n} - t|N_1(t) = r_1) \leq_{lr} (Y_{j:n} - t|N_2(t) = r_2)$ , for  $i = r_1 + 1, \dots, n$  and  $j = r_2 + 1, \dots, n$  then  $(T_1 - t|T_1 > t, N_1(t) = r_1) \leq_{lr} (T_2 - t|T_2 > t, N_2(t) = r_2)$ .

**Proof.** In the proof of Theorem 2.8 if we use the Equation (2.4) the proof follows.

We see that for comparing of used systems the conditions of Theorem 2.9 are more weak than those of Theorem 2.8. Another useful property of  $\mathbf{w}^*$  is shown in the following example.

**Example 2.2.** Let  $\mathbf{s}_1 = (0, 0, 0, 4/5, 1/5)$  and  $\mathbf{s}_2 = (0, 0, 3/5, 1/5, 1/5)$  be the signatures of two coherent systems consisting of common i.i.d. components and suppose  $N_1(t) = 2$  and  $N_2(t) = 3$ . From Equation (1.2) we have  $\mathbf{w}_1 = (0, 4/5, 1/5)$  and  $\mathbf{w}_2 = (1/2, 1/2)$ . Also from Theorem 2.5 it is easy to show that the equivalent vectors of  $\mathbf{w}_1$  and  $\mathbf{w}_2$  are  $\mathbf{w}_1^{(5)} = (0, 24/100, 34/100, 30/100, 12/100)$  and  $\mathbf{w}_2^{(5)} = (1/5, 1/5, 1/5, 1/5, 1/5)$ , respectively. We note that in view of the usual stochastic order,  $\mathbf{w}_1^{(5)}$  and  $\mathbf{w}_2^{(5)}$  are not ordered. Therefore by using part (a) in Theorem 2.8, the mentioned used systems can not be ordered. But it is easy to see that  $\mathbf{w}_1^* = (0, 0, 0, 4/5, 1/5) \leq_{st} \mathbf{w}_2^* = (0, 0, 0, 1/2, 1/2)$ . Therefore by using part (a) of Theorem 2.9 in i.i.d. case we have  $(T_1 - t|T_1 > t, N_1(t) = 2) \leq_{st} (T_2 - t|T_2 > t, N_2(t) = 3)$ .

## References

- [1] Barlow, R.E. and Proschan, F. (1975), *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart and Winston.
- [2] Khanjari Sadegh, M.(2016), Dynamic signature of a coherent system with exchangeable components. *13th Iranian statistical conference*.
- [3] Khanjari Sadegh M. and Tavasolian T. (2010), A note on the comparisons among coherent systems, *Journal of Statistical Research of Iran*. 7, 1, 1-10.
- [4] Kochhar S., Mukherjee H. and Samaniego F.J. (1999), The signature of a coherent system and its application to comparisons among systems, *Naval Research Logistic*. 46, 507-523.
- [5] Marichal, J.L., Mathonet, P. and Waldhauser, T.(2011), On signature-based expressions of system reliability. *Journal of Multivariate Analysis*, **102**, 1410-1416.
- [6] Navarro, J., Balakrishnan, N. and Samaniego, F.J.(2011), Signature-based representations for the reliability of systems with heterogeneous components. *Journal of Applied Probability*, **48**, 856-867.
- [7] Navarro, J. and Rubio, R. (2011), A note on necessary and sufficient conditions for ordering properties of coherent systems with exchangeable components. *Naval Research Logistic*, **58(5)**, 478-489.
- [8] Samaniego, F.J. (1985), on the closure of the IFR class under formation of coherent systems. *IEEE Trans. on Rel.*, **34**, 69-72.



## On the point reconstruction of order statistics from geometric distribution

Khatib Astaneh, B.<sup>1</sup>

Department of Mathematics and Statistics, University of Neyshabur

### Abstract

Sometimes the life testing experiment may be studied in a discrete set up. Here, the problem of reconstructing missing order statistics from Geometric distribution is investigated; some classic and Bayes reconstructors are presented.

**Keywords:** Order statistics, Missing order statistics, Conditional distribution, Geometric distribution.

## 1 Introduction

There are some situations in life-testing and reliability experiments in which a part of the sample (or subjects) are lost or removed from the experiment. For example, in a life-testing experiment, suppose  $n$  items are placed on the test simultaneously. Sometimes some of observations may be observed at the end of the experiment. In such a situation, the experimenter may not obtain complete information on failure times for all experimental units and might be interested in getting inferences about the history of the system, e.g., when the individual components have failed. So, one question that arises here is “How can one reconstruct the missing observations?”.

Recently, the problem of reconstructing past data points in continuous set-up has been studied by some authors such as Klimczak and Rychlik (2005), Balakrishnan et al. (2009), Razmkhah et al. (2010), Asgharzadeh et al. (2011), Khatib et al. (2012, 2013). In these cases the bayesian approach is investigated by several authors such as Fernndez (2004),

<sup>1</sup>khatib\_b@neyshabur.ac.ir

Khatib et al. (2011), Ahmad et al. (2011) and etc.

In some situations, the life testing experiment must be investigated in a discrete set up. For example, suppose that the life times of the units in an experiment depend on the number of times the units are switched on and off, the number of pages a printer prints or the number of rotations of a machine and etc. Let  $w$  be the number of aforementioned shocks the units receive until they fail, so,  $w$  is considered as the associated failure time. Censored samples in discrete set up have been studied by some authors. For example, Rezaei and Arghami (2002) investigated Type-I and Type-II right censoring in a discrete life model. Davarzani and Parsian (2011) considered discrete middle censored samples and used both classical and Bayesian approaches to do inference about the parameter of interest. Balakrishnan et al. (2011) studied characterizations of geometric distribution based on some properties of progressively Type-II right censored order statistics.

If a component in a system is capable of just two modes of performance, functioning and non-functioning, the mode of performance of the component may be represented by Bernoulli random variable  $X$  with the values 0 and 1 depending on whether the component is functioning or non-functioning with the probability distribution

$$P(X = 1) = 1 - P(X = 0) = \theta, \quad 0 < \theta < 1,$$

The geometric distribution with probability mass function,

$$f(x, \theta) = \theta(1 - \theta)^{x-1}, \quad x = 1, 2, \dots \quad 0 < \theta < 1, \quad (1)$$

is a discrete analogue of exponential distribution, is used as discrete lifetime failure model, which has been studied by Shah and Patel (2007), Patel and Patel (2005), Lui (1997), Patel and Gajjar (1990) and Yaqub and Khan (1981).

The rest of this paper is as follows: In Section 2, some preliminaries are presented. By considering geometric distribution, a reconstruction for the past order statistics is presented in Section 3; in this section, the expected value of the mean of missing order statistics given the observed data will be derived. Section 4 focuses on the Bayesian reconstruction of the lost order statistics. Finally, some conclusions are stated in Section 5.

## 2 Some preliminaries

Here, we present some properties of order statistics from discrete distribution which will be used to obtain the results in the next sections. The most important concept in order statistics of discrete distribution is tie-run. This concept is defined below, which was utilized by Rüschendorf (1985) implicitly.

**Definition.** A subchain  $t_{i_1} \leq t_{i_2} \leq \dots \leq t_{i_k}$  of real numbers is said to have  $r$  tie-runs ( $1 \leq r \leq k$ ) with length  $z_j$  ( $1 \leq j \leq r$ ) for the  $j$ th one, if

$$t_{i_1} = \dots = t_{i_{z_1}} < t_{i_{z_1+1}} = \dots = t_{i_{z_1+z_2}} < \dots < t_{i_{n-z_r+1}} = \dots = t_{i_k},$$

with  $\sum_{j=1}^r z_j = k$ .

- The joint pmf of the vector of order statistics in a random sample of size  $n$ ,  $\mathbf{X} = (X_{(1)}, \dots, X_{(n)})$  is given by

$$p_{\mathbf{X}}(\mathbf{x}) = n! \left( \prod_{j=1}^r z_j! \right)^{-1} \prod_{i=1}^n p(x_i), \quad (2)$$

where  $x_1 \leq x_2 \leq \dots \leq x_n$  has  $r$  tie-runs with length  $z_j$  for the  $j$ th one,  $j = 1, \dots, r$  (see, Arnold et al., 2008, pp. 42).

- Moreover, similar on lines to Gan and Bain(1995), it can be shown that the joint pmf of  $\mathbf{X}_{(m-n)} = (X_{(m)}, X_{(m+1)}, \dots, X_{(n)})$  is given by

$$p_{\mathbf{X}_{(m-n)}}(\mathbf{x}) = \frac{n! \prod_{i=m}^n p(x_i)}{\prod_{j=2}^r z_j!} \sum_{\nu=0}^{m-1} \frac{(p(x_m))^{\nu} (F(x_{m-}))^{m-1-\nu}}{(z_1 + \nu)! (m - 1 - \nu)!}, \quad (3)$$

where  $x_m \leq x_2 \leq \dots \leq x_n$  has  $r$  tie-runs with length  $z_j$  for the  $j$ th one,  $j = 1, \dots, r$ .

- Furthermore, it can be shown that for  $1 \leq l < m \leq n$ , the pmf of  $\mathbf{X}_{(l,m-n)} = (X_{(l)}, X_{(m)}, X_{(m+1)}, \dots, X_{(n)})$  is presented as follows

$$\begin{aligned} p_{\mathbf{X}_{(l,m-n)}}(x_l, \mathbf{x}) &= \frac{n! \prod_{i=m}^n p(x_i)}{\prod_{j=2}^r z_j!} \sum_{\nu=0}^{l-1} \frac{(p(x_m))^{m-1-\nu} (F(x_{m-}))^{\nu}}{(z_1 + m - 1 - \nu)! \nu!} \quad (\text{if } x_l = x_m) \\ &= \frac{n! \prod_{i=m}^n p(x_i)}{\prod_{j=2}^r z_j!} \sum_{u=0}^{m-l-1} \sum_{h=0}^{m-l-1-u} \sum_{\nu=0}^{l-1} \frac{(p(x_m))^u (F(x_{m-}) - F(x_l))^h}{(u + z_1)! h!} \\ &\quad \times \frac{(p(x_l))^{m-1-u-h-\nu} (F(x_{l-}))^{\nu}}{(m - 1 - u - h - \nu)! \nu!} \quad (\text{if } x_l < x_m) \\ &= 0 \quad (\text{if } x_l > x_m), \end{aligned}$$

where  $\mathbf{x} = (x_m, \dots, x_n)$  is the observed value of  $\mathbf{X}$ , and  $x_m \leq x_{m+1} \leq \dots \leq x_n$  has  $r$  tie-runs with length  $z_j$  for the  $j$ th one,  $j = 1, 2, \dots, r$ .

- The conditional distribution of  $X_{(l)}$  given  $\mathbf{X}_{(m-n)} = \mathbf{x}$  is as follow

Let  $1 \leq l < m \leq n$  and  $p(x_j) > 0$  for  $j = m, \dots, n$ ,

$$P[X_{(l)} = x_l | \mathbf{X}_{(m-n)} = \mathbf{x}] = \begin{cases} A(x_m, z_1) & (\text{if } x_l = x_m) \\ \frac{B(x_m, z_1)}{C(x_m, z_1)} & (\text{if } x_l < x_m) \\ 0 & (\text{if } x_l > x_m), \end{cases} \quad (4)$$

where

$$A(x_m, z_1) = \frac{\sum_{\nu=0}^{l-1} \frac{(m-1+z_1)!}{(z_1+m-1-\nu)!\nu!} (p(x_m))^{z_1+m-1-\nu} (F(x_{m-}))^{\nu}}{\sum_{\nu=0}^{m-1} \frac{(m-1+z_1)!}{(z_1+m-1-\nu)!\nu!} (p(x_m))^{z_1+m-1-\nu} (F(x_{m-}))^{\nu}},$$

and

$$B(x_m, z_1) = \sum_{u=0}^{m-1-l} \sum_{h=0}^{m-1-l-u} \sum_{\nu=0}^{l-1} \frac{(m-1+z_1)!}{(u+z_1)!h!(m-1-u-h-\nu)!\nu!} \\ \times (p(x_m))^{u+z_1} (F(x_m-) - F(x_l))^h (p(x_l))^{m-1-u-h-\nu} (F(x_l-))^\nu,$$

and

$$C(x_m, z_1) = \sum_{\nu=0}^{m-1} \frac{(m-1+z_1)!}{(z_1+\nu)!(m-1-\nu)!} (p(x_m))^{z_1+\nu} (F(x_m-))^{m-1-\nu}. \quad (5)$$

By attention to (4) it can be seen that the conditional distribution of  $X_{(l)}$  given  $\mathbf{X}_{(m-n)} = \mathbf{x}$  depends only on the first tie-run and  $X_{(m)} = x_m$ .

### 3 Point reconstruction

In order to reconstruct the  $l$ th ( $l < m$ ) order statistic based on the data set  $\mathbf{X}_{(n-m)}$ , such that  $x_m \leq x_{m+1} \leq \dots \leq x_n$  has  $r$  tie-runs with length  $z_j$  for the  $j$ th one,  $j = 1, \dots, r$ , we use the conditional distribution of  $X_{(l)}$  given  $\mathbf{X}_{(m-n)} = \mathbf{x}$ . A logical reconstructor of  $X_{(l)}$  having observed  $\mathbf{X}_{(n-m)}$ , may be  $E[X_{(l)}|\mathbf{X}_{(n-m)}]$ . With this in mind, using (4), we consider

$$\sum_{y < x_m} y P[X_{(l)} = y | \mathbf{X}_{(n-m)} = \mathbf{x}]$$

as an unbiased conditional reconstrucor of  $X_{(l)}$ .

Thus, using (4), the unbiased conditional reconstructor of  $X_{(l)}$  having observed  $\mathbf{X}_{(m)}$  for geometric distribution is as follow

$$\hat{X}_{(l)} = \sum_{y=1}^{x_m-1} y P[X_{(l)} = y | \mathbf{X}_{(n-m)} = \mathbf{x}] \\ = \frac{\psi(x_m, z_1)}{C(x_m, z_1)},$$

where

$$\psi(x_m, z_1) = \sum_{u=0}^{m-1-l} \sum_{h=0}^{m-1-l-u} \sum_{\nu=0}^{l-1} \sum_{i=0}^h \sum_{j=0}^{\nu} \frac{(m-1+z_1)!}{(u+z_1)!h!(m-1-u-h-\nu)!\nu!} \\ \binom{h}{i} (-1)^i \binom{\nu}{j} (-1)^j \times \frac{\theta^{m-1+z_1-h-\nu} (1-\theta)^{(x_m-1)(z_1+u+i)+h-i}}{(1-(1-\theta)^{m-1-u-\nu+j-i})^2} \\ \times \left( 1 - x_m (1-\theta)^{(m-1-u-\nu+j-i)x_m} (1 - (1-\theta)^{m-1-u-\nu+j-i}) \right. \\ \left. - (1-\theta)^{(m-1-u-\nu+j-i)x_m} \right),$$

and  $C(x_m, z_1)$  is as defined in (5).

## 4 Bayesian approach

In Bayesian analysis, we use not only the sample information but also some information about the parameter. Usually, in the Bayesian framework, given the status of a random variable  $\mathbf{X}$ , a conditional probability is attached to this variable, say  $f(x|\theta)$ , and a prior density of the parameter, say  $\pi(\theta)$ , is specified based on previous knowledge. So, from a Bayes viewpoint, reconstruction problem mainly lies on the posterior distribution of the parameter  $\theta$  given the sample information  $\mathbf{X}$ ,  $\pi(\theta|x)$ , which can be obtained from the combination of the prior density and the sample information via Bayes theorem.

In this section we obtain bayes reconstructors for missing order statistics while the underlying distribution is geometric. Let us denote the prior and posterior distributions by  $\pi(\theta)$  and  $\pi(\theta|\mathbf{x})$ , respectively. Then, the Bayes reconstructive mass function of  $X_{(l)}$  given  $\mathbf{X} = \mathbf{x}$  is

$$h(x_l|\mathbf{x}) = \int f(x_l|\mathbf{x}, \theta) \pi(\theta|\mathbf{x}) d\theta. \quad (6)$$

In what follows, we suppose that  $X$  has a geometric distribution as given in (1) and that the parameter  $\theta$  has beta distribution with the pdf

$$\pi(\theta) = \frac{1}{\beta(a, b)} \theta^{a-1} (1-\theta)^{b-1}, \quad 0 < \theta < 1. \quad (7)$$

By Bayes theorem and using (3) and (7) for the geometric distribution, the posterior distribution of  $\theta$  given  $\mathbf{X}_{(m-n)} = (X_{(m)}, X_{(m+1)}, \dots, X_{(n)})$ , for  $1 \leq l < m \leq n$ , is given by

$$\pi(\theta|\mathbf{x}) = \frac{\varphi_1(x_m, z_1, a, b)}{\varphi_2(x_m, z_1, a, b)}, \quad (8)$$

where

$$\varphi_1(x_m, z_1, a, b) = \sum_{\nu=0}^{m-1} \sum_{i=0}^{m-1-\nu} \binom{m-1-\nu}{i} (-1)^i \frac{\theta^{z_1+\nu+a-1} (1-\theta)^{(x_m-1)(z_1+\nu+i)+b-1}}{(z_1+\nu)!(n-m-\nu)!},$$

and

$$\begin{aligned} \varphi_2(x_m, z_1, a, b) &= \sum_{\nu=0}^{m-1} \sum_{i=0}^{m-1-\nu} \frac{\binom{m-1-\nu}{i} (-1)^i}{(z_1+\nu)!(n-m-\nu)!} \\ &\quad \times \frac{\Gamma(z_1+\nu+a) \Gamma((x_m-1)(z_1+\nu+i)+b)}{\Gamma(z_1+\nu+(x_m-1)(z_1+\nu+i)+a+b)}, \end{aligned} \quad (9)$$

Substituting (4) and (8) in (6) and upon integrating with respect to  $\theta$ , we obtain the Bayes reconstructive mass function of  $X_{(l)}$  (for,  $1 \leq l < m \leq n$ ) given  $\mathbf{X}_{(n-m)}$ , as follows

$$h(x_l|\mathbf{x}) = \frac{\varphi_3(x_m, x_l, z_1, a, b)}{\varphi_2(x_m, z_1, a, b)},$$

where

$$\varphi_3(x_m, x_l, z_1, a, b) = \sum_{u=0}^{m-1-l} \sum_{h=0}^{m-1-l-u} \sum_{\nu=0}^{l-1} \sum_{i=0}^h \sum_{j=0}^{\nu} \frac{\binom{h}{i} (-1)^i \binom{\nu}{j} (-1)^j}{(u+z_1)! h! (m-1-u-h-\nu)! \nu!} \frac{\Gamma(z_1 + m - 1 - h - \nu + a) \Gamma((x_m - 1)(z_1 + u + i) + (x_l - 1)(m - 1 - u - \nu - i) + h - i + b)}{\Gamma(z_1 + m - 1 - \nu + (x_m - 1)(z_1 + u + i) + (x_l - 1)(m - 1 - u - \nu - i) - i + a + b)},$$

and  $\varphi_2(x_m, z_1, a, b)$  is defined in (9).

The Bayes point reconstructor of  $X_{(l)}$  under the squared error loss (SEL) function is the expectation of the Bayes reconstructive mass function of  $X_{(l)}$  given  $\mathbf{X}$ , i.e.  $E(X_{(l)}|\mathbf{X})$ .

## 5 Concluding remark

In this paper we considered order statistics from geometric distribution and derived the probabilities mass functions of some order statistics and related functions. Unbiased reconstructors as well as Bayes reconstructors were presented for missing order statistics. Some numerical computation may be used to clarity the performance of the proposed procedures in this paper.

## References

- [1] Ahmad, K. E. Jaheen, Z. F. Mohammed, Heba S.(2012), Bayesian prediction based on type-I censored data from a mixture of Burr type XII distribution and its reciprocal, *Statistics*, **46**, 787-797.
- [2] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (2008). A First Course in Order Statistics. *Philadelphia: SIAM*.
- [3] Berger, JO. (1985), Statistical decision theory and Bayesian analysis. *Springer, Berlin Heidelberg New York*
- [4] J.B. Shah and M.N. Patel,(2007), Bayes estimation of shift point in geometric sequence, *Comm. Statist. Theory Methods*, **36** pp. 1139-1151.
- [5] Asgharzadeh, A., Ahmadi, J., Mirzazadeh Ganji, Z. and Valiollahi, R. (2012). Reconstruction of the past failure times for the proportional reversed hazard rate model, *Journal of Statistical Computation and Simulation*, **82**, 475-489.
- [6] Balakrishnan, N., Doostparast, M., Ahmadi, J. (2009). Reconstruction of past records. *Metrika*, **70**, 89–109.
- [7] Balakrishnan, N., Cramer, E. and Dembińska, A. (2011), Characterizations of geometric distribution through progressively Type-II right-censored order statistics, *Statistics*, **45**, 559–573.

- [8] Davarzani, N. and Parsian, A. (2011), Statistical Inference for Discrete Middle-Censored Data, *J. Statist. Plann. Inference*, **141**, 1455–1462.
- [9] Gan, G. and Bain, L. J. (1995), Distribution of order statistics for discrete parents with applications to censored sampling, *Journal of Statistical Planning and Inference*, **44**, 37–46.
- [10] Khatib, B., Ahmadi, J. and Razmkhah, M. (2013). Bayesian reconstruction of the missing failure times in exponential distribution. *Journal of Statistical Computation and Simulation*, **83**, 501–517.
- [11] Khatib, B., Ahmadi, J. and Razmkhah, M. (2012). Reconstruction of the past lower record values in a proportional reversed hazard rate model. *to appear in Statistics*.
- [12] Khatib, B. and Ahmadi, J. (2014), Best linear unbiased and invariant reconstructor for the past records, *Bullet. Malay. Math. Sci. Soc.*, **37**, 1017-1028.
- [13] Razmkhah, M., Khatib, B. and Ahmadi, J. (2010). Reconstruction of order statistics in exponential distribution, *Journal of the Iranian Statistical Society*, **9**, 21-40.
- [14] Razmkhah, M., Khatib, B. and Ahmadi, J. (2011). On performance of reconstructed middle order statistics in exponential distribution, *Journal of Statistical research of Iran*, **11**, 49–62.
- [15] Rezaei, A. H. and Arghami, N. R. (2002) Right censoring in a discrete life model, *Metrika*, **55**, 151–160.
- [16] Rüschenendorf, L. (1985), Two remarks on order statistics. *J. Statist. Plann. Inference*, **11**, 71–74.
- [17] Patel, N.W. and Patel, M.N.(2005), A problem in discrete life testing with progressive changing failure rate, *Indian Statist* , **43**, 51-62.
- [18] Lui, K.J. (1997), Estimation of the geometric survival distribution, *IEEE Trans. Reliab.*, **46**, 134-144.
- [19] Patel, M.N. and Gajjar, A.V. (1990), Geometric failure law in life testing, *Aligarh J. Statist.* **10**, 1-8.
- [20] Yaqub, M. and Khan, A.H. (1981), Geometric failure law in life testing, *Pure Appl. Math. Sci.*, **XIV**, 69-76.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



# Bayesian inference for $R = P(X < Y)$ under adaptive type-II hybrid progressive censored samples in generalized inverted exponential distribution

Kohansal, A.<sup>1</sup>

Department of Statistics, Imam Khomeini International University, Qazvin

## Abstract

In this paper, the Bayesian inference of  $R = P(X < Y)$  for generalized inverted exponential distribution under the adaptive Type-II hybrid progressive censored samples is considered. We solve the problem in three cases. In first case, assuming that  $X$  and  $Y$  have the unknown common scale parameter and different shape parameters, the Bayes estimate of  $R$  is derived by two approximation method: Lindley's approximation and MCMC method. In second case, assuming that  $X$  and  $Y$  have the known common scale parameter and unknown different shape parameters, the exact Bayes estimate of  $R$  is derived. In third case, assuming that all parameters are different and unknown, the Bayesian inference of  $R$  is derived by MCMC method. We use one Monte Carlo simulation study to compare the performance of different methods.

**Keywords:** Adaptive Type-II hybrid progressive censored sample , Stress-strength model, Generalized inverted exponential distribution, Bayesian inference.

## 1 Introduction

Statistical inference about the stress-strength parameter,  $R = P(X < Y)$ , is one of the most important problem in reliability theory and statistics and has been done from the frequentist and Bayesian viewpoints. In spite of the fact that, many papers have studied

---

<sup>1</sup>kohansal@sci.ihiu.ac.ir

the stress-strength models in complete samples, much consideration has not been paid to censored data (see [3]).

Type-I and Type-II censoring schemes are two most fundamental schemes and by mixing of these two schemes, hybrid scheme is derived. Unfortunately, none of above schemes cannot remove active units during the experiment. So, the progressive censoring scheme is mentioned. Combining hybrid and progressive schemes, hybrid progressive scheme is provided which introduced by Kundu and Joarder [4]. One of the most objectionable of this scheme is that the sample size is random and a very small number may be turn out under this scheme. So, Ng *et al.* [6] introduced the adaptive hybrid progressive scheme, so that in this scheme the sample size is fixed. The adaptive Type-II hybrid progressive censoring scheme (AT-II HPC) can be described as follows: Suppose that  $X_{1:n:N}, \dots, X_{n:n:N}$  be a progressive censoring sample and  $T > 0$  is fixed. In this condition, if  $X_{n:n:N} < T$ , the experiment ends at time  $X_{n:n:N}$  and  $n$  failures,  $X_{1:n:N}, \dots, X_{n:n:N}$ , with the progressive censoring scheme  $(R_1, \dots, R_n)$ , achieved. Also, if  $X_{J:n:N} < T < X_{J+1:n:N}$ , then we will not withdraw any items from the experiment by setting  $R_{J+1} = \dots = R_{n-1} = 0$ ,  $R_n = N - n - \sum_{i=1}^J R_i$ . We denote an AT-II HPC sample with  $\{X_1, \dots, X_n\}$  under the scheme  $\{N, n, T, J, R_1, \dots, R_n\}$ . The likelihood function of the AT-II HPC samples is as follows:

$$L(\theta) \propto \prod_{i=1}^n f(x_i) \prod_{i=1}^J [1 - F(x_i)]^{R_i} [1 - F(x_n)]^{R_n}.$$

Generalized Inverted Exponential (GIE) distribution with the shape and scale parameters  $\alpha$  and  $\lambda$ , respectively, has the probability density function as  $f(x) = \frac{\alpha\lambda}{x^2} e^{-\frac{\lambda}{x}} (1 - e^{-\frac{\lambda}{x}})^{\alpha-1}$ ,  $x, \alpha, \lambda > 0$ . In this paper, we obtain the Bayesian inference of the  $R = P(X < Y)$  based on AT-II HPC sample, when  $X$  and  $Y$  are two independent random variables from the GIE distribution.

## 2 Bayesian inference of $R$ with unknown common $\lambda$

If  $X \sim GIE(\alpha, \lambda)$  and  $Y \sim GIE(\beta, \lambda)$ , then the stress-strength parameter can be obtained as

$$R = P(X < Y) = \frac{\alpha}{\alpha + \beta}.$$

In this section, the Bayesian inference of  $R$  is considered under squared error loss functions, when  $\alpha$ ,  $\beta$  and  $\lambda$  are independent gamma random variables. Based on the observed censoring samples, the joint posterior density function is as follows:

$$\pi(\alpha, \beta, \lambda | \text{data}) \propto L(\text{data} | \alpha, \beta, \lambda) \pi_1(\alpha) \pi_2(\beta) \pi_3(\lambda) \quad (1)$$

where  $\pi_1(\alpha) \propto \alpha^{a_1-1} e^{-b_1\alpha}$ ,  $\alpha, a_1, b_1 > 0$ ,  $\pi_2(\beta) \propto \beta^{a_2-1} e^{-b_2\beta}$ ,  $\beta, a_2, b_2 > 0$  and  $\pi_3(\lambda) \propto \lambda^{a_3-1} e^{-b_3\lambda}$ ,  $\lambda, a_3, b_3 > 0$ . As we see, from equation (1), the Bayes estimate cannot be obtained in a closed form. So, we should approximate it by applying two methods:

- Lindley's approximation,
- MCMC method.

## 2.1 Lindley's approximation

Lindley [5] introduced one of the most numerical techniques to derive the Bayes estimate. If  $U(\theta)$  be a function of  $\theta = (\theta_1, \theta_2, \theta_3)$ , Lindley's approximation of it,  $\mathbb{I}(\text{data})$ , is

$$\begin{aligned}\mathbb{I}(\text{data}) &= u + (u_1 d_1 + u_2 d_2 + u_3 d_3 + d_4 + d_5) + \frac{1}{2}[A(u_1 \sigma_{11} + u_2 \sigma_{12} + u_3 \sigma_{13}) \\ &\quad + B(u_1 \sigma_{21} + u_2 \sigma_{22} + u_3 \sigma_{23}) + C(u_1 \sigma_{31} + u_2 \sigma_{32} + u_3 \sigma_{33})],\end{aligned}$$

calculated at  $\widehat{\theta} = (\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3)$ , where  $\ell(\theta)$  is the logarithm of the likelihood function, and  $\rho(\theta)$  is the logarithm of the prior density of  $\theta$ . Also,  $u_i = \partial u(\theta)/\partial\theta_i$ ,  $u_{ij} = \partial^2 u(\theta)/\partial\theta_i\partial\theta_j$ ,  $\ell_{ijk} = \partial^3 \ell(\theta)/\partial\theta_i\partial\theta_j\partial\theta_k$ ,  $\rho_j = \partial\rho(\theta)/\partial\theta_j$ , and  $\sigma_{ij} = (i, j)$ th element in the inverse of matrix  $[-\ell_{ij}]$  all evaluated at the MLE of the parameters. Moreover,

$$\begin{aligned}d_i &= \rho_1 \sigma_{i1} + \rho_2 \sigma_{i2} + \rho_3 \sigma_{i3}, \quad i = 1, 2, 3, \\ d_4 &= u_{12} \sigma_{12} + u_{13} \sigma_{13} + u_{23} \sigma_{23}, \\ d_5 &= \frac{1}{2}(u_{11} \sigma_{11} + u_{22} \sigma_{22} + u_{33} \sigma_{33}), \\ A &= \ell_{111} \sigma_{11} + 2\ell_{121} \sigma_{12} + 2\ell_{131} \sigma_{13} + 2\ell_{231} \sigma_{23} + \ell_{221} \sigma_{22} + \ell_{331} \sigma_{33}, \\ B &= \ell_{112} \sigma_{11} + 2\ell_{122} \sigma_{12} + 2\ell_{132} \sigma_{13} + 2\ell_{232} \sigma_{23} + \ell_{222} \sigma_{22} + \ell_{332} \sigma_{33}, \\ C &= \ell_{113} \sigma_{11} + 2\ell_{123} \sigma_{12} + 2\ell_{133} \sigma_{13} + 2\ell_{233} \sigma_{23} + \ell_{223} \sigma_{22} + \ell_{333} \sigma_{33}.\end{aligned}$$

In our case, for  $(\theta_1, \theta_2, \theta_3) \equiv (\alpha, \beta, \lambda)$ , we have

$$\begin{aligned}\rho_1 &= \frac{a_1 - 1}{\alpha} - b_1, \quad \rho_2 = \frac{a_2 - 1}{\beta} - b_2, \quad \rho_3 = \frac{a_3 - 1}{\lambda} - b_3, \\ \ell_{11} &= -\frac{n}{\alpha^2}, \quad \ell_{22} = -\frac{m}{\beta^2}, \quad \ell_{12} = \ell_{21} = 0, \\ \ell_{13} &= \ell_{31} = -\sum_{i=1}^n \frac{1}{x_i(1 - e^{\frac{\lambda}{x_i}})} - \sum_{i=1}^{J_1} \frac{R_i}{x_i(1 - e^{\frac{\lambda}{x_i}})} - \frac{R_n}{x_n(1 - e^{\frac{\lambda}{x_n}})}, \\ \ell_{23} &= \ell_{32} = -\sum_{j=1}^m \frac{1}{y_j(1 - e^{\frac{\lambda}{y_j}})} - \sum_{j=1}^{J_2} \frac{S_j}{y_j(1 - e^{\frac{\lambda}{y_j}})} - \frac{S_m}{y_m(1 - e^{\frac{\lambda}{y_m}})}, \\ \ell_{33} &= \frac{-n}{\lambda^2} - (\alpha - 1) \sum_{i=1}^n \frac{e^{\frac{\lambda}{x_i}}}{x_i^2(1 - e^{\frac{\lambda}{x_i}})^2} - \alpha \left[ \sum_{i=1}^{J_1} \frac{R_i e^{\frac{\lambda}{x_i}}}{x_i^2(1 - e^{\frac{\lambda}{x_i}})^2} + \frac{R_n e^{\frac{\lambda}{x_n}}}{x_n^2(1 - e^{\frac{\lambda}{x_n}})^2} \right] \\ &\quad - \frac{m}{\lambda^2} - (\beta - 1) \sum_{j=1}^m \frac{e^{\frac{\lambda}{y_j}}}{y_j^2(1 - e^{\frac{\lambda}{y_j}})^2} - \beta \left[ \sum_{j=1}^{J_2} \frac{S_j e^{\frac{\lambda}{y_j}}}{y_j^2(1 - e^{\frac{\lambda}{y_j}})^2} + \frac{S_m e^{\frac{\lambda}{y_m}}}{y_m^2(1 - e^{\frac{\lambda}{y_m}})^2} \right],\end{aligned}$$

$\sigma_{ij}$ ,  $i, j = 1, 2, 3$  are obtained by using  $\ell_{ij}$ ,  $i, j = 1, 2, 3$  and

$$\begin{aligned}\ell_{111} &= \frac{2n}{\alpha^3}, & \ell_{222} &= \frac{2m}{\beta^3} \\ \ell_{133} = \ell_{331} = \ell_{313} &= -\sum_{i=1}^n \frac{e^{\frac{\lambda}{x_i}}}{x_i^2(1-e^{\frac{\lambda}{x_i}})^2} - \sum_{i=1}^{J_1} \frac{R_i e^{\frac{\lambda}{x_i}}}{x_i^2(1-e^{\frac{\lambda}{x_i}})^2} - \frac{R_n e^{\frac{\lambda}{x_n}}}{x_n^2(1-e^{\frac{\lambda}{x_n}})^2}, \\ \ell_{233} = \ell_{332} = \ell_{323} &= -\sum_{i=1}^n \frac{e^{\frac{\lambda}{y_j}}}{y_j^2(1-e^{\frac{\lambda}{y_j}})^2} - \sum_{j=1}^{J_2} \frac{S_j e^{\frac{\lambda}{y_j}}}{y_j^2(1-e^{\frac{\lambda}{y_j}})^2} + \frac{S_m e^{\frac{\lambda}{y_m}}}{y_m^2(1-e^{\frac{\lambda}{y_m}})^2}, \\ \ell_{333} &= \frac{2n}{\lambda^3} - (\alpha-1) \sum_{i=1}^n \frac{e^{\frac{\lambda}{x_i}}(e^{\frac{\lambda}{x_i}}+1)}{x_i^3(1-e^{\frac{\lambda}{x_i}})^3} - \alpha \left[ \sum_{i=1}^{J_1} R_i \frac{e^{\frac{\lambda}{x_i}}(e^{\frac{\lambda}{x_i}}+1)}{x_i^3(1-e^{\frac{\lambda}{x_i}})^3} + R_n \frac{e^{\frac{\lambda}{x_n}}(e^{\frac{\lambda}{x_n}}+1)}{x_n^3(1-e^{\frac{\lambda}{x_n}})^3} \right] \\ &\quad + \frac{2m}{\lambda^3} - (\beta-1) \sum_{j=1}^m \frac{e^{\frac{\lambda}{y_j}}(e^{\frac{\lambda}{y_j}}+1)}{y_j^3(1-e^{\frac{\lambda}{y_j}})^3} - \beta \left[ \sum_{j=1}^{J_2} S_j \frac{e^{\frac{\lambda}{y_j}}(e^{\frac{\lambda}{y_j}}+1)}{y_j^3(1-e^{\frac{\lambda}{y_j}})^3} + S_m \frac{e^{\frac{\lambda}{y_m}}(e^{\frac{\lambda}{y_m}}+1)}{y_m^3(1-e^{\frac{\lambda}{y_m}})^3} \right],\end{aligned}$$

and other  $\ell_{ijk} = 0$ . Hence,

$$\begin{aligned}A &= \ell_{111}\sigma_{11} + \ell_{331}\sigma_{33}, & B &= \ell_{222}\sigma_{22} + \ell_{332}\sigma_{33}, \\ C &= 2\ell_{133}\sigma_{13} + 2\ell_{233}\sigma_{23} + \ell_{333}\sigma_{33}d_4 = u_{12}\sigma_{12}, & d_5 &= \frac{1}{2}(u_{11}\sigma_{11} + u_{22}\sigma_{22}).\end{aligned}$$

So, the approximate Bayes estimate of  $R$ , under the squared error loss function is obtained by setting  $u(\theta) = R = \frac{\alpha}{\alpha+\beta}$ . Also,  $u_3 = 0$ ,  $u_{i3} = 0$ ,  $i = 1, 2, 3$  and

$$u_1 = \frac{\beta}{(\alpha+\beta)^2}, \quad u_2 = \frac{-\alpha}{(\alpha+\beta)^2}, \quad u_{11} = \frac{-2\beta}{(\alpha+\beta)^3}, \quad u_{12} = \frac{-2(\alpha-\beta)}{(\alpha+\beta)^3}, \quad u_{22} = \frac{2\alpha}{(\alpha+\beta)^3}.$$

Consequently, under the squared error loss function, the Bayes estimate of  $R$  is

$$\begin{aligned}\widehat{R}^{LB} &= \mathbb{E}(u(\theta)|\text{data}) = u(\theta) + [u_1 d_1 + u_2 d_2 + d_4 + d_5] + \frac{1}{2}[A(u_1\sigma_{11} + u_2\sigma_{12}) \\ &\quad + B(u_1\sigma_{21} + u_2\sigma_{22}) + C(u_1\sigma_{31} + u_2\sigma_{32})].\end{aligned}\tag{2}$$

Notice that all parameters are calculated at  $(\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda})$ .

## 2.2 MCMC method

From the equation (1), the posterior pdfs of  $\alpha$ ,  $\beta$  and  $\lambda$  can be derived as:

$$\begin{aligned}\alpha|\lambda, \text{data} &\sim \Gamma(n+a_1, b_1 - V(\lambda)), \\ \beta|\lambda, \text{data} &\sim \Gamma(m+a_2, b_2 - U(\lambda)), \\ \pi(\lambda|\alpha, \beta, \text{data}) &\propto \lambda^{n+m+a_3-1} e^{-\lambda \left( \sum_{i=1}^n \frac{1}{x_i} + \sum_{j=1}^m \frac{1}{y_j} + b_3 \right)} \\ &\quad \times e^{\alpha V(\lambda) + \beta U(\lambda)} \prod_{i=1}^n (1 - e^{-\frac{\lambda}{x_i}})^{-1} \prod_{j=1}^m (1 - e^{-\frac{\lambda}{y_j}})^{-1},\end{aligned}$$

where

$$\begin{aligned} V(\lambda) &= \sum_{i=1}^n \log(1 - e^{-\frac{\lambda}{x_i}}) + \sum_{i=1}^{J_1} R_i \log(1 - e^{-\frac{\lambda}{x_i}}) + R_n \log(1 - e^{-\frac{\lambda}{x_n}}), \\ U(\lambda) &= \sum_{j=1}^m \log(1 - e^{-\frac{\lambda}{y_j}}) + \sum_{j=1}^{J_2} S_j \log(1 - e^{-\frac{\lambda}{y_j}}) + S_m \log(1 - e^{-\frac{\lambda}{y_m}}). \end{aligned} \quad (3)$$

It is observed that generating samples from the posterior pdf of  $\lambda$  should be done by the Metropolis-Hastings method. So, we propose the following algorithm of Gibbs sampling:

1. Start with initial values  $(\alpha_{(0)}, \beta_{(0)}, \lambda_{(0)})$ .
2. Set  $t = 1$ .
3. Generate  $\lambda_{(t)}$  from  $\pi(\lambda | \alpha_{(t-1)}, \beta_{(t-1)}, \text{data})$ , using Metropolis-Hastings method.
4. Generate  $\alpha_{(t)}$  from  $\Gamma(n + a_1, b_1 - V(\lambda_{(t-1)}))$ .
5. Generate  $\beta_{(t)}$  from  $\Gamma(m + a_2, b_2 - U(\lambda_{(t-1)}))$ .
6. Evaluate  $R_t = \frac{\alpha_t}{\alpha_t + \beta_t}$ .
7. Set  $t = t + 1$ .
8. Repeat T times, steps 3-7.

Therefore, the Bayes estimate of  $R$ , under the squared error loss functions is:

$$\widehat{R}^{MB} = \frac{1}{T} \sum_{t=1}^T R_t. \quad (4)$$

Also, the  $100(1 - \gamma)\%$  HPD credible interval of  $R$  can be constructed, using the method of Chen and Shao [1].

### 3 Bayesian inference of $R$ with known common $\lambda$

In this section, the Bayesian inference of  $R$  is considered under the squared error loss function, when  $\alpha$  and  $\beta$  are independent gamma random variables. Based on the observed censoring samples, the joint posterior density function is as follows:

$$\pi(\alpha, \beta | \lambda, \text{data}) = \frac{(\alpha(V(\lambda) + b_1))^{n+a_1} (\beta(U(\lambda) + b_2))^{m+a_2}}{\alpha \beta \Gamma(n + a_1) \Gamma(m + a_2)} e^{-\alpha(V(\lambda) + b_1) - \beta(U(\lambda) + b_2)}, \quad (5)$$

where  $V$  and  $U$  are given in (3). So, the Bayes estimate of  $R$  under the squared error loss function, should be achieved by solving the following integral:

$$\widehat{R}^B = \int_0^\infty \int_0^\infty \frac{\alpha}{\alpha + \beta} \times \pi(\alpha, \beta | \lambda, \text{data}) d\alpha d\beta.$$

By applying the idea of Kizilaslan and Nadar [2], the exact Bayes estimate is obtained as:

$$\hat{R}^B = \begin{cases} \frac{(1-z)^{n+a_1}(n+a_1)}{w} {}_2F_1(w, n+a_1+1; w+1, z) & \text{if } |z| < 1, \\ \frac{(n+a_1)}{(1-z)^{m+a_2}w} {}_2F_1(w, m+a_2; w+1, \frac{z}{1-z}) & \text{if } z < -1, \end{cases} \quad (6)$$

where  $w = n + m + a_1 + a_2$ ,  $z = 1 - \frac{V(\lambda) + b_1}{U(\lambda) + b_2}$  and

$${}_2F_1(\alpha, \beta; \gamma, z) = \frac{1}{B(\beta, \gamma - \beta)} \int_0^1 t^{\beta-1} (1-t)^{\gamma-\beta-1} (1-tz)^{-\alpha} dt, \quad |z| < 1$$

is hypergeometric series, which is quickly calculated and readily available in standard software such as MATLAB. Furthermore, the  $100(1-\gamma)\%$  Bayesian interval of  $R$  can be constructed as  $(L, U)$ , where  $L$  and  $U$  should be satisfied, respectively, in

$$\int_0^L f_R(R) dR = \frac{\gamma}{2}, \quad \int_0^U f_R(R) dR = 1 - \frac{\gamma}{2}, \quad (7)$$

where  $f_R(R)$ , using change-of-variable method, can be earned by (5) as

$$f_R(R) = \frac{(1-z)^{n+a_1} R^{n+a_1-1} (1-R)^{m+a_2-1} (1-Rz)^{-w}}{B(n+a_1, m+a_2)}, \quad 0 < R < 1.$$

## 4 Bayesian inference of $R$ in general case

If  $X \sim GIE(\alpha, \lambda_1)$  and  $Y \sim GIE(\beta, \lambda_2)$ , then the stress-strength parameter can be obtained as

$$R = P(X < Y) = 1 - \int_0^\infty \frac{\beta \lambda_2}{y^2} e^{-\frac{\lambda_2}{y}} (1 - e^{-\frac{\lambda_2}{y}})^{\beta-1} (1 - e^{-\frac{\lambda_1}{y}})^\alpha dy.$$

In this section, the Bayesian inference of  $R$  is considered under squared error loss functions, when  $\alpha, \beta, \lambda_1$  and  $\lambda_2$  are independent gamma random variables. Like in section 2, as the Bayes estimate of  $R$  can not be evaluated in a closed form, it is approximated by MCMC method. From the joint posterior density function, we can be derived the posterior pdfs of  $\alpha, \beta, \lambda_1$  and  $\lambda_2$  as follows:

$$\begin{aligned} \alpha | \lambda_1, \text{data} &\sim \Gamma(n + a_1, b_1 - V(\lambda_1)), \\ \beta | \lambda_2, \text{data} &\sim \Gamma(m + a_2, b_2 - U(\lambda_2)), \\ \pi(\lambda_1 | \alpha, \text{data}) &\propto \lambda_1^{n+a_3-1} e^{-\lambda_1 \left( \sum_{i=1}^n \frac{1}{x_i} + b_3 \right)} e^{\alpha V(\lambda_1)} \prod_{i=1}^n (1 - e^{-\frac{\lambda_1}{x_i}})^{-1}, \\ \pi(\lambda_2 | \beta, \text{data}) &\propto \lambda_2^{m+a_4-1} e^{-\lambda_2 \left( \sum_{j=1}^m \frac{1}{y_j} + b_4 \right)} e^{\beta U(\lambda_2)} \prod_{j=1}^m (1 - e^{-\frac{\lambda_2}{y_j}})^{-1}. \end{aligned}$$

It is observed that generating samples from the posterior pdfs of  $\lambda_1$  and  $\lambda_2$  should be done by the Metropolis-Hastings method. So, we propose the following algorithm of Gibbs sampling:

1. Start with initial values  $(\alpha_{(0)}, \beta_{(0)}, \lambda_{1(0)}, \lambda_{2(0)})$ .
2. Set  $t = 1$ .
3. Generate  $\lambda_{1(t)}$  from  $\pi(\lambda_1 | \alpha_{(t-1)}, \text{data})$ , using Metropolis-Hastings method.
4. Generate  $\lambda_{2(t)}$  from  $\pi(\lambda_2 | \beta_{(t-1)}, \text{data})$ , using Metropolis-Hastings method.
5. Generate  $\alpha_{(t)}$  from  $\Gamma(n + a_1, b_1 - V(\lambda_{1(t-1)}))$ .
6. Generate  $\beta_{(t)}$  from  $\Gamma(m + a_2, b_2 - U(\lambda_{2(t-1)}))$ .
7. Evaluate  $R_t = 1 - \int_0^\infty \frac{\beta_{(t)} \lambda_{2(t)}}{y^2} e^{-\frac{\lambda_{2(t)}}{y}} (1 - e^{-\frac{\lambda_{2(t)}}{y}})^{\beta_{(t)}-1} (1 - e^{-\frac{\lambda_{1(t)}}{y}})^{\alpha_{(t)}} dy$ .
8. Set  $t = t + 1$ .
9. Repeat T times, steps 3-8.

Therefore, the Bayes estimate of  $R$ , under the squared error loss functions is:

$$\widehat{R}^{MB} = \frac{1}{T} \sum_{t=1}^T R_t. \quad (8)$$

Also, the  $100(1 - \gamma)\%$  HPD credible interval of  $R$  can be constructed, using the method of Chen and Shao [1].

## 5 Simulation study

We consider the performance of different Bayes estimates, under AT-II HPC schemes by using the Monte Carlo simulations. The different estimates, in terms of mean squared errors (MSEs) are compared together and the different confidence intervals, in terms of average confidence lengths and coverage percentages are compared together. Based on 3000 replications and  $T = 0.9, 1.5$ , all results are gathered. Also, the used censoring schemes are as:

$$\begin{aligned} (N, n) &= (20, 10), \quad r_1 : (1^{*10}), \quad r_2 : (2^{*5}, 0^{*5}), \quad r_3 : (0^{*5}, 2^{*5}), \\ (N, n) &= (30, 10), \quad r_4 : (2^{*10}), \quad r_5 : (4^{*5}, 0^{*5}), \quad r_6 : (0^{*5}, 4^{*5}), \\ (N, n) &= (30, 20), \quad r_7 : (1^{*10}, 0^{*10}), \quad r_8 : (2^{*5}, 0^{*15}), \quad r_9 : (0^{*10}, 1^{*10}), \\ (N, n) &= (40, 30), \quad r_{10} : (1^{*10}, 0^{*20}), \quad r_{11} : (0^{*10}, 1^{*10}, 0^{*10}), \quad r_{12} : (0^{*20}, 1^{*10}). \end{aligned}$$

In the first case, with unknown common  $\lambda$ , the parameter values  $(\alpha, \beta, \lambda) = (2, 2, 2)$  are used to obtain the simulation results. Also, the Bayesian inference is considered by assuming three priors as Prior 1:  $a_j = 0, b_j = 0, j = 1, 2, 3$ , Prior 2:  $a_j = 1, b_j = 0.1, j = 1, 2, 3$ . Under the above hypotheses, the MSEs of Bayesian estimates of  $R$ , via Linldey's approximation and MCMC method are derived by (2) and (4), respectively. Also, we derived the 95% HPD intervals for  $R$ . The simulation results are given in Table 1.

Table 1: Simulation results

$(n, N, T)$	C.S.	Unknown common $\lambda$							
		Bayes (MCMC)			Prior 2			Prior 1	Prior 2
		MSE	C.I	C.P	MSE	C.I	C.P	MSE	MSE
(10, 20, 0.9)	$(r_1, r_1)$	0.0084	0.4074	0.904	0.0080	0.3879	0.915	0.0091	0.0089
	$(r_2, r_2)$	0.0080	0.4229	0.919	0.0074	0.3955	0.919	0.0115	0.0108
	$(r_3, r_3)$	0.0061	0.4056	0.901	0.0058	0.3867	0.907	0.0066	0.0058
	$(r_1, r_2)$	0.0145	0.4045	0.908	0.0135	0.3885	0.908	0.0173	0.0152
	$(r_1, r_3)$	0.0106	0.4087	0.903	0.0099	0.3986	0.912	0.0102	0.0092
(10, 30, 1.5)	$(r_4, r_4)$	0.0043	0.4069	0.901	0.0041	0.3821	0.922	0.0060	0.0057
	$(r_5, r_5)$	0.0083	0.4102	0.902	0.0078	0.3941	0.920	0.0127	0.0098
	$(r_6, r_6)$	0.0071	0.4068	0.914	0.0069	0.3872	0.929	0.0110	0.0092
	$(r_4, r_5)$	0.0276	0.4000	0.909	0.0267	0.3851	0.911	0.0370	0.0355
	$(r_4, r_6)$	0.0201	0.3844	0.924	0.0195	0.3688	0.922	0.0263	0.0239
(20, 30, 0.9)	$(r_7, r_7)$	0.0030	0.3023	0.940	0.0027	0.2954	0.944	0.0071	0.0047
	$(r_8, r_8)$	0.0049	0.2981	0.931	0.0045	0.2922	0.942	0.0110	0.0095
	$(r_9, r_9)$	0.0016	0.3003	0.936	0.0015	0.2920	0.937	0.0016	0.0016
	$(r_7, r_8)$	0.0063	0.3005	0.932	0.0058	0.2936	0.939	0.0230	0.0106
	$(r_7, r_9)$	0.0085	0.3020	0.931	0.0080	0.2934	0.941	0.0115	0.0091
(30, 40, 1.5)	$(r_{10}, r_{10})$	0.0018	0.2444	0.951	0.0017	0.2434	0.9520	0.0031	0.0026
	$(r_{11}, r_{11})$	0.0017	0.2509	0.948	0.0016	0.2449	0.9490	0.0029	0.0026
	$(r_{12}, r_{12})$	0.0006	0.2501	0.943	0.0006	0.2461	0.9500	0.0007	0.0007
	$(r_{10}, r_{11})$	0.0042	0.2407	0.943	0.0041	0.2382	0.9450	0.0066	0.0060
	$(r_{10}, r_{12})$	0.0068	0.2330	0.943	0.0066	0.2315	0.9490	0.0104	0.0101
$(n, N, T)$	C.S.	Known common $\lambda$							
		Bayes(exact)							
		MSE	C.I	C.P	MSE	C.I	C.P		
(10, 20, 1.5)	$(r_1, r_1)$	0.0116	0.4163	0.919	0.0106	0.3986	0.921		
	$(r_2, r_2)$	0.0072	0.4100	0.917	0.0066	0.3932	0.928		
	$(r_3, r_3)$	0.0042	0.4157	0.915	0.0040	0.3974	0.929		
	$(r_1, r_2)$	0.0147	0.3988	0.917	0.0138	0.3825	0.923		
	$(r_1, r_3)$	0.0101	0.4063	0.919	0.0096	0.3894	0.929		
(10, 30, 0.9)	$(r_4, r_4)$	0.0046	0.4153	0.916	0.0043	0.3979	0.926		
	$(r_5, r_5)$	0.0102	0.4071	0.915	0.0091	0.3904	0.925		
	$(r_6, r_6)$	0.0030	0.4178	0.918	0.0028	0.3990	0.919		
	$(r_4, r_5)$	0.0282	0.3797	0.916	0.0258	0.3668	0.928		
	$(r_4, r_6)$	0.0149	0.3996	0.918	0.0140	0.3846	0.920		
(20, 30, 1.5)	$(r_7, r_7)$	0.0034	0.3011	0.939	0.0032	0.2938	0.942		
	$(r_8, r_8)$	0.0032	0.3005	0.937	0.0031	0.2934	0.940		
	$(r_9, r_9)$	0.0015	0.3026	0.932	0.0014	0.2957	0.938		
	$(r_7, r_8)$	0.0040	0.3000	0.937	0.0038	0.2930	0.940		
	$(r_7, r_9)$	0.0121	0.2894	0.933	0.0117	0.2829	0.943		
(30, 40, 0.9)	$(r_{10}, r_{10})$	0.0029	0.2473	0.946	0.0027	0.2433	0.947		
	$(r_{11}, r_{11})$	0.0017	0.2480	0.941	0.0016	0.2449	0.943		
	$(r_{12}, r_{12})$	0.0016	0.2482	0.946	0.0015	0.2449	0.948		
	$(r_{10}, r_{11})$	0.0065	0.2439	0.940	0.0060	0.2404	0.946		
	$(r_{10}, r_{12})$	0.0056	0.2442	0.942	0.0052	0.2408	0.945		
$(n, N, T)$	C.S.	General Case							
		Bayes(MCMC)							
		MSE	C.I	C.P	MSE	C.I	C.P		
(10, 20, 1.5)	$(r_1, r_1)$	0.0026	0.4185	0.909	0.0025	0.4000	0.915		
	$(r_2, r_2)$	0.0073	0.4107	0.905	0.0066	0.3963	0.918		
	$(r_3, r_3)$	0.0033	0.4184	0.906	0.0031	0.3973	0.913		
	$(r_1, r_2)$	0.0111	0.4020	0.910	0.0105	0.3859	0.917		
	$(r_1, r_3)$	0.0070	0.4115	0.902	0.0068	0.3939	0.913		
(20, 30, 1.5)	$(r_7, r_7)$	0.0022	0.3001	0.928	0.0021	0.2944	0.935		
	$(r_8, r_8)$	0.0025	0.3018	0.923	0.0023	0.2947	0.928		
	$(r_9, r_9)$	0.0010	0.3046	0.928	0.0010	0.2952	0.931		
	$(r_7, r_8)$	0.0034	0.2997	0.926	0.0032	0.2959	0.931		
	$(r_7, r_9)$	0.0098	0.2945	0.924	0.0094	0.2884	0.934		
(30, 40, 0.9)	$(r_{10}, r_{10})$	0.0021	0.2474	0.934	0.0020	0.2424	0.942		
	$(r_{11}, r_{11})$	0.0016	0.2487	0.933	0.0015	0.2453	0.945		
	$(r_{12}, r_{12})$	0.0010	0.2481	0.937	0.0009	0.2442	0.942		
	$(r_{10}, r_{11})$	0.0067	0.2430	0.930	0.0063	0.2400	0.944		
	$(r_{10}, r_{12})$	0.0066	0.2426	0.935	0.0063	0.2399	0.945		

In the second case, with known common  $\lambda$ , the parameter values  $(\alpha, \beta) = (2, 2)$  are used to obtain the simulation results. Also, the Bayesian inference is considered by assuming three priors as Prior 3:  $a_j = 0, b_j = 0, j = 1, 2$ , Prior 4:  $a_j = 1, b_j = 0.1, j = 1, 2$ . Under the above hypotheses, the Bayes estimate and Bayesian intervals of  $R$  are derived by (6) and (7), respectively. The results are provided in Table 1.

In the third case, with unknown different  $\lambda_1$  and  $\lambda_2$ , the parameter values  $(\alpha, \beta, \lambda_1, \lambda_2) = (2, 2, 2, 2)$  are used to obtain the simulation results. Also, the Bayesian inference are considered by assuming three priors as Prior 5:  $a_j = 0, b_j = 0, j = 1, 2, 3, 4$ , Prior 6:  $a_j = 1, b_j = 0.1, j = 1, 2, 3, 4$ . Under the above hypotheses, the MSEs of Bayesian estimates of  $R$  are derived by (8). Also, we derived the 95% HPD intervals for  $R$ . The simulation results are given in Table 1.

From Table 1, we observed that the best performance, in terms of MSE, belong to informative priors (priors 2, 4 and 6). Furthermore, in first case, performance of Bayes estimates which obtained by MCMC method are generally better than those obtained by Lindleys approximation. Also, we observed that the best performance among the different intervals belong to HPD intervals based on informative priors (priors 2, 4 and 6).

## References

- [1] Chen, M. H. and Shao, Q. M. (1999), Monte Carlo estimation of Bayesian Credible and HPD intervals. *Journal of Computational and Graphical Statistics*, **8**, 69-92.
- [2] Kizilaslan, F. and Nadar, M. (2016), Estimation of reliability in a multicomponent stress-strength model based on a bivariate Kumaraswamy distribution. *Statistical Papers*, Accepted. DOI: 10.1007/s00362-016-0765-8.
- [3] Kohansal, A. (2017), On estimation of reliability in a multicomponent stress-strength model for a Kumaraswamy distribution based on progressively censored sample, *Statistical Papers*, Accepted. DOI: 10.1007/s00362-017-0916-6.
- [4] Kundu, D. and Joarder, A. (2006), Analysis of type-II progressively hybrid censored data, *Computational Statistics and Data Analysis*, **50**, 2509-2528.
- [5] Lindley, D. V. (1980), Approximate Bayesian methods. *Trabajos de Estadistica*, **3**, 281-288.
- [6] Ng, H. K. T., Kundu, D. and Chan, P. S. (2009), Statistical analysis of exponential lifetimes under an adaptive Type-II progressively censoring scheme. *Naval Research Logistics*, **56**, 687-698.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



# An additive-multiplicative mean past life regression model

Mansourvar, Z.<sup>1</sup>

Department of Statistics, University of Isfahan

## Abstract

The mean past lifetime measures the expected time elapsed since the failure of a subject till the time of observation. In this paper, we propose an additive-multiplicative mean past life model to study the association between the mean past life function and potential regression covariates in the presence of left censoring. This model extends the proportional mean past life model using an additive model as its covariate dependent baseline. For the suggested model, some covariate effects are allowed to be time-varying. To estimate the model parameters, martingale estimating equations are developed, and the asymptotic properties of the resulting estimators are established.

**Keywords:** additive model, counting process, left censoring, martingale estimating equation, mean past lifetime, proportional model, reversed hazard rate, time-varying effect.

## 1 Introduction

In survival analysis and reliability researches, the hazard rate and the mean residual lifetime are the basic tools for studying the characteristics of lifetime distribution under right censored data. However in survival studies, there are some occasions where lifetime data are left censored. For example, when we are following persons who have undergone a medical test to check for a certain disease, we may record a failure when the test is positive.

---

<sup>1</sup>z.mansourvar@sci.ui.ac.ir

But we may not know exactly the time at which the patients have been infected by this disease, and therefore do not know exactly when the failure occurred. Thus, the survival time is censored on the left side. The reversed hazard rate (RHR) is more appropriate than the hazard rate to analyze such left censored data, which is defined by

$$r(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t - \Delta t < T < \Delta t | T \leq t)}{\Delta t}.$$

The RHR specifies the instantaneous rate of failure of a subject at time  $t$  given that it failed before time  $t$ . Another useful quantity to facilitate the analysis of left censored data is mean past lifetime (MPL) which is defined as

$$k(t) = E(t - T | T \leq t).$$

The MPL corresponds to the mean time elapsed since the failure of  $T$ , given that  $T \leq t$ . For instance, assume that, at time  $t$ , one has undergone a medical test to check for a certain disease, and the test is positive. Let  $T$  denote the time he(she) has been infected by this disease. Hence, it is known that  $T \leq t$ . Now the question is, how much time has elapsed since he(she) had been infected by this disease. As an another example, suppose that an item has been put under test by an engineer at time  $t = 0$ . Assume that when the engineer checks the item at time  $t$ , he (she) discovers that it has already failed. Now, the same question as of the first example can be asked. In both examples, the random variable of interest is  $T_t = t - T | T < t$ , and the MPL function is its expected value, see [3]. The MPL also known as reversed mean residual life, mean inactivity time or mean waiting time, has recently been a topic of increasing interest in the literature. [7] gave the necessary conditions for a function to be MPL. [11] defined some new classes of distributions based on the MPL, and obtained some results on the MPL ordering. Later, [8], [2], [10], [6] studied various stochastic comparisons based on MPL function and associated properties. [3] investigated the properties of the MPL in connection with other reliability measures. They also gave some results on partial ordering and characterization and finally estimated the MPL. [13] studied some implications of stochastic orders and aging notions for the proportional MPL model and for its extended mixture model.

In the presence of covariates, regression models are needed to study the association between the MPL function and potential regression covariates. As an alternative to the widely-used Cox proportional hazards model [5], the proportional MPL model can be used for fitting survival data under left censoring that takes the form

$$k(t|X) = k_0(t) \exp(X^T \beta), \quad (1)$$

where  $k(t|X) = E(t - T | T \leq t, X)$  is the MPL function corresponding to a vector covariate  $X$ ,  $k_0(t)$  is some unspecified baseline MPL function when  $X = 0$ , and  $\beta$  is an unknown vector of regression parameters. When the effects of covariates are additive rather than multiplicative, a plausible alternative is the additve MPL model, which assumes that

$$k(t|X) = k_0(t) + X^T \beta. \quad (2)$$

In this study, we propose a flexible model that combines the proportional and additive MPL model that takes the form

$$k(t|X, Z) = \{X^T \alpha(t)\} \exp(Z^T \beta), \quad (3)$$

where  $k(t|X, Z) = E(t - T|T \leq t; X, Z)$ ,  $X$ , and  $Z$  are bounded covariate vectors of dimensions  $(p+1)$  and  $q$ , respectively, and  $\alpha(t)$  and  $\beta$  are conformal vectors of unknown time-dependent and time-independent regression parameters. The first component of  $X$  is set to be 1, which contributes to the baseline MPL function. Model (3) extends the proportional MPL model (1) by allowing the baseline MPL function to depend on covariates through an additive model. It further extends the additive MPL model (2) by considering time-varying effects and therefore provides a very flexible class of models. In the presence of left censoring, we apply martingale estimating equations to estimate the parameters of model (3). The rest of the paper is organized as follows. Section 2 is devoted to semiparametric inference procedures for estimating the non-parametric component  $\alpha(t)$  and parametric component  $\beta$  in model (3) by applying martingale estimating equations in the presence of left censoring. Section 3 gives the asymptotic properties of the proposed estimators with theoretical proofs.

## 2 Inference procedures

Suppose that the lifetime random variable  $\tilde{T}$  is left censored by the random variable  $C$ . We assume that  $\tilde{T}$  is independent of  $C$  given the covariates  $X$  and  $Z$ . Under left censoring, we observe random vector  $(T, \Delta, X, Z)$  where  $T = \max(\tilde{T}, C)$  and  $\Delta = I(\tilde{T} \geq C)$  with  $I(\cdot)$  is indicator function. Let  $(T_i, \Delta_i, X_i, Z_i)$  be independent and identically distributed copies of  $(T, \Delta, X, Z)$ ,  $i = 1, 2, \dots, n$ . Furthermore, let  $N_i(t) = I(T_i \geq t, \Delta_i = 1)$ , be a counting process and  $Y_i(t) = I(T_i \leq t)$  be an at risk indicator process. Define the sigma field  $\mathcal{F}_t = \sigma\{N_i(u), Y_i(u), X_i, Z_i : 0 \leq u \leq t; i = 1, \dots, n\}$ . We denote history at an instant just after to time  $t$  by  $\mathcal{F}_{t+}$ .

For left censored data, under the assumption of independent censoring, we have

$$E\{dN_i(t)|\mathcal{F}_{t+}; \alpha_0(\cdot), \beta_0\} = Y_i(t)dR_i(t; \alpha_0, \beta_0),$$

where  $\alpha_0(t)$  and  $\beta_0$  are the true values of the parameters  $\alpha(t)$  and  $\beta$  in (3), respectively, and  $R(t) = \int_t^\tau r(u)du$  and  $\tau = \sup(t; F(t) < 1)$ . Let

$$dM_i(t; \alpha_0, \beta_0) = dN_i(t) - Y_i(t)dR_i(t; X_i, Z_i; \alpha_0, \beta_0), \quad (i = 1, \dots, n),$$

where  $\{M_i(t; \alpha_0, \beta_0), t \geq 0\}$ , are zero-mean martingale with respect to  $\mathcal{F}_t$ . Therefore without assuming any particular form for the time-varying effect  $\alpha(t)$ , it is natural to

estimate  $\alpha_0(t)$  and  $\beta_0$  from estimating equations parallel to the partial score equations

$$\begin{aligned} \sum_{i=1}^n Y_i(t) X_i \{dN_i(t) - Y_i(t)dR_i(t|X_i, Z_i; \alpha, \beta)\} &= 0, \quad (0 \leq t < \tau), \\ \sum_{i=1}^n \int_0^\tau Y_i(t) Z_i \{dN_i(t) - Y_i(t)dR_i(t|X_i, Z_i; \alpha, \beta)\} &= 0. \end{aligned} \tag{4}$$

It can be shown that the distribution function of  $T$  given  $X$  and  $Z$  is

$$F(t|X, Z) = \exp \left\{ - \int_t^\infty \frac{1 - k'(s|X, Z)}{k(s|X, Z)} ds \right\}, \tag{5}$$

where  $k'(s|X, Z) = dk(s|X, Z)/ds$ . Then under model (3), one obtains that

$$X_i^T \alpha(t) dR_i(t|X_i, Z_i; \alpha, \beta) = \exp(-Z_i^T \beta) dt - X_i^T d\alpha(t), \quad (i = 1, \dots, n).$$

Therefore, by analogy with (4), the following estimating equations can be used for estimating  $\alpha_0(t)$  and  $\beta_0$  in model (3), respectively

$$\sum_{i=1}^n Y_i(t) X_i [X_i^T \alpha(t) dN_i(t) - Y_i(t) \{\exp(-Z_i^T \beta) dt - X_i^T d\alpha(t)\}] = 0, \tag{6}$$

$$\sum_{i=1}^n \int_0^\tau Y_i(t) Z_i [X_i^T \alpha(t) dN_i(t) - Y_i(t) \{\exp(-Z_i^T \beta) dt - X_i^T d\alpha(t)\}] = 0. \tag{7}$$

In view of equation (6), it can be obtained that

$$d\alpha(t) = \{X^T(t)X(t)\}^{-1} X^T(t) \left\{ \exp(-Z\beta) dt - d\tilde{N}(t)X(t)\alpha(t) \right\}, \tag{8}$$

where  $X(t) = \{Y_1(t)X_1, \dots, Y_n(t)X_n\}^T$  is a design matrix of dimension  $n \times (p+1)$ ,  $d\tilde{N}(t) = \text{diag}\{dN_i(t)\}$  and  $\exp(-Z\beta) = \{\exp(-Z_1^T \beta), \dots, \exp(-Z_n^T \beta)\}^T$ . We make the convention that  $\{X^T(t)X(t)\}^{-1}$  is 0 when the inverse does not exist. Since  $\alpha(0) = 0$  in model (3), we suggest the following recursive estimating equation for estimation of  $\alpha_0(t)$  (given  $\beta$ )

$$\hat{\alpha}(t) = \int_0^t \{X^T(s)X(s)\}^{-1} X^T(s) \left\{ \exp(-Z\beta) ds - d\tilde{N}(s)X(s)\alpha(s^-) \right\}, \tag{9}$$

where  $\hat{\alpha}_0(s^-)$  is the left hand limit of  $\hat{\alpha}(t)$  at  $s$ . The equation displayed in (9) has a recursive structure, but admits an explicit solution. The Volterra equation [1, pp. 90-91] can easily be adapted to obtain the solution for equation (9) as

$$\hat{\alpha}(t; \beta) = Q(t)^{-1} \int_0^t Q(s^-) \{X^T(s)X(s)\}^{-1} X^T(s) \exp(-Z\beta) ds, \tag{10}$$

where  $Q(t) = \prod_{[t,\tau]} [I - \{X^T(s)X(s)\}^{-1} X^T(s)d\tilde{N}(s)X(s)]$ , with  $\prod_{[t,\tau]}$  denoting a product-integral, see [1, Scet. II.6] and  $I$  is the identity matrix.

To estimate  $\beta_0$ , we replace  $\alpha(t)$  with  $\hat{\alpha}(t; \beta)$  in (7) and solve that for  $\beta$ . Hence, using expression (8), the score equation for  $\beta_0$  can be simplified as

$$U\{\tau, \beta; \hat{\alpha}(t; \beta)\} = \int_0^\tau \{Z^T(t) - H(t)X^T(t)\} \left\{ d\tilde{N}(t)X(t)\hat{\alpha}(t; \beta) - \exp(-Z\beta)dt \right\}, \quad (11)$$

where  $Z(t) = \{Y_1(t)Z_1, \dots, Y_n(t)Z_n\}^T$  is a design matrix of dimension  $n \times q$  and  $H(t) = Z^T(t)X(t)\{X^T(t)X(t)\}^{-1}$ . Let  $\hat{\beta}$  denote the solution to  $U\{\tau, \beta; \hat{\alpha}(t; \beta)\} = 0$ . The corresponding estimator of  $\alpha_0(t)$  given  $\hat{\beta}$  is  $\hat{\alpha}(t) = \hat{\alpha}(t; \hat{\beta})$ .

### 3 Asymptotic properties

In this section, we establish the asymptotic properties of the estimators given in the previous section. First we consider the existence, uniqueness and strong consistency of  $\hat{\beta}$  together with the strong consistency of  $\hat{\alpha}(t)$  and then we drive asymptotic normality results. To study the asymptotic properties of the proposed estimators, we need the following regularity conditions:

- (C1)  $P(C \geq \tau) > 0$ , and  $N(\tau)$  is bounded almost surely.
- (C2) The covariates  $X$  and  $Z$  are bounded.
- (C3)  $\alpha_0(t)$  is right continuous with left-hand limits, and has bounded total variation on  $[0, \tau]$ .

The first step is to show that  $\hat{\alpha}(t; \beta)$  for given  $\beta$  converges uniformly in probability to a limit, say  $\alpha(t)$ . Using expression (9) and Volterra equation [1, pp. 90-91], it can be written that

$$\hat{\alpha}(t; \beta) - \alpha(t) = -Q(t)^{-1} \int_0^t Q(s^-) \{X^T(s)X(s)\}^{-1} X^T(s)d\tilde{M}(s)X(s)\alpha(s). \quad (12)$$

where  $d\tilde{M}(t) = \text{diag}\{dM_i(t)\}$ . Hence according to the standard martingale theory for counting process, (12) implies that  $\hat{\alpha}(t; \beta_0)$  converges almost surely to  $\alpha(t; \beta_0) = \alpha_0(t)$  in  $t \in [0, \tau]$ .

The next step is to show that the proposed estimator  $\hat{\beta}$  is unique and consistent. Denote minus the derivative of  $U\{\tau, \beta; \hat{\alpha}(t; \beta)\}$  with respect to  $\beta$  by  $I(\tau, \beta)$ , where it can be obtained that

$$I(\tau, \beta) = - \int_0^\tau \{Z^T(t) - H(t)X^T(t) - G(t)X^T(t)\} \text{diag}\{\exp(-Z^T\beta)\} Z(t)dt,$$

where

$$G(t) = \left[ \int_0^t \{Z^T(s) - H(s)X^T(s)\} d\tilde{N}(s)X(s)Q(s)^{-1} \right] Q(t^-) \{X^T(t)X(t)\}^{-1}.$$

By the uniform strong law of large numbers [12, p. 41], it follows that  $n^{-1}I(\tau, \beta)$  converges almost surely to a non-random function  $D(\tau, \beta)$  uniformly in  $\beta$  where

$$D(\tau, \beta) = -E \left[ \int_0^\tau \{Z_i(t) - h(t)X_i(t) - g(t)X_i(t)\} \exp(-Z_i^T \beta) Z_i^T(t) dt \right],$$

with  $h(t)$  and  $g(t)$  are the limit in probability of  $H(t)$  and  $G(t)$ , respectively. Assuming  $D(\tau, \beta)$  is non-singular, it follows from the inverse function theorem ([14], p. 221) that within a small neighborhood of  $\beta_0$ , there exists a unique solution  $\hat{\beta}$  to  $U(\tau, \beta) = 0$  for all sufficiently large  $n$ . Since this neighborhood of  $\beta_0$  can be arbitrarily small, the preceding proof also implies that  $\hat{\beta}$  is strongly consistent. It then follows from the uniform convergence of  $\hat{\alpha}(t; \beta_0)$  to  $\alpha(t; \beta_0)$  that  $\hat{\alpha}(t) \equiv \hat{\alpha}(t; \hat{\beta}) \rightarrow \alpha(t; \beta_0) \equiv \alpha_0(t)$  almost surely uniformly in  $t \in [0, \tau]$ .

An algebraic manipulation of (11) yields

$$n^{-1/2}U\{\tau, \beta_0; \hat{\alpha}(t; \beta_0)\} = n^{-1/2} \sum_{i=1}^n \int_0^\tau \{Z_i(t) - H(t)X_i(t) - G(t)X_i(t)\} X_i^T(t) \alpha_0(t) dM_i(t).$$

Hence by the multivariate central limit theorem,  $n^{-1/2}U\{\tau, \beta_0; \hat{\alpha}(t; \beta_0)\}$  converges in distribution to zero-mean normal distribution whose variance-covariance matrix  $\Sigma = E(\xi_i^{\otimes 2})$  can be consistently estimated by  $\hat{\Sigma} = n^{-1} \sum_{i=1}^n \hat{\xi}_i^{\otimes 2}$ , defined as

$$\hat{\xi}_i = \int_0^\tau \{Z_i(t) - H(t)X_i(t) - G(t)X_i(t)\} X_i^T(t) \hat{\alpha}(t) d\hat{M}_i(t),$$

where  $d\hat{M}_i(t)$  is the corresponding estimator  $dM_i(t)$  with all unknown parameters replaced by their estimates and  $v^{\otimes 2} = vv^T$  for a vector  $v$ . Furthermore, a straightforward Taylor expansion of score function (11) around  $\hat{\beta}$  gives

$$n^{1/2}(\hat{\beta} - \beta_0) = \{n^{-1}I(\tau, \beta_*)\}^{-1} n^{-1/2}U\{\tau, \beta_0; \hat{\alpha}(t; \beta_0)\},$$

where  $\beta_*$  is on the line segment between  $\beta_0$  and  $\hat{\beta}$ . Thus, it follows that  $n^{1/2}(\hat{\beta} - \beta_0)$  is asymptotically normal with mean zero and a variance-covariance matrix  $D(\tau, \beta_0)^{-1} \Sigma D(\tau, \beta_0)^{-1}$  that can be consistently estimated by  $\hat{D}(\tau, \hat{\beta})^{-1} \hat{\Sigma} \hat{D}(\tau, \hat{\beta})^{-1}$ , where

$$\hat{D}(\tau, \hat{\beta}) = -n^{-1} \sum_{i=1}^n \int_0^\tau \{Z_i(t) - H(t)X_i(t) - G(t)X_i(t)\} \exp(-Z_i^T \hat{\beta}) Z_i^T(t) dt.$$

## References

- [1] Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N. (2012), *Statistical models based on counting processes*, Springer Science and Business Media.
- [2] Asadi, M. (2006), On the mean past lifetime of the components of a parallel system, *Journal of Statistical Planning and Inference*, **136**(4), 1197-1206.

- [3] Asadi, M. and Berred, A. (2012), Properties and estimation of the mean past lifetime, *Statistics*, **46**(3), 405-417.
- [4] Bilius, Y., Gu, M. and Ying, Z. (1997), Towards a general asymptotic theory for Cox model with staggered entry, *The Annals of Statistics*, **25**(2), 662-682.
- [5] Cox, D. (1972), Regression models and life tables (with discussion), *Journal of the Royal Statistical Society*, **34**, 187-220.
- [6] Eryilmaz, S. (2010), Mean residual and mean past lifetime of multi-state systems with identical components, *IEEE Transactions on Reliability*, **59**(4), 644-649.
- [7] Finkelstein, M. S. (2002), On the reversed hazard rate, *Reliability Engineering and System Safety*, **78**(1), 71-75.
- [8] Kayid, M. and Ahmad, I. A. (2004), On the mean inactivity time ordering with reliability applications, *Probability in the Engineering and Informational Sciences*, **18**(3), 395-409.
- [9] Lin, D. Y. (2000), On fitting Cox's proportional hazards models to survey data, *Biometrika*, **87**(1), 37-47.
- [10] Nanda, A. K., Bhattacharjee, S. and Alam, S. S. (2006), On upshifted reversed mean residual life order, *Communications in Statistics Theory and Methods*, **35**(8), 1513-1523.
- [11] Nanda, A. K., Singh, H., Misra, N. and Paul, P. (2003), Reliability properties of reversed residual lifetime, *Communications in Statistics-Theory and Methods*, **32**(10), 2031-2042.
- [12] Pollard, D. (1990), *Empirical processes: theory and applications*, Institute of Mathematical Statistics and the American Statistical Association.
- [13] Rezaei, M. (2016), On proportional mean past lifetimes model, *Communications in Statistics-Theory and Methods*, **45**(14), 4035-4047.
- [14] Rudin, W. (1964), *Principles of mathematical analysis*, Vol. 3, New York: McGraw-Hill.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



# Parametric inference from system lifetime data under a proportional hazard rate model based on the minimum Jensen-Gini estimator

Mehrali, Y.<sup>1</sup>

Department of Statistics, University of Isfahan

## Abstract

In this paper, we discuss the statistical inference of the lifetime distribution of components based on observing the system lifetimes when the system structure is known. A general proportional hazard rate model for the lifetime of the components is considered, which includes some commonly used lifetime distributions. the minimum Jensen-Gini estimator, that recently presented by [7], for the proportionality parameter is discussed. We focus on a special case when the lifetime distributions of the components are exponential. Computational formulas for point and interval estimations of the unknown mean lifetime of the components are provided. A Monte Carlo simulation study is used to compare the performance of our method with some other estimation methods that recently considered by [8].

**Keywords:** Coherent systems, Exponential distribution, Order statistics.

## 1 Introduction

Parameter estimation is one of most important parts of statistical inference. There are various methods of estimation that presented with statisticians. We refer to [5], [9] and references therein. In recent years some authors proposed some parameter estimators which investigate by minimizing different divergence measures between the population density and kernel estimators of the density (see, for example, [1], [6] and [2]).

---

<sup>1</sup>yasermehrali@gmail.com

Consider the Gini's mean difference index of a random variable  $X$  with the cumulative distribution function (*c.d.f.*)  $F$  as

$$GMD(F) = \int_{-\infty}^{\infty} F(x)\bar{F}(x)dx = \frac{E(|X_1 - X_2|)}{2},$$

where  $X_1$  and  $X_2$  are *i.i.d.* with  $X$ . The Gini index, that introduced by Gini in 1914, is an index of statistical dispersion to measure the income distribution of individuals in a population (see, also [4]).

Recently, [7] presented an estimator of parameters of distribution based on their proposed measure which called the Jensen-Gini (*JG*) measure of divergence. We have the following definition of the *JG* from them.

**Definition 1.1.** Assume that  $\alpha_1, \alpha_2 > 0$ ,  $\alpha_1 + \alpha_2 = 1$ , are weights of two populations with *c.d.f.*'s  $F_1$  and  $F_2$ , respectively. Let also  $F = \alpha_1 F_1 + \alpha_2 F_2$  be a mixture of  $F_1$  and  $F_2$ . We define the *JG* measure of divergence as

$$JG(F_1, F_2) = GMD(F) - \alpha_1 GMD(F_1) - \alpha_2 GMD(F_2).$$

Let  $X$  is a random variable with *c.d.f.*  $F(x; \boldsymbol{\theta})$ , where  $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^k$  is unknown parameter. Let also  $F_n$  is the empirical *c.d.f.* based on  $n$  *i.i.d.* observations from  $X$ . The cited authors estimated  $\boldsymbol{\theta}$  based on  $JG(F(x; \boldsymbol{\theta}), F_n(x))$ . Considering the part of  $JG(F(x; \boldsymbol{\theta}), F_n(x)) / \alpha_1 \alpha_2$  that depends on  $\boldsymbol{\theta}$  as  $g_n(\boldsymbol{\theta})$ , they defined the minimum Jensen-Gini estimator (*MJGE*) of  $\boldsymbol{\theta}$  as

$$\hat{\boldsymbol{\theta}}_n = \arg \min_{\boldsymbol{\theta} \in \Theta} g_n(\boldsymbol{\theta}).$$

where

$$\begin{aligned} g_n(\boldsymbol{\theta}) &= \int_{-\infty}^0 F^2(x; \boldsymbol{\theta}) dx + \int_0^{\infty} \bar{F}^2(x; \boldsymbol{\theta}) dx - \overline{s(x; \boldsymbol{\theta})} \\ &= E_{\boldsymbol{\theta}} |X| - GMD_{\boldsymbol{\theta}}(X) - \overline{s(x; \boldsymbol{\theta})}, \end{aligned} \quad (1)$$

with

$$\begin{aligned} s(x; \boldsymbol{\theta}) &= I_{(-\infty, 0)}(x) u(x; \boldsymbol{\theta}) + I_{[0, \infty)}(x) h(x; \boldsymbol{\theta}) \\ u(x; \boldsymbol{\theta}) &= 2 \int_x^0 F(y; \boldsymbol{\theta}) dy, \\ h(x; \boldsymbol{\theta}) &= 2 \int_0^x \bar{F}(y; \boldsymbol{\theta}) dy, \end{aligned}$$

and

$$\overline{s(x; \boldsymbol{\theta})} = \frac{1}{n} \sum_{i=1}^n s(x_i; \boldsymbol{\theta}).$$

They employed this estimator in generalized Pareto distribution (*GPD*) and found that in *GPD*, for some values of parameter, the *MLE* does not exist while the *MJGE* exists and has a good performance relative to some other existing estimators. They also

studied properties of their estimator and proved that their estimator is in the class of the generalized estimating equations (*GEE*) estimators; see, for example [9]. Using this fact, they showed asymptotic normality of *MJGE* that in general case, when  $\boldsymbol{\theta}$  is vector, presents as the following Lemma.

Under the conditions of Theorem 5.14 of [9],

$$V_n^{-1/2} (\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{d} N_k(0, I_k),$$

where  $V_n = \frac{1}{n} B^{-1} A B^{-1}$ , provided that  $B$  is invertible matrix, where

$$\begin{aligned} A &= E \left\{ \boldsymbol{\psi}(X, \boldsymbol{\theta}) [\boldsymbol{\psi}(X, \boldsymbol{\theta})]^T \right\} \\ &= \left[ \frac{\partial}{\partial \boldsymbol{\theta}} s(X; \boldsymbol{\theta}) \right] \left[ \frac{\partial}{\partial \boldsymbol{\theta}} s(X; \boldsymbol{\theta}) \right]^T \\ &\quad - \left[ \frac{\partial}{\partial \boldsymbol{\theta}} E_{\boldsymbol{\theta}} |X| - \frac{\partial}{\partial \boldsymbol{\theta}} GMD_{\boldsymbol{\theta}}(X) \right] \left[ \frac{\partial}{\partial \boldsymbol{\theta}} E_{\boldsymbol{\theta}} |X| - \frac{\partial}{\partial \boldsymbol{\theta}} GMD_{\boldsymbol{\theta}}(X) \right]^T, \\ B &= \boldsymbol{\Psi}'(\boldsymbol{\theta}) = 2 \int_{-\infty}^{\infty} \left[ \frac{\partial}{\partial \boldsymbol{\theta}} F(x; \boldsymbol{\theta}) \right] \left[ \frac{\partial}{\partial \boldsymbol{\theta}} F(x; \boldsymbol{\theta}) \right]^T dx, \\ \boldsymbol{\psi}(x, \boldsymbol{\theta}) &= \frac{\partial}{\partial \boldsymbol{\theta}} E_{\boldsymbol{\theta}} |X| - \frac{\partial}{\partial \boldsymbol{\theta}} GMD_{\boldsymbol{\theta}}(X) - \frac{\partial}{\partial \boldsymbol{\theta}} s(x; \boldsymbol{\theta}), \end{aligned}$$

and  $\boldsymbol{\Psi}(\boldsymbol{\theta}) = E[\boldsymbol{\psi}(X, \boldsymbol{\theta})]$ .

The cited authors also expressed the sample version of the asymptotic variance of the *MJGE* as follows. Let  $x_1, \dots, x_n$  be observed values of the sample and

$$J = \left\{ \overline{\frac{\partial}{\partial \boldsymbol{\theta}} s(x; \boldsymbol{\theta})} \right\} \left\{ \overline{\frac{\partial}{\partial \boldsymbol{\theta}} s(x; \boldsymbol{\theta})} \right\}^T \Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_n} - \left\{ \overline{\frac{\partial}{\partial \boldsymbol{\theta}} s(x; \boldsymbol{\theta})} \right\} \left\{ \overline{\frac{\partial}{\partial \boldsymbol{\theta}} s(x; \boldsymbol{\theta})} \right\}^T \Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_n},$$

and

$$I = - \frac{\partial^2}{\partial \boldsymbol{\theta}^2} E_{\boldsymbol{\theta}} |X| \Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_n} + \frac{\partial^2}{\partial \boldsymbol{\theta}^2} GMD_{\boldsymbol{\theta}}(X) \Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_n} + \overline{\frac{\partial^2}{\partial \boldsymbol{\theta}^2} s(x; \boldsymbol{\theta})} \Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_n}.$$

Then,

$$\widehat{V}_n^{-1/2} (\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{d} N_k(0, I_k), \quad (2)$$

where

$$\widehat{V}_n = \frac{1}{n} I^{-1} J I^{-1},$$

provided that  $I$  is invertible matrix, or equivalently  $g_n(\boldsymbol{\theta})$  has its infimum value on  $\Theta$ . For special case that  $X$  is nonnegative,  $J$  and  $I$  will be replaced, respectively, with

$$J = \left\{ \overline{\frac{\partial}{\partial \boldsymbol{\theta}} h(x; \boldsymbol{\theta})} \right\} \left\{ \overline{\frac{\partial}{\partial \boldsymbol{\theta}} h(x; \boldsymbol{\theta})} \right\}^T \Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_n} - \left\{ \overline{\frac{\partial}{\partial \boldsymbol{\theta}} h(x; \boldsymbol{\theta})} \right\} \left\{ \overline{\frac{\partial}{\partial \boldsymbol{\theta}} h(x; \boldsymbol{\theta})} \right\}^T \Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}_n},$$

and

$$I = - \frac{\partial^2}{\partial \boldsymbol{\theta}^2} E_{\boldsymbol{\theta}}(X) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_n} + \frac{\partial^2}{\partial \boldsymbol{\theta}^2} GMD_{\boldsymbol{\theta}}(X) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_n} + \overline{\frac{\partial^2}{\partial \boldsymbol{\theta}^2} h(x; \boldsymbol{\theta})} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_n}.$$

The aim of this paper is to study the parametric statistical estimation based on the *MJGE*, from system lifetime data under a proportional hazard rate model. The rest of this paper is organized as follows: In Section 2 we present the model and parametric statistical estimation based on *MJGE*. In Section 3 we perform some simulation studies and compare our estimator with other estimators.

## 2 Model and parametric statistical estimation based on *MJGE*

Let  $\mathbf{p} = (p_1, p_2, \dots, p_n)$  be the signature of a coherent system with  $n$  independent components. we have the survival function (*SF*) of the system lifetime  $T$  as

$$\bar{F}_T(t) = \sum_{i=1}^n p_i \bar{F}_{i:n}(t).$$

On the other hand, using minimal signature of system  $\mathbf{a} = (a_1, a_2, \dots, a_n)$ , the *SF* of the system lifetime  $T$  can be expressed as

$$\bar{F}_T(t) = \sum_{i=1}^n a_i \bar{F}_{1:i}(t).$$

We consider the proportional hazard rate (*PHR*) model for the common distribution of the *IID* lifetimes of the components as

$$\bar{F}_X(t) = [\bar{G}(t)]^\alpha.$$

Based on the above model we have

$$\bar{F}_{1:i}(t) = \bar{G}^{i\alpha}(t),$$

and

$$\bar{F}_T(t) = \sum_{i=1}^n a_i \bar{G}^{i\alpha}(t),$$

Now, based on the  $m$  observed lifetimes  $t_1, t_2, \dots, t_m$  of  $n$ -component systems with signature  $\mathbf{p} = (p_1, p_2, \dots, p_n)$  (minimal signature  $\mathbf{a} = (a_1, a_2, \dots, a_n)$ ), using (1), we can evaluate the function  $g_m(\alpha)$  as

$$g_m(\alpha) = \int_0^\infty \bar{F}_T^2(t) dt - \bar{h}(t; \alpha),$$

In order to find an explicit form for  $g_m(\alpha)$ , using some algebra we have

$$\begin{aligned}\int_0^\infty \bar{F}_T^2(t) dt &= \int_0^\infty \left( \sum_{i=1}^n a_i \bar{G}^{i\alpha}(t) \right)^2 dt \\ &= \int_0^\infty \sum_{i=1}^n \sum_{j=1}^n a_i a_j \bar{G}^{(i+j)\alpha}(t) dt \\ &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j \int_0^\infty \bar{G}^{(i+j)\alpha}(t) dt,\end{aligned}$$

and

$$\begin{aligned}h(t; \alpha) &= 2 \int_0^t \bar{F}_T(y) dy \\ &= 2 \int_0^t \sum_{i=1}^n a_i \bar{G}^{i\alpha}(y) dy \\ &= 2 \sum_{i=1}^n a_i \int_0^t \bar{G}^{i\alpha}(y) dy.\end{aligned}$$

So we get

$$g_m(\alpha) = \sum_{i=1}^n \sum_{j=1}^n a_i a_j \int_0^\infty \bar{G}^{(i+j)\alpha}(t) dt - \frac{2}{m} \sum_{k=1}^m \sum_{i=1}^n a_i \int_0^{t_k} \bar{G}^{i\alpha}(y) dy,$$

and the *MJGE* of  $\alpha$  is defined as

$$\hat{\alpha}_n = \arg \min_{\alpha \in \mathbb{R}^+} g_m(\alpha).$$

Now, using (2), the asymptotic distribution of  $\hat{\alpha}_n$  is given by

$$\hat{V}_m^{-1/2} (\hat{\alpha}_n - \alpha) \xrightarrow{d} N(0, 1), \quad (3)$$

where

$$\hat{V}_m = \frac{1}{m} \frac{J}{I^2},$$

with

$$J = \left\{ \overline{\left\{ \frac{\partial}{\partial \alpha} h(t; \alpha) \right\}}^2 \right|_{\alpha=\hat{\alpha}_m} - \left\{ \overline{\left\{ \frac{\partial}{\partial \alpha} h(t; \alpha) \right\}}^2 \right|_{\alpha=\hat{\alpha}_m},$$

and

$$I = - \frac{\partial^2}{\partial \alpha^2} \int_0^\infty \bar{F}_T^2(t) dt \Big|_{\alpha=\hat{\alpha}_m} + \overline{\frac{\partial^2}{\partial \alpha^2} h(t; \alpha)} \Big|_{\alpha=\hat{\alpha}_m}.$$

An asymptotic  $100(1 - \gamma)\%$  confidence interval for  $\alpha$  (namely *AMJGE*) is then

$$\hat{\alpha}_m \pm z_{1-\gamma/2} \sqrt{\hat{V}_m},$$

where  $z_q$  is the  $q$ -th upper percentile of the standard normal distribution.

## 2.1 Systems with exponentially distributed components

Consider that the lifetime of the  $n$  components in a system are *IID* exponentially distributed with *PDF* and *CDF*

$$f_X(x) = \frac{1}{\theta} \exp\left(-\frac{x}{\theta}\right) \text{ and } F_X(x) = 1 - \exp\left(-\frac{x}{\theta}\right), \quad x > 0, \quad \theta > 0,$$

respectively. This model is equivalent to setting  $\bar{G}(t) = e^{-t}$  and  $\alpha = 1/\theta$ . After some algebra we have

$$\begin{aligned} \int_0^\infty \bar{F}_T^2(t) dt &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j \int_0^\infty e^{-(i+j)\alpha t} dt \\ &= \sum_{i=1}^n \sum_{j=1}^n \frac{a_i a_j}{(i+j)\alpha} \int_0^\infty (i+j)\alpha e^{-(i+j)\alpha t} dt \\ &= \frac{1}{\alpha} \sum_{i=1}^n \sum_{j=1}^n \frac{a_i a_j}{i+j}, \\ h(t; \alpha) &= 2 \sum_{i=1}^n a_i \int_0^t e^{-i\alpha y} dy \\ &= 2 \sum_{i=1}^n \frac{a_i}{i\alpha} \int_0^t i\alpha e^{-i\alpha y} dy \\ &= \frac{2}{\alpha} \sum_{i=1}^n \frac{a_i}{i} (1 - e^{-i\alpha t}), \end{aligned}$$

So, we get

$$g_m(\alpha) = \frac{1}{\alpha} \sum_{i=1}^n \sum_{j=1}^n \frac{a_i a_j}{i+j} - \frac{2}{\alpha m} \sum_{k=1}^m \sum_{i=1}^n \frac{a_i}{i} (1 - e^{-i\alpha t}).$$

After minimizing  $g_m(\alpha)$  with respect to  $\alpha$ , we can evaluate the *MJGE* of  $\alpha$  numerically. The asymptotic distribution of  $\hat{\alpha}_n$  is given as (2) where

$$\begin{aligned} \frac{\partial}{\partial \alpha} \int_0^\infty \bar{F}_T^2(t) dt &= -\frac{1}{\alpha^2} \sum_{i=1}^n \sum_{j=1}^n \frac{a_i a_j}{i+j}, \\ \frac{\partial^2}{\partial \alpha^2} \int_0^\infty \bar{F}_T^2(t) dt &= \frac{2}{\alpha^3} \sum_{i=1}^n \sum_{j=1}^n \frac{a_i a_j}{i+j}, \\ \frac{\partial}{\partial \alpha} h(t; \alpha) &= -\frac{2}{\alpha^2} \sum_{i=1}^n \frac{a_i}{i} (1 - e^{-i\alpha t}) + \frac{2t}{\alpha} \sum_{i=1}^n a_i e^{-i\alpha t}, \\ \frac{\partial^2}{\partial \alpha^2} h(t; \alpha) &= \frac{4}{\alpha^3} \sum_{i=1}^n \frac{a_i}{i} (1 - e^{-i\alpha t}) - \frac{4t}{\alpha^2} \sum_{i=1}^n a_i e^{-i\alpha t} - \frac{2t^2}{\alpha} \sum_{i=1}^n i a_i e^{-i\alpha t}. \end{aligned}$$

In the next section we perform a simulation study.

### 3 Simulation study

Table 1: Simulated biases of method of moments estimators (*MOM*), maximum likelihood estimators (*MLE*), least squares estimators (*LSE*), weighted least squares estimators (*WLSE*) and minimum Jensen-Gini estimators (*MJGE*) of  $\alpha$  for different 4-component systems with  $m = 5, 10, 15$  and  $25$

m	System No.	MOM	MLE	LSE	WLS	MJGE
5	1	0.280	0.280	0.342	0.242	0.254
	2	0.198	0.214	0.240	0.245	0.189
	3	0.188	0.141	0.116	0.120	0.142
	4	0.163	0.159	0.213	0.161	0.145
	5	0.119	0.104	0.076	0.094	0.095
	6	0.082	0.077	0.066	0.105	0.072
10	1	0.098	0.098	0.078	0.070	0.088
	2	0.088	0.096	0.073	0.087	0.082
	3	0.086	0.066	0.048	0.054	0.063
	4	0.093	0.092	0.076	0.083	0.089
	5	0.060	0.052	0.033	0.051	0.046
	6	0.042	0.039	0.027	0.050	0.036
20	1	0.049	0.049	0.042	0.036	0.045
	2	0.042	0.047	0.028	0.034	0.039
	3	0.026	0.017	0.012	0.014	0.019
	4	0.036	0.035	0.022	0.026	0.032
	5	0.029	0.025	0.014	0.022	0.023
	6	0.014	0.013	0.005	0.016	0.011
25	1	0.044	0.044	0.030	0.028	0.038
	2	0.021	0.023	0.018	0.022	0.022
	3	0.024	0.014	0.006	0.008	0.015
	4	0.029	0.028	0.021	0.024	0.026
	5	0.020	0.018	0.014	0.021	0.017
	6	0.024	0.023	0.019	0.028	0.022

In this section, we present a Monte Carlo simulation study to compare the performance of our method with some other estimation methods that recently considered by [8]. These methods are method of moments (*MOM*) estimation , maximum likelihood estimation (*MLE*), least squares estimation (*LSE*) and weighted least squares (*WLS*) estimation. For 6 different 4-component systems that mentioned by [8] with  $m = 5, 10, 15$  and  $25$ , we generated 1000 sets of system lifetimes from proportional components with  $\alpha = 1$ . Using these simulated system lifetimes, we estimated parameter  $\alpha$ . Tables 1 and 2 represent simulated biases and mean squares errors (*MSEs*) of these estimators along with our estimator (*MJGE*). It can be seen from table 1 that in term of lesser bias, between all 24 cases, the *MJGE* is in 2 cases (8.3%) in order 1, in 8 cases (33.3%) in order 2, in 10 cases (41.6%) in order 3, and in 4 cases (16.6%) in order 4. But in no cases *MJGE* has most bias. On the other hands, it can be seen from table 2 that in term of lesser *MSE*,

Table 2: Simulated mean squares errors ( $MSE$ 's) of method of moments estimators ( $MOM$ ), maximum likelihood estimators ( $MLE$ ), least squares estimators ( $LSE$ ), weighted least squares estimators ( $WLSE$ ) and minimum Jensen-Gini estimators ( $MJGE$ ) of  $\alpha$  for different 4-component systems with  $m = 5, 10, 15$  and  $25$

m	System No.	MOM	MLE	LSE	WLS	MJGE
5	1	0.593	0.593	2.598	1.715	0.612
	2	0.387	0.398	0.913	0.791	0.391
	3	0.341	0.289	0.312	0.308	0.291
	4	0.333	0.324	2.064	0.583	0.337
	5	0.190	0.181	0.225	0.216	0.182
	6	0.108	0.106	0.147	0.161	0.109
10	1	0.146	0.146	0.191	0.177	0.154
	2	0.111	0.112	0.155	0.170	0.112
	3	0.111	0.099	0.112	0.112	0.102
	4	0.117	0.117	0.141	0.144	0.122
	5	0.074	0.072	0.078	0.085	0.071
	6	0.041	0.040	0.047	0.060	0.041
20	1	0.065	0.065	0.083	0.075	0.068
	2	0.047	0.046	0.060	0.062	0.050
	3	0.041	0.039	0.046	0.044	0.041
	4	0.048	0.048	0.058	0.057	0.050
	5	0.030	0.029	0.034	0.035	0.030
	6	0.018	0.018	0.020	0.021	0.018
25	1	0.051	0.051	0.065	0.059	0.053
	2	0.032	0.031	0.043	0.043	0.034
	3	0.038	0.034	0.036	0.035	0.035
	4	0.034	0.034	0.044	0.043	0.036
	5	0.024	0.024	0.030	0.031	0.025
	6	0.016	0.015	0.017	0.018	0.016

the  $MJGE$  is in 1 cases (4.16%) in order 1, in 6 cases (25%) in order 2, and in 17 cases (70.83%) in order 3. But in no cases  $MJGE$  is in order 4 or 5 in term  $MSE$ . So, the  $MJGE$  can be considered as an estimator beside other traditional estimators.

## References

- [1] Basu, A., & Lindsay, B. G. (1994). Minimum disparity estimation for continuous models: efficiency, distributions and robustness. *Annals of the Institute of Statistical Mathematics*, **46**(4), 683-705.
- [2] Basu, A., Shioya, H., & Park, C. (2011). *Statistical inference: the minimum distance approach*. CRC Press.

- [3] Gini, C. (1914). Sulla misura della concentrazione e della variabilità dei caratteri. *Atti del R . Istituto Veneto di SS.LL.AA ., a .a. 1913-14, tomo LXXIII, parte II*, 1203-1248.
- [4] Giorgi, G. M. (1993). A fresh look at the topical interest of the Gini concentration ratio. *Metron*, **51(1-2)**, 83-98.
- [5] Lehmann, E. L., & Casella, G. (1998). *Theory of Point Estimation* (Springer Texts in Statistics) (2 ed.).
- [6] Lindsay, B. G. (1994). Efficiency versus robustness: the case for minimum Hellinger distance and related methods. *The Annals of Statistics*, **22(2)**, 1081-1114.
- [7] Mehrali, Y., Asadi, M., & Kharazmi, O. (2017). A Jensen-Gini Measure of Divergence with Application in Parameter Estimation. *Metron*, Accepted.
- [8] Ng, H. K. T., Navarro, J., & Balakrishnan, N. (2012). Parametric inference from system lifetime data under a proportional hazard rate model. *Metrika*, **75(3)**, 367-388.
- [9] Shao, J. (2003). *Mathematical Statistics (2 ed.)*. Springer, New York, USA.



*Forth Seminar on*  
**Reliability Theory and its Applications**

*April 25-26, 2018*



## Survival analysis of dependent competing risks with masked failure causes

Mireh, S.<sup>1</sup>      Khodadadi, A.<sup>2</sup>      Haghghi, F.<sup>3</sup>

<sup>1,2</sup> Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran

<sup>3</sup> Department of Statistics, Faculty of Mathematics, Statistics and Computer Sciences, Tehran University

### Abstract

The competing risks model is useful in settings in which units may fail for different reasons. Traditionally, statistical inference is performed under the assumption that the failure causes act independently on each item. In this paper we propose a copula-based approach which allows for dependent competing risks. In addition to failure time data, we also use degradation data that has not been investigated in previous studies. A complication arises when some of the failures are masked within a group of possible causes. Following a failure, immediate procedures are used in an attempt to reach a definitive diagnosis (stage-1). The cause of failure may not have been identified but has only been narrowed down to a subset of all potential risks. Stage-2 procedures, such as failure analysis, provide definitive diagnosis for a sample of the masked cases. We show how stage-1 and stage-2 information can be combined to provide statistical inference about three things: survival functions, the proportions of failures associated with individual risks and probability that each of the masked competing risks is responsible for the failure. We suppose that the intensity function corresponding to each cause of failure be a known function.

**Keywords:** Dependent competing risks, Masked cause, Copula function.

<sup>1</sup>mireh@sbu.ac.ir

<sup>2</sup>khodadadi@sbu.ac.ir

<sup>3</sup>haghghi@khayam.ut.ac.ir

# 1 Introduction

In reliability engineering, the problem of estimating the reliability of system is very important. The failure time data is often used to estimate component reliability. Ideally, lifetime data consist of a failure time and the associated cause of failure. However, sometimes the exact cause of failure may not be known due to some unavoidable reasons such as lack of proper diagnostic equipment or cost and time constraints, etc. Such data are termed as masked data. As mentioned, in this paper we used degradation data in addition to failure time data.

In the analysis of the survival function, it is usually the case that the system is damaged by only one factor; failure for each item being due to only one failure cause. Crowder [3] studied a problem for which one needs to estimate the failure rates for each cause. Items that do not fail during the experiment have no failure time associated with them, but it is also possible that some of the items that fail during this period have a cause of failure that is only known to belong to a certain subset of all possible failures, in other words, their cause of failure is group masked.

Under the assumption of independency of components of a system, many authors studied the statistical analysis of masked data. Sarhan and El-Bassiouny [7] considered the reliability estimation of components in a parallel system using masked system life data. Fan and Hsu [4] discussed the statistical analysis of masked interval data with exponential distributed components in step-stress accelerated life test (SSALT). Xu et al. [9] developed a Bayesian procedure using Gibbs sampling for masked system life data in SSALT. Wang et al. [8] provided the maximum likelihood method for estimating the parameter in a hybrid system.

It is well known that the independence between the competing risks cannot be tested using the failure times and the failure causes even if these are completely known for all items in the study. Kalbfleisch and Prentice [6] emphasize that in many examples (e.g., life time data) dependence between the competing risks is reasonable although a parametric model for the dependency is hard to specify. In this paper, we show that copula function can be used for estimation in a general situation in which the competing risks do not act independently. We also used degradation data along with failure time data.

Literature about the statistical analysis of dependent component system is rare. Actually, a copula function provides a method to model multivariate survival data, and it has become interestingly popular. Xu and Tang [10] discussed the statistical analysis of competing failure modes in accelerated life testing based on the Gumbel copula. Zhang et al. [11] developed the statistical inference of accelerated life test data with dependent competing risks model based on copula theory. In our article, a copula function is used to describe the dependent lifetimes of components in a system, and the statistical analysis of masked data along with degradation data in such a system is discussed.

Frequently, engineering and medical considerations require that some subset of masked failures be subject to a second stage of study to arrive at a definitive resolution of the failure cause. This occurs in failure analysis laboratories. This stage-2 data, appropriately collected, can be used to improve survival estimation. This article focuses on parametric survival analysis in systems with dependent competing risks (given degradation) and masked failures. We show how stage-1 and stage-2 information can be combined to

produce statistical inference on:

- Survival (reliability) functions
- The proportion of failures associated with individual components
- Probability that each of the masked competing risks is responsible for the failure.

For the competing risks model with masked causes of failure, some authors have derived semiparametric and nonparametric inference procedures for the case with two failure causes and no second-stage analysis, which often occurs in carcinogenicity bioassays: Craiu and Duchesne [1, 2] use a semiparametric model with piecewise constant hazard functions which presents robust properties and can be adapted to most situations in which some second stage data is available. In the case of a general number of failure causes and availability of second-stage analysis data, Flehinger et al.[5] propose maximum likelihood estimation under a model with proportional cause-specific hazards and a model with completely parametric cause-specific hazards.

In this article, we mainly discuss the statistical analysis of systems with dependent competing risks along with degradation data and two stages masked data. The paper is organized as follows. In Section 2 we introduce the copula theory. In Section 3, we give the basic notation and assumptions, model description and maximum likelihood estimators for the parameters. Numerical simulation is presented in Section 4. Conclusion is discussed in Section 5.

## 2 Copula function

There is a need for more realistic modeling of a stochastic dependency structure that contain all information about dependent random variables, which goes beyond the measure of linear correlation coefficients and has the capability to capture co-dependency outside the world of elliptical distributions, particularly in situations where marginal are known and joint distributions may be unknown. Copula function allows one to model the dependence structure independently of marginal distributions. This approach provides a frequently more useful representation of multivariate distribution compared to traditional approaches such as multivariate normality. Note that an independent copula implies zero correlation but the opposite is false. Formally, copulas can be defined as follows. Suppose that we have two marginal CDFs,  $F_{X_1}(x_1)$ ,  $F_{X_2}(x_2)$  where  $X_1$ ,  $X_2$  are the random variables. Sklar's theorem states that every bivariate cumulative distribution function

$$H(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2),$$

of a random vector  $(X_1, X_2)$  can be expressed by involving only the marginals  $F_{X_1}(x_1)$  and  $F_{X_2}(x_2)$  as

$$H(x_1, x_2) = C(F_{X_1}(x_1), F_{X_2}(x_2)),$$

where  $C$  is a copula. Differentiating with respect to  $x_1, x_2$  leads to:

$$h(x_1, x_2) = \frac{\partial H(x_1, x_2)}{\partial x_1 \partial x_2} = f_{X_1}(x_1).f_{X_2}(x_2).c(F_{X_1}(x_1), F_{X_2}(x_2)), \quad (1)$$

where  $c(u_1, u_2) = \frac{\partial C(u_1, u_2)}{\partial u_1 \partial u_2}$  is the density of the copula  $C$ . The random variables  $X_1$  and  $X_2$  are said to be independent if:

$$C(F_{X_1}(x_1), F_{X_2}(x_2)) = F_{X_1}(x_1).F_{X_2}(x_2). \quad (2)$$

### 3 The proposed model

Suppose we have  $N$  systems that are subject to failure due to dependent competing causes (we consider only two cases.). Each system either fails or censored at a specified time  $c$ . Let  $T_j$  be the failure time of system  $j$ . Thus,  $T_j = \min\{T_j^{(1)}, T_j^{(2)}, c\}$  is the system failure time. We suppose that each system failure is due to only one of the Failure modes; however, our stage 1 test procedures may result in masking. A sample of cases corresponding to each masked subset is taken to stage 2 testing depend on the observed stage 1 data, but it does not depend on the model parameters. In the following, notation and assumptions are expressed and then the likelihood function is introduced.

#### 3.1 Notation and assumptions

In the following topics, the following notations are used:

$g$ : masked group

$n_i$ : times to failure identified as being caused by cause  $i$  have lifetimes  $t_j^{(i)}$

$n_g$ : times to failure restricted to masking group  $g$  have lifetimes  $t_j^{(g)}$

$n_c$ : the number of censored systems with censoring (survival) times  $t_j^{(c)}$

$i \subset g$ : cause  $i$  is contained in masked group  $g$

$n_{g,i}$ : the number of system failures restricted to masking group  $g$  in stage 1 and identified with cause  $i$  in stage 2

$n_j^+ = \sum_{i=1}^k n_{g,i}$ : the total number of system failures restricted to masking group  $g$  in stage 1 that are identified in stage 2

$\tilde{n}_j$ : the number of system failures restricted to masking group  $g$  in stage 1 that are not taken to stage 2 for further identification ( $n_j^+ + \tilde{n}_j = n_g$ )

$n_i^* = n_i + \sum_{g \supset i} n_{g,i}$ : the total number of system failures identified with cause  $i$  in either stage 1 or 2

$\bar{t}_j$ : the  $j$ -th event time that corresponds either to a failure or a censoring time

$P_i$ : the  $i$ th identification probability (the probability that a system failure induced by cause  $i$  is correctly identified in stage 1)

$P_{g|i}$ : the masking probability (the probability that a system failure induced by cause  $i$  is restricted to masking group  $g$  in stage 1)

Suppose that we have two dependent failure modes with intensity function  $h^{(i)}(z) = (\theta_i z)^{\nu_i}$  which depends on the amount of degradation ( $z$ ). Also assume that the degradation is linear,  $Z = t/A$ , and  $A$  has Weibull distribution with parameters  $(\alpha_0, \beta_0)$ . The dependence structure is generated by the Frank copula. Thus, the joint survival (reliability) distribution of  $(T^{(1)}, T^{(2)})$  is defined as follows:

$$S(t^{(1)}, t^{(2)}|\theta) = C(F_1(t^{(1)}), F_2(t^{(2)})|\theta) = -\frac{1}{\theta} \log \left( 1 + \frac{(e^{-\theta F_1(t^{(1)})} - 1)(e^{-\theta F_2(t^{(2)})} - 1)}{(e^{-\theta} - 1)} \right)$$

where  $\theta \in R \setminus \{0\}$  is the copula parameter that measures the strength of dependency.

Finally, suppose that  $\pi_{i|g}(t)$  be the probability that a failure at time  $t$  that is restricted to a masking group  $g$  is due to component  $i$ . According to Bayes theorem:

$$\pi_{i|g}(t) = \frac{P_{g|i}f^{(i)}(t)}{\sum_{r \subset g} P_{g|r}f^{(r)}(t)}, \quad i = 1, \dots, k,$$

where  $f^{(i)}(t) = \frac{\partial}{\partial t^{(i)}} C(F_1(t^{(1)}), F_2(t^{(2)}))|_{t^{(i)}=t}$  is 'crude density function' and  $\pi_{i|g}(t)$  is 'diagnostic probabilities'.

### 3.2 Likelihood function

To write the likelihood function, we must know that the data are divided into four categories: (1) failure modes identified in stage 1, (2) failure mode masked and never identified, (3) failure mode identified in stage 2 and (4) failures that have been censored. Therefore the likelihood is constructed by the multiplication of each of these categories. The likelihood function based on both stage-1 and stage-2 data and the above notation and assumptions can be seen as:

$$\begin{aligned} L = & \prod_{j=1}^{n_c} S_T(t_j^{(c)}, Z = Z_j^{(c)}) \prod_{i=1}^k \prod_{j=1}^{n_i} \left\{ P_i \cdot P(T^{(i)} = t_j^{(i)}, T^{(i' \neq i)} > t_j^{(i)}, Z = Z_j^{(i)}) \right\} \\ & \times \prod_g \prod_{j=1}^{\tilde{n}_g} \left[ \sum_{r \subset g} P_{g|r} \cdot P(T^{(r)} = \tilde{t}_j^{(g)}, T^{(l \neq r)} > \tilde{t}_j^{(g)}, Z = \tilde{Z}_j^{(g)}) \right] \\ & \times \prod_{i=1}^k \prod_{g \supset i} \prod_{j=1}^{n_{g,i}} \left[ P_{g|i} \cdot P(T^{(i)} = t_j^{(g,i)}, T^{(l \neq i)} > t_j^{(g,i)}, Z = Z_j^{(g,i)}) \right] \end{aligned}$$

After doing some algebraic operations, the likelihood function is written as follows:

$$\begin{aligned} L = & \prod_{j=1}^N S_T(\bar{t}_j, Z = Z_j^{(c)}) \prod_{i=1}^k \left( P_i^{n_i} \prod_{g \supset i} P_{g|i}^{n_{g,i}} \right) \times \prod_{i=1}^k \prod_{j=1}^{n_i^*} h^{(i)}(t_j^{(i*)}, Z = Z_j^{(i*)}) \\ & \times \prod_g \prod_{j=1}^{\tilde{n}_g} \sum_{r \subset g} \left[ P_{g|r} \cdot h^{(r)}(\tilde{t}_j^{(g)}, Z = \tilde{Z}_j^{(g)}) \right] \end{aligned} \quad (3)$$

where  $h^{(r)}(t)$  is the crude hazard function.

To estimate the parameters, we maximize the likelihood function (4), using the numerical method. When the distributional parameters  $(\alpha_0, \beta_0, \theta_1, \nu_1, \theta_2, \nu_2, \theta)$ , the identification probabilities  $P_1, P_2$ , and the masking probabilities  $P_{g|i}$ , for all possible  $i$  and  $g$ , can not be treated as known quantities, we are interested in maximizing (4) over all of them. An iterative approach to this problem will be taken. At first, we estimate the identification and masking probabilities for given values of  $(\alpha_0, \beta_0, \theta_1, \nu_1, \theta_2, \nu_2, \theta)$ . Then we discuss the estimation of  $(\alpha_0, \beta_0, \theta_1, \nu_1, \theta_2, \nu_2, \theta)$  given the identification and masking probabilities. We start with some initial values of the  $P_i$  and  $P_{g|i}$ . The iteration process stops when the estimation of the parameters in two successive steps are close together.

### 3.3 Estimation of $P_i$ and $P_{g|i}$

To determine the estimation of  $P_i$  and  $P_{g|i}$ , we set the derivatives of the Lagrangian,  $(\log L + \sum_{i=1}^k \lambda_i \times (1 - P_i - \sum_{g \supseteq i} P_{g|i}))$ , by  $P_i$  and  $P_{g|i}$  to zero, where  $\lambda_i = n_i^* + \sum_{g \supseteq i} \sum_{j=1}^{\tilde{n}_g} \pi_{i|g}(\tilde{t}_j^{(g)})$ . In summary, the estimation of  $P_i$  and  $P_{g|i}$  given  $\beta_i$  and  $\theta$  is as follows:

1. Set:  $P_i = P_{g|i} = 1/(1 + \sum_{g \supseteq i} 1)$
2. Compute the estimated values of  $\lambda_i$ :  $\lambda_i = n_i^* + \sum_{g \supseteq i} \sum_{j=1}^{\tilde{n}_g} \pi_{i|g}(\tilde{t}_j^{(g)})$
3. Update the identification and masking probabilities:

$$\begin{cases} P_i^{(new)} = n_i / \lambda_i \\ P_{g|i}^{(new)} = (1/\lambda_i)[n_{g,i} + \sum_{j=1}^{\tilde{n}_g} \pi_{i|g}(\tilde{t}_j^{(g)})] \end{cases}$$

and return to Step 1.

## 4 Simulation study

As mentioned before, we supposed that the intensity function of  $T^{(i)}$  is  $h^{(i)}(z) = (\theta_i z)^{\nu_i}$ ,  $i = 1, 2$ .  $A$  in  $Z = t/A$  also has a Weibull distribution with parameters  $(\alpha_0, \beta_0)$ . We simulated  $N = 50$  dependent data given  $\theta_1 = 0.06$ ,  $\nu_1 = 5$ ,  $\theta_2 = 0.06$ ,  $\nu_2 = 5$ ,  $\alpha_0 = 5$ ,  $\beta_0 = 4$  and  $\theta = 20$ . Data censored at  $c = 50$ .

To investigate the effect of masked data on parameter estimation and survival analysis, we simulated data with different number of masked data. Therefore, the four following items are compared: 1)  $P_{g|1} = 0.1, P_{g|2} = 0.2$ , 2)  $P_{g|1} = 0.3, P_{g|2} = 0.5$ , 3)  $P_{g|1} = 0.7, P_{g|2} = 0.3$ , 4)  $P_{g|1} = 0.8, P_{g|2} = 0.8$ . The estimation of parameters are given in Table 1 and the estimation of survival function are shown in Figure 1. As can be seen in Table 1,  $P_i + P_{g|i} = 1, i = 1, 2$ , and lower number of masked data lead to more precise parameter estimation.

## 5 Conclusion

In this paper we have emphasized the use of parametric modelling for the dependent competing risks with masked cases. For more precision, in addition to failure time data, degradation data is also used. Different values of  $P_{g|i}$  are considered to examine the effect of the number of masked data on parameter estimation. Then the estimation of the parameters and the survival function are obtained.

Table 1: The MLEs of the parameters for different number of masked data ( $\alpha_1 = 0.06$ ,  $\nu_1 = 5$ ,  $\alpha_2 = 0.06$ ,  $\nu_2 = 5$ ,  $\theta = 20$ )

n	Parameters	$(P_g 1, P_g 2)$			
		(0.1,0.2)	(0.3,0.5)	(0.7,0.3)	(0.8,0.8)
n=50	$\alpha_0$	5.285	5.285	5.285	5.285
	$\beta_0$	4.014	4.014	4.014	4.014
	$\theta_1$	0.0608	0.0612	0.0605	0.064
	$\nu_1$	5.254	5.318	5.263	6.609
	$\theta_2$	0.0615	0.0615	0.0619	0.057
	$\nu_2$	5.461	5.473	5.461	4.876
	$P_1$	0.825	0.679	0.264	0.244
	$P_2$	0.707	0.449	0.684	0.193
	$P_g 1$	0.175	0.321	0.736	0.756
	$P_g 2$	0.293	0.551	0.316	0.807
	$\theta$	20.083	19.457	19.350	17.494
	$\tau_\theta$	0.718	0.7173	0.715	0.6058

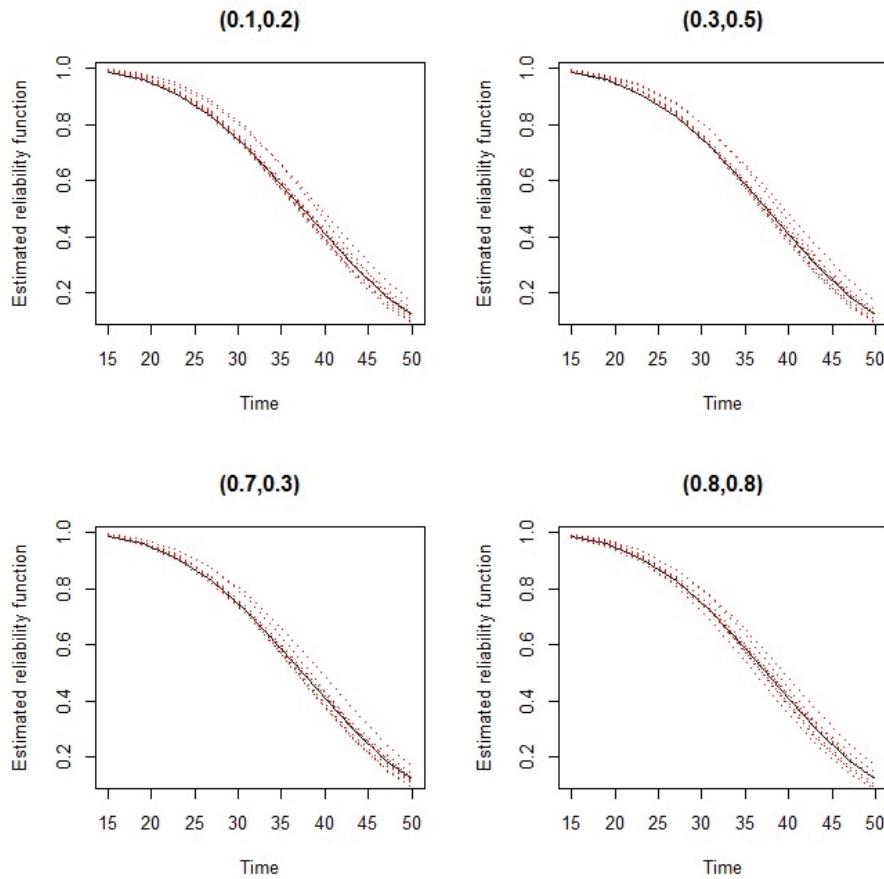


Figure 1: The estimation of survival function in different cases

## References

- [1] Craiu, R.V. and Duchesne, T. (2004a), Inference based on the EM algorithm for the competing risk model with masked causes of failure. *Biometrika*, **91**, 543-558.
- [2] Craiu, R.V. and Duchesne, T. (2004b), *Using EM and data augmentation for the competing risks model*. In: Gelman A, Meng XL (eds), Applied Bayesian modeling and causal inference from an incomplete-data perspective. Wiley
- [3] Crowder, M. (2001), *Classical competing risks*, Chapman & Hall.
- [4] Fan, T.H. and Hsu, T.M. (2012), Accelerated life tests of a series system with masked interval data under exponential lifetime distributions. *IEEE Transactions on Reliability*, **61**, 798-808.
- [5] Flehinger, B.J., Reiser, B. and Yashchin, E. (2002). Parametric modeling for survival with competing risks and masked failure causes. *Lifetime Data Analysis*, **8**, 177-203.
- [6] Kalbfleisch, J.D. and Prentice, R.L. (2002), *The statistical analysis of failure time data*, 2nd ed. John Wiley & Sons.
- [7] Sarhan, A.M. and El-Bassiouny, A. H. (2003), Estimation of components reliability in parallel system using masked system life data. *Applied Mathematics and Computation*, **138**(1), 61-75.
- [8] Wang, R., Sha, N., Gu, B. and Xu, X. (2015), Parameter Inference in a Hybrid System with Masked Data. *IEEE Transactions on Reliability*, **64**, (2), 636-644.
- [9] Xu, A., Basu, S. and Tang, Y. (2014), A full Bayesian approach for masked data in step-stress accelerated life testing. *IEEE Transactions on Reliability*, **63**, (3), 798-806.
- [10] Xu, A.C. and Tang, Y.C. (2012), Statistical analysis of competing failure modes in accelerated life testing based on assumed copulas. *Chinese Journal of Applied Probability and Statistics*, **28**,(1), 5162.
- [11] Zhang, X.P., Shang, J.Z., Chen, X., Zhang, C.H. and Wang, Y.S. (2014), Statistical inference of accelerated life testing with dependent competing failures based on copula theory. *IEEE Transactions on reliability*, **63**, (3), 764-780.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## Bi-objective reliability-redundancy allocation problem with cold standby strategy

Mohammadi, Z. <sup>1</sup>      Zarezadeh, S. <sup>2</sup>      Dehbozorgi, M. <sup>3</sup>

<sup>1</sup> Department of Statistics, Jahrom University

<sup>2</sup> Department of Statistics, Shiraz University

<sup>3</sup> Department of Medical physics and biomedical Engineering, Shiraz university of medical science.

### Abstract

This paper is concerned with the problem of optimization of reliability redundancy allocation with cold standby strategy in series-parallel systems. Most of the existing approaches in the reliability-redundancy allocation problems (RRAPs) literature concentrate only on maximizing the reliability function. Since there is an uncertainty in the estimates of components reliability, we does not only optimize the system reliability and also consider the associated variance structure to reduce the variability in system reliability.

**Keywords:** Reliability-redundancy allocation, Cold-standby strategy, Genetic algorithm.

## 1 Introduction

Reliability optimization is an important problem that has been received considerable attention by researchers in last decades. To enhance the system reliability, the following approaches can be used:

- Increasing the reliability of components;

---

<sup>1</sup>z.mohammadi@jahromu.ac.ir

- b) Handling redundant components in parallel which is called redundancy allocation problem (RAP);
- c) Merging the cases of (a) and (b) which is called reliability-redundancy allocation problem (RRAP);
- d) Reassignment of interchangeable components;

see, [6]. In RAP, it is assumed that the reliability of components are predesignated and the aim is to find the type and number of components that should be used in each subsystem to maximize the system reliability. RRAP considers component reliability as a decision variable.

RAP uses two different strategies: active and standby. In the active strategy, although one component is necessary to function, all redundant components simultaneously works from time zero. The standby strategy is categorized to three types as called cold, warm, and hot standby. In cold-standby redundancy, the redundant components are not affected by the operational stresses due to system operation and hence the failure of system is not occurred before its start. When the warm-standby redundancy is used, the components are subject to operational stresses. In the hot-standby redundancy, the failure of component does not depend on whether the component is idle or in operation. The mathematical formulation for the hot-standby strategy is the same as the case where the active redundancy strategy is applied.

Most of the studies that have been done in RRAP have considered the active redundancy strategy. In the literature, different solution methods have been proposed such as dynamic programming [5] and genetic algorithm [7]. For more studies on RRAP with active strategy, see, e.g., [9].

For the first time, Ardakan and Hamadani [1] considered RRAP with cold-standby strategy. They presented a modified genetic algorithm to maximize the system reliability under the non-linear constraints of weight, cost, and volume.

As mentioned, in RRAPs the reliability of components are uncertain and they are estimated from the field or by test data. Therefore, we must consider the estimation of its reliability and the uncertainty associated with such estimates in the selection of an optimal design configuration. Consequently, in this paper, the system design problem is formulated with two objectives:

1. Maximizing the system reliability estimate;
2. Minimizing the estimation of variance of system reliability.

We apply a bi-objective genetic algorithm to solve the proposed model. Section 2 is devoted to present the structure of the problem. In section 3, the genetic algorithm is presented for solving the problem. Last section considers some experimental results.

## 2 Problem definition and formulation

Consider a series-parallel system with the structure as depicted in Figure 1, where  $s$  subsystems connected to each other in series and each subsystem is designed as a parallel structure. It is assumed that each subsystem has at least one component. Further, the subsystems are working independently and the components of each subsystem have independent and identically distributed (i.i.d.) lifetimes.

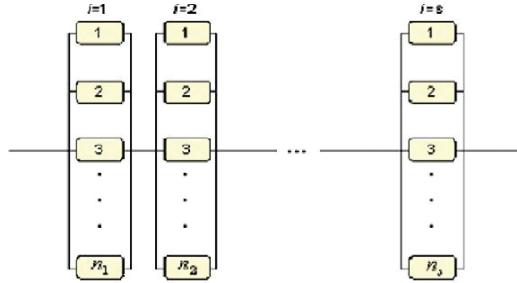


Figure 1: Series-parallel system.

Table 1 describes the used notations in model.

Table 1: Nomenclature

Notation	Description
$t$	The mission time
$\lambda_i$	The failure rate of components in subsystem $i$
$i$	Index for subsystem
$n_i$	Number of components in subsystem $i$
$R$	Reliability of the system
$R_i$	Reliability of subsystem $i$
$s$	The number of subsystems
$var(.)$	The variance of the quantity in the parenthesis
$r_i$	Common reliability of components in subsystem $i$

The reliability of a series-parallel system is written as follows:

$$R(t) = \prod_{i=1}^s R_i(t). \quad (1)$$

From [3], the reliability of a cold-standby subsystem with perfect switching is given by

$$R_i(t) = r_i(t) + \sum_{k=1}^{n_i-1} \int_0^t r_i(t-u) f_i^{(k)}(u) du, \quad (2)$$

where  $f_i^{(k)}(t)$  is the probability density function of the  $k$ th failure arrival time in subsystem  $i$ , i.e. the sum of  $k$  i.i.d. component failure times. In this paper, we assume that the

components of subsystem  $i$  have exponential distribution with failure rate  $\lambda_i$ . Then Eq. (2) can be rewritten as bellow:

$$R_i(t) = r_i(t) + \sum_{k=1}^{n_i-1} \frac{e^{-\lambda_i t} (\lambda_i t)^k}{k!} \quad (3)$$

It is remarkable that  $r_i(t)$  is calculated based on the amount of  $\lambda_i$ . Therefore,  $\lambda_i$  and  $n_i$  are two decision variables in the model.

From [2] the system reliability estimates and its variance are given by:

$$\hat{R}(t) = \prod_{i=1}^s \hat{R}_i(t) \quad (4)$$

and

$$\text{var}(\hat{R}(t)) = \prod_{i=1}^s \left\{ R_i^2(t) + \text{var}(\hat{R}_i(t)) \right\} - \prod_{i=1}^s R_i^2(t) \quad (5)$$

Suppose that  $t_{i1}^*, t_{i2}^*, \dots, t_{im}^*$  are the observed values of a sample of size  $m$  from exponential distribution with failure rate  $\lambda_i$ . The maximum likelihood estimator (MLE) of  $\lambda_i$  can be easily obtained as

$$\hat{\lambda}_i = \frac{m}{\sum_{j=1}^m t_{ij}^*}. \quad (6)$$

From delta method and Zehna's theorem, we have

$$\hat{R}_i(t) = \sum_{k=0}^{n_i-1} \frac{e^{-\hat{\lambda}_i t} (\hat{\lambda}_i t)^k}{k!} \quad (7)$$

and

$$\widehat{\text{var}}(\hat{R}(t)) \simeq \prod_{i=1}^s \left\{ \hat{R}_i^2(t) + \widehat{\text{var}}(\hat{R}_i(t)) \right\} - \prod_{i=1}^s \hat{R}_i^2(t) \quad (8)$$

where

$$\widehat{\text{var}}(\hat{R}_i(t)) \simeq \frac{\hat{\lambda}_i^2}{m} \left( \frac{\hat{\lambda}^{(n_i-1)} t^{n_i} e^{-\hat{\lambda}_i t}}{(n_i - 1)!} \right)^2. \quad (9)$$

Therefore, the RRAP can be written as follows:

$$\begin{aligned} & \text{Maximize } \hat{R}(t) \\ & \text{Minimize } \widehat{\text{var}}(\hat{R}(t)) \end{aligned}$$

### 3 Genetic algorithm

Genetic algorithms (GA) are a family of optimization methods based on evolution. In GA, Data structures are considered as simple chromosomes. The first step in any genetic algorithm is to generate an initial population, which is a set of various recombination of

the genes of each chromosome. Applying each individual to an objective function (evolution function) determines the suitability of each member of the population. Individuals with the best response to the objective function are chosen to be parents of the next generation [8]. Next step would be to generate offsprings (children) and also to mutate the children to maintain genetic diversity. Mutation is often a small random change in some of the genes of the offsprings. The next population consists of: 1) a subset of best chromosomes of the past population which at least involves the parents, 2) offsprings, 3) mutated individuals, and 4) a set of randomly generated members. This procedure goes on in an iterative sequence to find the optimal or sub-optimal solution to the problem [4]. The GA problem implemented here is described in the following:

1. Generating an initial population with size  $P = 100$ . Each individual has two chromosomes  $\mathbf{n}$  and  $\boldsymbol{\lambda}$ , each of which consists of  $s$  components (genes)  $(n_1, n_2, \dots, n_s)$  and  $(\lambda_1, \lambda_2, \dots, \lambda_s)$  where  $s$  is the number of subsystems.

$\mathbf{n}$  has  $s$  random integer values between 1 and  $N$ , where  $N$  is the maximum number of redundant components  $n_i = m, m \in [1, N]$ .

The values for  $\boldsymbol{\lambda}$  are randomly generated based on a exponential distribution. To calculate these values, we took the following steps:

- (a)  $l_i = \text{random} \in [0, 1], i = 1, \dots, s$  where  $l_i$  belongs to  $i$ th subsystem.
  - (b)  $t_{i1}, \dots, t_{im} \sim \text{Exponential}(l_i)$ .
  - (c)  $\lambda_i = \frac{m}{\sum_{j=1}^m t_{ij}}$
2. Since we have two assessment criteria for the optimization problem, we ought to consider a multi-objective GA. Therefore, each individual is applied to Eq. (4) and Eq. (8) separately and the results are combined together as one objective value. It should be noted that our goal is to maximize the reliability and minimize the variance; also the scope of change for these two variables are not the same. Accordingly, we have to normalize all the values for reliability and variance first, then, to minimize reliability, it should be multiplied by -1.

$$R_S = -a\hat{R}(t) + (1 - a)\widehat{\text{var}}(\hat{R}(t)),$$

where  $a$  is the ownership ratio.

3. All  $P$  values for  $R_S$  are sorted in an ascending order. Since our optimization problem is a minimization problem, the two first values are chosen as parents ( $n_{p1.2}$  and  $\lambda_{p1.2}$ ).
4. Breeding:

$$n_{c1} = \text{fix}(rn_{p1} + (1 - r)n_{p2}) , \quad n_{c2} = \text{fix}(rn_{p2} + (1 - r)n_{p1})$$

$$\lambda_{c1} = r\lambda_{p1} + (1 - r)\lambda_{p2} , \quad \lambda_{c2} = r\lambda_{p2} + (1 - r)\lambda_{p1},$$

where  $r$  is the offspring rate.

5. Mutation: in this stage some genes of the children are randomly chosen and changed with a mutation probability  $P_\mu$ .
6. Generating new population as described below:
  - (a) Include 20 percent of the best members of the previous population.
  - (b) Include children and mutated offsprings.
  - (c) Generate random individuals as described in Step 1.
7. Go to Step 2.
8. The loop continues for 1000 iterations.

## 4 Experimental results

In this section, we apply the proposed GA to obtain the optimal system configuration for series-parallel system with 5 subsystems. The minimum and maximum numbers of components in each subsystem are 1 and 5, respectively. Therefore, in the proposed algorithm, it is assumed that  $s = 5$  and  $N = 5$ . The computer program is developed in MATLAB environment. To run this program we use a intel core i5-3337U computer with a 1.8 GHz CPU and 6 GB RAM. Other parameters used in the proposed algorithm: the mission time  $t = 10$ , the population size  $P = 100$ , offspring rate  $r = 0.2$ , mutation probability  $P_\mu = 0.1$ , number of iteration 1000 and  $m = 5$ .

To show the advantage of the proposed bi-objective model, we compare this model with the model described by [1] as follows:

### One-objective Problem:

$$\text{Maximize} \quad \hat{R}(t)$$

Subject to:

$$\begin{aligned} \sum_{i=1}^s w_i v_i^2 n_i^2 &\leq V \\ \sum_{i=1}^s e^{0.25n_i} w_i n_i &\leq W \\ \sum_{i=1}^s \alpha_i (n_i + e^{0.25n_i}) (-\frac{1000}{ln r_i})^2 &\leq C \end{aligned}$$

where  $V$ ,  $W$  and  $C$  are the limitations on system volume, weight and cost, respectively. The input parameters for this problem were presented in Table 3 [1]. It should be mentioned that the proposed GA not exactly the same as GA was used in [1].

Tables 2 and 3 show the simulation results of the 5 independent runs of the two problems. Furthermore, the standard deviation (STD) values of the the objective functions in the 5 different runs are measured. The smaller STDs indicate the robustness of the used algorithm. It can be seen from the tables that the bi-objective problem works better than the one-objective problem in terms of the two criteria, reliability and STD.

Table 2: Bi-objective problem

Parameter	$i$	Run #1	Run #2	Run #3	Run #4	Run #5
R	1	1	1	1	0.9004	0.9692
$\lambda_i$	1	0.0046	0.0063	0.0029	0.0121	0.0113
	2	0.006	0.0164	0.0092	0.0041	0.0028
	3	0.0133	0.0028	0.0099	0.0058	0.0031
	4	0.0007	0.0102	0.0086	0.0082	0.0005
	5	0.0216	0.0073	0.0008	0.0193	0.0052
$R_i$	1	0.955041962	0.938943474	0.971416464	0.88603396	0.89315066
	2	0.941764534	0.848742022	0.91210515	0.95982913	0.972388367
	3	0.875465092	0.972388367	0.905742708	0.943649947	0.969475573
	4	0.993024443	0.903029552	0.917594231	0.921271959	0.995012479
	5	0.805735302	0.92960083	0.992031915	0.824481974	0.949328867
$n$		(4,4,5,3,5)	(4,4,5,5,4)	(4,5,5,4,4)	(4,3,4,4,5)	(4,4,4,3,4)
STD		0.0018669				

Table 3: One objective problem

Parameter	$i$	Run #1	Run #2	Run #3	Run #4	Run #5
R		0.6667	0.9983	0.6889	0.9649	0.5574
$\lambda_i$	1	0.0373	0.0024	0.0011	0.0163	0.0446
	2	0.0964	0.0011	0.0767	0.0132	0.0684
	3	0.0546	0.0021	0.0403	0.0073	0.0801
	4	0.0557	0.0049	0.0789	0.0287	0.0821
	5	0.0092	0.001	0.0089	0.0249	0.0857
$R_i$	1	0.688665233	0.97628571	0.989060279	0.849591188	0.640183772
	2	0.381364373	0.989060279	0.464404193	0.876340995	0.504594572
	3	0.579262231	0.979218965	0.668312099	0.92960083	0.44887986
	4	0.572925264	0.95218113	0.454298867	0.750511729	0.439991443
	5	0.91210515	0.990049834	0.914845574	0.779579973	0.424433475
$n$		(2,4,2,2,1)	(2,3,2,3,2)	(3,2,2,4,1)	(2,2,2,3,3)	(2,2,2,3,3)
STD		0.0381041				

## References

- [1] Ardakan, M.A. and Hamadani, Z.A. (2014), Reliability-redundancy allocation problem with cold-standby redundancy strategy, *Simulation Modelling Practice and Theory*, **42**, 107-118.
- [2] Coit, D.W. (1997), System-reliability confidence-intervals for complex-systems with estimated component reliability. *IEEE Transactions on Reliability*, **46** (4), 487-493.
- [3] Coit, D.W. (2001), Cold-standby redundancy optimization for non-repairable system. *IIE Transactions*, **33**, 471-478.
- [4] Coley, D.A. (1999), *An introduction to genetic algorithms for scientists and engineers*, World Scientific Publishing Co Inc.

- [5] Kuo, W. Prasad, R. Tillman,F.A. and Hwang, C.L. (2001), *Optimal Reliability Design: Fundamentals and Applications*, Cambridge University Press, Cambridge.
- [6] Kuo, W. and Prasad, R. (2000), An annotated overview of system reliability optimization. *IEEE Transactions on Reliability.*, **49** (2), 176-187.
- [7] Lins, I.D. and Drogueut, E.L. (2011), Redundancy allocation problems considering systems with imperfect repairs using multi-objective genetic algorithms and discrete event simulation. *Simul. Model. Pract. Theory.*, **19**, 362381.
- [8] Whitley, D. (1994) A genetic algorithm tutorial. *Statistics and computing.*, **4** (2), 65-85.
- [9] Yeh, W.C. Hsieh, T.J. (2011), Solving reliability redundancy allocation problems using an artificial bee colony algorithm. *Computers and Operations Research.*, **38**, 14651473.



# Bayesian designing of step-stress accelerated degradation test under an inverse Gaussian process with tampered failure rate model

Mosayebi E.<sup>1</sup>      Azizi F.<sup>2</sup>      Shemehsavar S.<sup>3</sup>      Haghghi F.<sup>4</sup>

<sup>1,2,3,4</sup> School of Mathematics, Statistics and Computer Sciences, University of Tehran

## Abstract

In this paper, we consider a step-stress accelerated degradation test (SSADT) when the degradation follows an inverse Gaussian (IG) process. It is assumed that change of stress from a level to another has a multiplicative effect on the degradation path and a tampered failure rate (TFR) model type is used. This study proposes a Bayesian method to obtain parameter estimates and optimal design for simple step-stress accelerated life tests by minimizing the expected pre-posterior variance of the  $p$ -quantile of the products lifetime distribution. Finally, a real-world example is analyzed to illustrate the application of the proposed methods.

**Keywords:** Step-Stress Accelerated Degradation Test, Tampered Failure Rate Model, Inverse Gaussian Process, Bayesian Optimal Test Plan.

## 1 Introduction

With great industrial improvement, the new products are more reliable and can operate for a long time before they fail. This makes it difficult for engineers to obtain sufficient failure data within a reasonable life-testing time. As a result, it is not easy to obtain lifetime distribution using the traditional life testing. Accelerated life test (ALT) is a technique to overcome this problem, in which products are subjected to conditions severe than normal conditions of testing to obtain the failure data sooner. Then, the failure

times observed under harsh conditions must be analyzed to draw inferences on the life distribution corresponding to the normal operating condition. There are several models to extrapolate the result of ALT: the tampered random variable (TRV) model, the cumulative exposure (CE) model and the tampered failure rate (TFR) model developed by [1].

Even, using ALT for some high reliable products may offer little help. If there exists a quality characteristic (QC) of the product whose degradation over time can be related to reliability, then collecting degradation data can provide useful information about the reliability of the product. Also, to hasten the procedure of collecting degradation data, accelerated degradation test (ADT) can be employed. An ADT can provide reliability information quickly, hence it has an important role in saving time and money in reliability. There are two main models for ADT data. The first model is known as general sample path and the second model is stochastic processes which are used to capture the time-dependent structure of the degradation over time. The most common stochastic processes employed for modeling degradation data are Wiener process and Gamma process. Because of simplicity, Wiener process was employed in many literatures. In the context of optimal ADT plan we can name works such as [2], and [5] when the degradation follows a Wiener process. Gamma process, a stochastic process with monotone sample path, is an appropriate model dealing with irreversible degradation data. Optimal ADT with considering gamma process as a degradation model can be found in [6], and [10].

Inverse Gaussian(IG) process is another stochastic process which is recently introduced to deal with degradation data. IG process has a monotone increasing behavior and has great properties when dealing with covariates and random effects. Regarding IG process as a degradation model, [9] considered problem of finding optimal constant stress accelerated degradation test (CSADT) under constraint that the experiment cost does not exceed a pre-specified budget. Recently, [7] obtained the optimal SSADT based on IG process using CE model.

In the ADT planning problems, the commonly used optimization criterion is to minimize the asymptotic variance of the maximum likelihood estimator (MLE) of a reliability measure at the normal conditions. The optimal design depends on the model parameters, and usually uses pre-estimates of the model parameters. In practice, it is possible that there is uncertainty in the values of the model parameters. Usually this uncertainty is considered by performing a sensitivity analysis by changing the values of the model parameters. An alternative approach to dealing with the uncertainty is to apply the Bayesian approach. This model expresses the uncertainty in the planning values by a joint prior distribution of the model parameters and minimizes the expected pre-posterior variance of the reliability measure at the normal conditions.

Bayesian optimality for the ADTs have been studied by some researchers. For example, [4] also discussed obtaining optimal CSADT with Bayesian perspective. [3] investigated the Bayesian planning method for SSADT using the IG process, with the objectives of relative entropy, quadratic loss function and D-optimality. In this paper, we apply the idea of the TFR model. Then, under the constraint that the total experimental cost does not exceed a predetermined budget, the optimal settings of an SSADT are obtained by minimizing the expected pre-posterior variance of the  $p$ -quantile of the products lifetime distribution.

The remainder of the paper is organized as follows. In Section 2, we describe the IG process, the SSADT and the required assumptions. The estimation problem is discussed in Section 3. In Section 4 an optimal test plan is presented. Finally, a real-world example is analyzed to illustrate the application of the proposed methods in Section 5.

## 2 Degradation modeling and assumptions

We consider the following basic assumptions:

**Assumption 1:** Let  $Z(t|S_0)$  denote the degradation path of the product under normal stress ( $S_0$ ). The product is assumed to fail when the degradation path  $Z(t|S_0)$  for the first time crosses a pre-specific threshold level  $\omega$ . The first passage time,  $T_\omega$ , is called the product's lifetime. In this paper,  $Z(t|S_0)$  is modeled as an IG degradation process, having the properties:

- $Z(t|S_0)$  has independent increments; that is,  $Z(t_{i+1}|S_0) - Z(t_i|S_0)$  and  $Z(t_{j+1}|S_0) - Z(t_j|S_0)$  are independent  $\forall t_{j+1} > t_j \geq t_{i+1} > t_i$ .
- Each increment,  $\Delta Z(t|S_0) = Z(t + \Delta t|S_0) - Z(t|S_0)$ , follows an Inverse Gaussian distribution  $IG(\mu_0 \Delta t, \lambda_0 \Delta t^2)$  with the probability density function (PDF) given by:

$$f_{\Delta Z}(y) = \sqrt{\frac{\lambda_0 \Delta t^2}{2\pi y^3}} \exp\left(-\frac{\lambda_0(y - \mu_0 \Delta t)^2}{2y\mu_0^2}\right), \quad y > 0,$$

where  $\mu_0 \Delta t$  and  $\lambda_0 \Delta t^2$  are the mean and scale parameter for fixed  $\Delta t$ , respectively.

The cumulative distribution function (CDF) of  $T_\omega = \inf\{t | Z(t|S_0) > \omega\}$  can be simply obtained by:

$$\begin{aligned} F_{T_\omega} &= P(T_\omega < t) = P(Z(t|S_0) > \omega) \\ &= \Phi[\sqrt{\lambda_0/\mu_0}(t - \omega/\mu_0)] - e^{2\lambda_0 t/\mu_0} \Phi[-\sqrt{\lambda_0/\mu_0}(t + \omega/\mu_0)], \end{aligned} \tag{1}$$

where  $\Phi$  is the standard normal CDF. Based on the CDF in Eq. 1, there is no closed form for the  $p$ -quantile of lifetime distribution and [9] noted that the CDF of  $T_\omega$  can be approximated as:

$$F_{T_\omega} = 1 - \Phi\left(\frac{\omega - \mu_0 t}{\sqrt{\mu_0^3 t / \lambda_0}}\right) = \Phi\left(\sqrt{\lambda_0 t / \mu_0} - \omega \sqrt{\frac{\lambda_0 / \mu_0^3}{t}}\right).$$

Using this approximation, the  $p$ -quantile of  $T_\omega$  is:

$$q_p = \frac{\mu_0}{4\lambda_0} \left( z_p + \sqrt{z_p^2 + 4\omega\lambda_0/\mu_0^2} \right)^2 \tag{2}$$

where  $z_p$  is the standard normal  $p$ -quantile.

**Assumption 2:** A simple step-stress ADT with two stress levels  $S_1$  and  $S_2$  ( $S_1 < S_2$ ), is considered where  $S_1$  and  $S_2$  are assumed to be pre-specified. The test is conducted as

follows. Suppose the test units are initially placed on stress  $S_1$  and the degradation data are recorded with a measurement frequency per  $f$  units of time. The stress level remains at  $S_1$  until stress change time  $\tau_1$ . Next, stress is increased to higher stress level  $S_2$  and held constant until the test termination time  $\tau_2$ . Therefore, we have  $\tau_1 = \ell_1 \times f$  and  $\tau_2 = (\ell_1 + \ell_2) \times f$  where  $\ell_1$  and  $\ell_2$  are the numbers of measurements under stress  $S_1$  and  $S_2$ , respectively.

**Assumption 3:** Based on the idea of [1], we represent a statistical model for simple step-stress ADT in which a change of stress has a multiplicative effect on the degradation path. According to this assumption, the degradation path of the simple step-stress ADT of the product at time  $t$ ,  $Z_{SS}(t)$ , can be expressed as:

$$Z_{SS}(t) = \begin{cases} Z(t|S_1) & 0 < t \leq \tau_1 \\ \alpha Z(t|S_1) & \tau_1 < t \leq \tau_2, \end{cases} \quad (3)$$

where  $Z(t|S_1)$  is the degradation path of the product under the stress  $S_1$  and  $\alpha \geq 1$  is the acceleration factor which is an unknown parameter of the model. Further, we assume that the independent increments of the degradation path at the first stress level,  $S_1$ , follow an IG distribution with the mean  $\mu_1 \Delta t$  and the scale parameter  $\lambda_1 \Delta t^2$  for fixed  $\Delta t$ . Then, according to the properties of the IG distribution, the distribution of the independent increments of the degradation path at the second stress level,  $S_2$ , will be an IG distribution with mean  $\alpha \mu_1 \Delta t$  and scale parameter  $\alpha \lambda_1 \Delta t^2$  for fixed  $\Delta t$ .

### 3 Parameter estimation

In this section, we deal with the estimation of the unknown parameters based on the maximum likelihood and Bayesian methods. Suppose  $n$  units are placed on a step-stress ADT and let  $Z_{SS}^{(i)}(t)$  denote the degradation value of the  $i$ -th unit at time  $t$ . For  $0 < t_{j-1} < t_j \leq \tau_1$ , define the degradation increment as follows:

$$\begin{aligned} Y_{ij} &= Z_{SS}^{(i)}(t_j) - Z_{SS}^{(i)}(t_{j-1}) \\ &= Z^{(i)}(t_j|S_1) - Z^{(i)}(t_{j-1}|S_1), \end{aligned}$$

for  $i = 1, \dots, n$  and  $j = 1, \dots, \ell_1$ . Then, from the above assumptions, we have:

$$Y_{ij} \sim IG(\mu_1 \Delta t_j, \lambda_1 \Delta t_j^2),$$

while for  $\tau_1 < t_{j-1} < t_j \leq \tau_2$ , we know that  $Y_{ij} \sim IG(\alpha \mu_1 \Delta t_j, \alpha \lambda_1 \Delta t_j^2)$ , for  $i = 1, \dots, n$  and  $j = \ell_1 + 1, \dots, \ell_1 + \ell_2$ .

Given the observed degradation values  $\{Z_{SS}^{(i)}(t_j)\}_{j=1}^{\ell_1+\ell_2}$ ,  $i = 1, \dots, n$ , the likelihood function based on the simple step-stress ADT model for an IG degradation process is obtained as:

$$\begin{aligned} L(\mu_1, \lambda_1, \alpha | \underline{y}) &= \prod_{i=1}^n \left\{ \prod_{j=1}^{\ell_1} \left( \frac{\lambda_1 \Delta t_j^2}{2\pi y_{ij}^3} \right)^{\frac{1}{2}} \exp\left(-\frac{\lambda_1(y_{ij} - \mu_1 \Delta t_j)^2}{2\mu_1^2 y_{ij}}\right) \right. \\ &\quad \times \left. \prod_{j=\ell_1+1}^{\ell_1+\ell_2} \left( \frac{\alpha \lambda_1 \Delta t_j^2}{2\pi y_{ij}^3} \right)^{\frac{1}{2}} \exp\left(-\frac{\lambda_1(y_{ij} - \alpha \mu_1 \Delta t_j)^2}{2\alpha \mu_1^2 y_{ij}}\right) \right\}. \end{aligned} \quad (4)$$

Let  $m_1 = \frac{1}{\mu_1}$ . From the likelihood function (4), we can derive the kernel functions of the unknown parameters  $m_1$ ,  $\lambda_1$ , and  $\alpha$ , as:

$$p(m_1|\lambda_1, \alpha, \underline{y}) \propto \exp \left( -\frac{\lambda_1}{2} \left( m_1^2 (A_1 + \frac{1}{\alpha} A_2) - 2m_1 \sum_{i=1}^n \sum_{j=1}^{\ell_1+\ell_2} \Delta t_j \right) \right), \quad (5)$$

$$p(\lambda_1|m_1, \alpha, \underline{y}) \propto \lambda_1^{\frac{n(\ell_1+\ell_2)}{2}} \exp \left( -\frac{\lambda_1}{2} \left( m_1^2 (A_1 + \frac{1}{\alpha} A_2) - 2m_1 \sum_{i=1}^n \sum_{j=1}^{\ell_1+\ell_2} \Delta t_j + B_1 + \alpha B_2 \right) \right), \quad (6)$$

and

$$p(\alpha|m_1, \lambda_1, \underline{y}) \propto \alpha^{\frac{n\ell_2}{2}} \exp \left( -\frac{\lambda_1}{2} \left( m_1^2 A_2 \frac{1}{\alpha} + \alpha B_2 \right) \right), \quad \alpha \geq 1, \quad (7)$$

where

$$A_1 = \sum_{i=1}^n \sum_{j=1}^{\ell_1} y_{ij}, \quad A_2 = \sum_{i=1}^n \sum_{j=\ell_1+1}^{\ell_1+\ell_2} y_{ij}, \quad B_1 = \sum_{i=1}^n \sum_{j=1}^{\ell_1} \frac{\Delta t_j^2}{y_{ij}}, \quad B_2 = \sum_{i=1}^n \sum_{j=\ell_1+1}^{\ell_1+\ell_2} \frac{\Delta t_j^2}{y_{ij}}.$$

Observe that  $p(m_1|\lambda, \alpha, \underline{y})$  and  $p(\lambda_1|m_1, \alpha, \underline{y})$  belong to the Normal and Gamma families, respectively.

Based on the kernel functions in Eq. 5 and 6, and having regard to  $\alpha \geq 1$ , the following prior distributions are considered for the parameters:

$$\begin{aligned} m_1 &\sim N(a_1, a_2), \\ \lambda_1 &\sim G(b_1, b_2), \\ \alpha &\sim TG(c_1, c_2), \end{aligned} \quad (8)$$

where  $a_1$ ,  $a_2$ ,  $b_1$ ,  $b_2$ ,  $c_1$ , and  $c_2$  are known and non-negative hyper-parameters. Hence, the full conditional posterior distributions of the parameters are obtained as:

$$\begin{aligned} \pi(m_1|\lambda_1, \alpha, \underline{y}) &= N\left(\frac{\lambda_1 \sum_{i=1}^n \sum_{j=1}^{\ell_1+\ell_2} \Delta t_j + \frac{a_1}{a_2}}{\lambda_1(A_1 + \frac{1}{\alpha} A_2) + \frac{1}{a_2}}, \frac{1}{\lambda_1(A_1 + \frac{1}{\alpha} A_2) + \frac{1}{a_2}}\right), \\ \pi(\lambda_1|m_1, \alpha, \underline{y}) &= G\left(\frac{(m_1^2 (A_1 + \frac{1}{\alpha} A_2) - 2m_1 \sum_{i=1}^n \sum_{j=1}^{\ell_1+\ell_2} \Delta t_j + B_1 + \alpha B_2)}{2} + b_1, \frac{n(\ell_1 + \ell_2)}{2} + b_2\right), \end{aligned}$$

and

$$\pi(\alpha|m_1, \lambda_1, \underline{y}) \propto p(\alpha|m_1, \lambda_1, \underline{y})TG(c_1, c_2), \quad \alpha \geq 1.$$

To draw samples from the posterior distributions of the interesting parameters, we use WinBUGS software and implement the Gibbs sampling procedure to obtain values from the above full conditional posterior distribution. How well our samples resemble the true posterior distribution depends heavily on the number of samples that we draw. We assume that the number of generated samples is large enough for reasonable precision.

Based on the Gibbs sampling method, we generate the samples as follows:

1. Choose  $m_1^{(0)} = \frac{1}{\hat{\mu}_1}$ ,  $\lambda_1^{(0)} = \hat{\lambda}_1$ , and  $\alpha^{(0)} = \hat{\alpha}$ .
2. Set  $k = 0$ .
3. Generate  $m_1^{(k+1)}$  from  $\pi(m_1|\lambda_1^{(k)}, \alpha^{(k)}, \underline{y})$ . Thus,  $\mu_1^{(k+1)} = \frac{1}{m_1^{(k+1)}}$ .
4. Generate  $\lambda_1^{(k+1)}$  from  $\pi(\lambda_1|m_1^{(k+1)}, \alpha^{(k)}, \underline{y})$ .
5. Generate  $\alpha^{(k+1)}$  from  $\pi(\alpha|m_1^{(k+1)}, \lambda_1^{(k+1)}, \underline{y})$ .
6. set  $k = k + 1$ .

Repeat steps 3 – 6 many times and eventually, obtain  $K$  samples from  $\pi(\mu_1, \lambda_1, \alpha|y)$ , having regard to the burn-in time in the Markov chains. The Bayes estimates of the parameters under the squared error loss function can be obtained as:

$$\tilde{\mu}_1 = \frac{1}{K} \sum_{k=1}^K \mu_1^{(k)}, \quad \tilde{\lambda}_1 = \frac{1}{K} \sum_{k=1}^K \lambda_1^{(k)}, \quad \tilde{\alpha}_1 = \frac{1}{K} \sum_{k=1}^K \alpha_1^{(k)}. \quad (9)$$

## 4 Optimization

Let define the design variable  $D = (n, \ell_1, \ell_2, f)$  where  $n$ ,  $f$ , and  $\ell_i$ ,  $i = 1, 2$  are respectively the sample size, measurement frequency, and the number of measurements for the stress level  $S_i$ ,  $i = 1, 2$ . For a given test plan  $D$  with the degradation observations  $\mathbf{Y}_i = \{y_i(t_j)\}_{j=1}^{\ell_1+\ell_2}$  where  $i = 1, \dots, n$ , the posterior distribution of  $\theta$  is obtained by using Bayes theorem:

$$p(\theta|\mathbf{Y}, D) = \frac{L(\mathbf{Y}|\theta, D)\pi(\theta)}{f_Y(\mathbf{Y}|D)}$$

where  $\pi(\theta)$  is the prior distribution of  $\theta$ ,  $L(\mathbf{Y}|\theta, D)$  is the likelihood function of degradation observations  $\mathbf{Y}$  under a specific test plan  $D$ , and  $f_Y(\mathbf{Y}|D)$  is defined as follows:

$$f_Y(\mathbf{Y}|D) = \int_{\theta} L(\mathbf{Y}|\theta, D)\pi(\theta)d\theta$$

Also, consider that  $Var(q_p(\theta)|\mathbf{Y}, D)$  be the posterior variance of the  $q_p(\theta)$  which is obviously depends on the degradation observations.

Then, to describe the relationship between the test plans and estimation precision, the average of the posterior variance  $Var(q_p(\theta)|\mathbf{Y}, D)$  over the degradation observations  $\mathbf{Y}$  is used as the Bayesian planning criterion. Such value is computed as:

$$AVar(q_p(\theta)|D) = \int Var(q_p(\theta)|\mathbf{Y}, D)f_Y(\mathbf{Y}|D)d\mathbf{Y}$$

Moreover, assume that  $TC(D)$  denotes the total experiment cost that should not exceed a pre-specified budget  $C_b$ . Clearly, both  $AVar(q_p(\theta)|D)$  and the total experiment cost are affected by design variables  $D = (n, \ell_1, \ell_2, f)$ . Then a typical problem of interest can be formulated as:

Minimize:  $AVar(q_p(\theta)|D)$ ,

subject to:  $TC(D) \leq C_b$ ,

where  $D \in \mathbb{N}^4$ . In the following subsections, we will develop a formula for the total experiment cost and discuss the steps to obtain the Bayesian optimal design.

## 4.1 The total cost of experimentation

The total cost of the experiment  $TC(D)$  consists of three parts:

- the cost of conducting an experiment is  $C_{op}f \sum_{i=1}^2 \ell_i$ ;
- the cost of measurement is  $C_{mea}n \sum_{i=1}^2 \ell_i$ ; and
- the cost of testing devices is  $nC_{it}$ ;

where  $C_{op}$ ,  $C_{mea}$ , and  $C_{it}$  are the unit cost of operation, the unit cost of measurement, and the unit cost of an item, respectively. Hence, the total cost of the experiment can be formulated as:

$$TC(D) = C_{op}f \sum_{i=1}^2 \ell_i + C_{mea}n \sum_{i=1}^2 \ell_i + nC_{it}.$$

## 4.2 Bayesian optimal design

Let  $N$  be the total number of the test designs  $D = (n, \ell_1, \ell_2, f)$  that meets the cost condition. For each design  $D_i = (n_i, \ell_{1i}, \ell_{2i}, f_i)$ ,  $i = 1, \dots, N$ , we can calculate the Bayesian planning criterion  $AVar(q_p(\theta)|D_i)$  with following steps.

- Step 1: Draw  $R$  sets of degradation observations independently (i.e.,  $\mathbf{Y}_{i,k}$ ,  $k = 1, \dots, R$ ). These sets are implemented by generating a set of model parameters  $\theta_{i,k}$  from the prior distributions given in the previous section, and then generating degradation observations  $\mathbf{Y}_{i,k} = \{y_{i,k}(t_j)\}_{j=1}^{\ell_1+\ell_2}$  with the test plan  $D_i = (n_i, \ell_{1i}, \ell_{2i}, f_i)$  from the IG process model.
- Step 2: For each set of degradation observations  $\mathbf{Y}_{i,k}$ ,  $k = 1, \dots, R$ , calculate the posterior distribution of the model parameters. Since the posterior distribution is not easy to obtain, Laplace's approximation can be used to give an appropriate approximation.
- Step 3: Find the Hessian matrix of the above approximated distribution and then use delta method to obtain  $Var(q_p(\theta)|\mathbf{Y}_{i,k}, D_i)$ .
- Step 4: Calculate the mean value of  $Var(q_p(\theta)|\mathbf{Y}_{i,k}, D_i)$  for the  $R$  sets of degradation observations to obtain  $AVar(q_p(\theta)|D_i)$  under the test plan  $D_i$ .

Finally, find the optimal design  $D^* = (n^*, \ell_1^*, \ell_2^*, f^*)$  is obtained as

$$D^* = \arg \min_{\{D_i\}_{i=1}^N} AVar(q_p(\theta)|D_i)$$

## 5 An illustrative example

In this section, we illustrate our proposed method through a numerical example based on the data from the stress relaxation problem described by [8]. Considering low-informative priors, the Bayese estimates are:

Table 1: bayes estimation Results and RMSEs (in parenthese) for the Model Parameters Based on Stress Relaxation Data

$\alpha$	$\mu_1$	$\lambda_1$
1.498476 (0.063391)	0.003447273 (0.001085)	4.76135e - 05 (0.000462)

Moreover, with the cost configration  $(C_{it}, C_{op}, C_{mea}) = (50, 2.5, 1.5)$ , the Bayesian optimal design based n the previous mentioned method is obtained for  $C_b = 3000$ . Then,

Table 2: Bayesian optimal test plan under  $C_b = 3000$ .

$n^*$	$\ell_1^*$	$\ell_2^*$	$f^*$	total cost	$AVar(q_p(\theta) D)$
15	18	15	18	2977.5	13257.06

under the pre mentioned cost configuration and the total affordable budeget  $C_b$ , the optimal items is 15, the optimal frequency of measurement is 18 and the optimal number of degradation measurement at stress level  $S_1$  and  $S_2$  are 18 and 15, respectively. The Bayesian optimal design for various  $C_b$  or different prior assumption can be done easily.

## References

- [1] Bhattacharyya, G. and Soejoeti, Z. (1989), A tampered failure rate model for step-stress accelerated life test. Communications in statistics- Theory and methods, vol. 18, no. 5, pp. 1627-1643.
- [2] Hu, H., Lee, M. Y. and Tang, J. (2015), Optimum step-stress accelerated degradation test for Wiener degradation process under constraints, European Journal of Operational Research, 241, 412-421.
- [3] Li, X., Hu, Y., Zio, E., and Kang, R. (2017). A Bayesian optimal design for accelerated degradation testing based on the inverse Gaussian process. IEEE Access, 5, 5690-5701.
- [4] Peng, W., Liu, Y., Li, Y. F., Zhu, S. P. and Huang, H. Z. (2014), A Bayesian optimal design for degradation tests based on the inverse Gaussian process, Journal of Mechanical Science and Technology, 28 (10), 3937-3946.
- [5] Sung, S. L. and Yum, B. J. (2016), Optimal design of step-stress accelerated degradation tests based on the Wiener degradation process. Quality Technology and Quantitative Management, 13(4), 367-393.

- [6] Tsai, C. C., Tseng, S. T. and Balakrishnan, N. (2012), Optimal Design for Degradation Tests Based on Gamma Processes With Random Effects, *IEEE Transactions on Reliability*, 61(2), 604-613.
- [7] Wang, H., Wang, G. J. and Duan, F. J. (2016), Planning of step-stress accelerated degradation test based on the inverse Gaussian process, *Reliability Engineering and System Safety*, 154, 97-105.
- [8] Yang, G. (2007), *Life Cycle Reliability Engineering*. Hoboken, NJ, USA: Wiley.
- [9] Ye, Z. S., Chen, L. P., Tang, L. C. and Xie, M. (2014). Accelerated degradation test planning using the inverse Gaussian process, *IEEE Trans Reliab*, 63(3),750-63.
- [10] Zhang, C., Lu, X., Tan, Y. and Wang, Y. (2015), Reliability demonstration methodology for products with Gamma Process by optimal accelerated degradation testing, *Reliability Engineering and System Safety*, 142, 369-377.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## Connections between diversion concepts and reliability measures

Nasr Esfahani, S. <sup>1</sup>      Asadi, M. <sup>2</sup>

<sup>1,2</sup> Department of Statistics, University of Isfahan

### Abstract

The aim of this paper is to investigate the connection between the Gini coefficient and some important reliability measures. After reviewing the recent results in connections between the reliability measures and well known measures of disparity, such as cumulative residual entropy and Gini coefficient, we show that the Gini coefficient can be represented as the sum of expectation of the mean residual life (MRL) of a series system and the mean inactivity time (MIT) of a parallel system lifetime with three components. We also show that the expectation of the MRL (MIT) of a stand by system consisting of  $n$  components can be decomposed in terms of the the expectation of MRLs (MITs) of its components. This leads to an upper bound for the cumulative residual (past) entropy of such system in terms of cumulative residual (past) entropies of the system components.

**Keywords:** Gini mean difference, mean residual life, cumulative residual entropy, mean inactivity time.

## 1 Introduction

Let  $X$  be continuous random variable with distribution function  $F$  and survival function  $\bar{F} = 1 - F$ . The Gini mean difference (GMD) associated to random variable  $X$  is defined as

$$GMD = E(|X_1 - X_2|) = 2 \int F(x)\bar{F}(x)dx,$$

---

<sup>2</sup>m.asadi@sci.ui.ac.ir

where  $X_1$  and  $X_2$  are independent random variables distributed as  $X$ . It is argued by [10] that the GMD, as a measure of variability, shares many properties of the variance of  $X$  and is more informative than the variance for the distributions that are far from normality. It should be noted that the GMD is the difference between the expected values of the maxima and the minima in a sample of two independent and identically distributed random variables  $X_1$  and  $X_2$ . That is

$$GMD = E[\max(X_1, X_2) - \min(X_1, X_2)],$$

(see, [10]). An interesting observation regarding the variance of  $X$  and the GMD is that both of them can be written as especial cases of covariance. In fact, one can write  $Var(X) = Cov(X, X)$ , while  $GMD = 4Cov(X, F(X))$ , (see, [11]). That is,  $GMD$  can be employed to measure the linear correlation between  $X$  and  $F(X)$  which is a measure of correlation between  $X$  and rank of the values of  $X$ .

$GMD$  has a close relation with Gini coefficient of inequality. The Gini coefficient of a wealth random variable with distribution  $F$  is mean  $\mu < \infty$  is defied as

$$G(X) = \frac{1}{\mu} \int_0^\infty F(x)\bar{F}(x)dx = \frac{GMD}{2\mu}.$$

This is the most important measure of inequality in economic studies and  $0 \leq G(X) \leq 1$ , where  $G(X) = 0$  when the wealth is distributed uniformly and  $G(X) = 1$  indicates the full concentration of the wealth at a single point. [2] have shown that GMD can be expressed as the sum of expectation of mean residual life and mean inactivity time functions. The MRL is an important tool in reliability and survival analysis to describe and analyze the lifetime data. For a nonnegative continuous random variable  $X$  the MRL at time  $t$  is defined as

$$m(t) = E(X - t|X > t) = \frac{\int_t^\infty \bar{F}(x)dx}{\bar{F}(t)}.$$

The mean inactivity time (MIT) of  $X$  is defined as

$$k(t) = E(t - X|X < t) = \frac{\int_0^t F(x)dx}{F(t)}.$$

The function  $k(t)$  is a dual measure of MRL and gives the mean inactivity time of a system with lifetime  $X$  given that the system has failed at a time before  $t$ .

[2] showed that  $G(X)$  can be represented in terms of the mean residual life of minima (series system) and mean inactivity time of maxim (parallel system) of two nonnegative random variables (see Section 2).

Another important measure of disparity which has applications in different areas is cumulative residual entropy (CRE). Let  $X$  be a nonnegative continuous random variable with DF  $F$  and survival function  $\bar{F}$ . The CRE is a measure of uncertainty defined as (see, [5])

$$\mathcal{E}_1(X) = - \int_0^\infty \bar{F}(x) \log \bar{F}(x)dx.$$

It is shown that the CRE  $\mathcal{E}_1(X)$  has applications in computer vision and image alignment (see, [5] and [8]). [1] showed the CRE can be represented in terms of MRL. In fact, the cited authors proved that the CRE can be represented as  $\mathcal{E}_1(X) = E(m(X))$ . Recently Asadi (2017) proved that the CRE has the following representation:

$$\mathcal{E}_1(X) = Cov(X, H(X)), \quad (1)$$

where  $H(x) = -\log \bar{F}(x)$ , is called the cumulative hazard rate. It has to be pointed out here that the covariance representation (1) is valid for random variables with support  $(a, b)$  where  $-\infty < a < b \leq \infty$ . On the other hand, a dual of uncertainty measure  $\mathcal{E}_1(X)$  can be defined based on distribution function  $F$  which we call it cumulative past entropy (CPE). The CPE is defined as

$$\mathcal{E}_2(X) = - \int_0^\infty F(x) \log F(x) dx,$$

(see, also [3]). It is easy to show that  $\mathcal{E}_2(X)$  can be expressed as  $\mathcal{E}_2(X) = E(k(X))$ , where  $k(t)$  is mean inactivity time (MIT) of  $X$ . Similar to  $\mathcal{E}_1(X)$ , it can be easily shown that  $\mathcal{E}_2(X)$  has a covariance representation as follows.

$$\mathcal{E}_2(X) = Cov(X, \log F(X)), \quad (2)$$

where this covariance representation is valid for random variables with support  $(a, b)$ ,  $-\infty < a < b \leq \infty$ .

The purpose of the present communication is to study the relation between reliability measures and disparity measures. The rest of the paper is organized as follows: In Section 2 we show that The GMD can be represented as the sum expectation of the MRL of a series system and the MIT of a parallel system consists of three iid components. In Section 3, we show that the expectation of the MRL (MIT) of a stand by system consisting of  $n$  components can be decomposed in terms of the the expectation of MRLs (MITs) of its components. This leads to an upper bound for the cumulative residual (past) entropy of such system in terms of cumulative residual (past) entropies of the system components.

## 2 Relation between GMD and reliability measures

Asadi et al. (2017) have shown that the GMD can be represented in terms of the MRL of a series system and the MIT of a parallel system with two independent components as follows:

$$E(m_2(X)) = \frac{1}{2}GMD,$$

where

$$m_2(t) = \frac{\int_t^\infty \bar{F}^2(x) dx}{\bar{F}^2(t)}$$

is the MRL of a two component series system and

$$E(k_2(X)) = \frac{1}{2}GMD,$$

where

$$k_2(t) = \frac{\int_0^t F^2(x)dx}{F^2(t)}$$

is the MIT of a two component parallel system. This in turn shows that

$$E(m_2(X) + k_2(X)) = \text{GMD}. \quad (3)$$

A natural question here is that: Is there any relation between the GMD and the MRL (or MIT) of the systems with more than two components? Let us consider a series system with three iid components where the component lifetimes has a common continuous distribution  $F$  with reliability function  $\bar{F}$ . Then for the series system with three iid components the MRL is:

$$m_3(t) = \int_t^\infty \frac{\bar{F}^3(u)}{\bar{F}^3(t)} du,$$

and for a parallel system with three iid components the MIT is

$$k_3(t) = \int_0^t \frac{F^3(u)}{F^3(t)} du. \quad (4)$$

It is easy to show that

$$\begin{aligned} E[m_3(X)] &= \frac{1}{2} \left( \int_0^\infty \bar{F}(x)dx - \int_0^\infty \bar{F}^3(x)dx \right) \\ &= \frac{1}{2} \int_0^\infty \bar{F}(x)F(x)(1 + \bar{F}(x))dx \\ &\leq \frac{1}{2} \text{GMD}. \end{aligned}$$

Similarly we have  $E[k_3(X)] \leq \frac{1}{2} \text{GMD}$ .

In the following theorem we prove a result similar to (3) for the systems with three components.

**Theorem 1.** *Let  $m_3(t)$  and  $k_3(t)$  be the MRL and MIT of the lifetime of a series systems with three iid lifetimes. Then*

$$E(m_3(X) + k_3(X)) = \frac{3}{4} \text{GMD},$$

where the expectation is with respect to the density of  $F$ .

*Proof.* We have

$$\begin{aligned} E[m_3(X)] &= \int_0^\infty \int_t^\infty \frac{\bar{F}^3(x)dx}{\bar{F}^3(t)} f(t)dt, \\ &= \int_0^\infty \bar{F}^3(x) \left( \int_0^x \frac{f(t)dt}{\bar{F}^3(t)} \right) dx \\ &= \frac{1}{2} \left( \int_0^\infty \bar{F}(x)dx - \int_0^\infty \bar{F}^3(x)dx \right). \end{aligned} \quad (5)$$

Similarly one can show that

$$\begin{aligned}
E[k_3(X)] &= \int_0^\infty \int_0^t \frac{F^3(x)dx}{F^3(t)} f(t)dt, \\
&= \int_0^\infty F^3(x) \left( \int_x^\infty \frac{f(t)dt}{F^3(t)} \right) dx \\
&= \frac{1}{2} \left( \int_0^\infty F(x)dx - \int_0^\infty F^3(x)dx \right). \tag{6}
\end{aligned}$$

From (5) and (6) we get

$$\begin{aligned}
E[m_3(X) + k_3(X)] &= \frac{1}{2} \int_0^\infty [F(x) + \bar{F}(x) - \bar{F}^3(x) - F^3(x)]dx. \\
&= \frac{3}{2} \int_0^\infty \bar{F}(x)(1 - \bar{F}(x))dx \\
&= \frac{3}{4} GMD.
\end{aligned}$$

□

Asadi et al (2017) defined a general form of CRE as follows: The generalized entropy functional of order  $\alpha$  is defined as

$$h_\alpha(\bar{F}) = - \int_0^\infty \bar{F}^\alpha(x) L_\alpha(\bar{F}(x))dx, \quad \alpha > 0, \tag{7}$$

where

$$L_\alpha(x) = \begin{cases} \frac{x^{1-\alpha}-1}{1-\alpha} & x \geq 0, \alpha \in \{0\} \cup \Omega \\ \log x & x > 0, \alpha = 1. \end{cases}$$

A dual of the generalized entropy can be defined if one replace the survival function  $F$  with distribution function  $F$ . The cited authors showed that for series system with  $n$  components and mean residual life  $m_n(t)$  we have

$$h_n(\bar{F}) = E(m_n(X)),$$

where the expectation is with respect to the density of  $F$ . According to the result we have obtained in Theorem 1, we can immediately conclude that the GMD can be represented in terms of generalized entropy of order 3 as:

$$GMD = \frac{4}{3}(h_3(\bar{F}) + h_3(F)).$$

An interesting open problem here is that whether one can extend similar results of Theorem 1, and what we have just observed on generalized entropy of order  $n$ , regarding the systems with more than three components. We leave this question for a future study.

### 3 A decomposition formula

As it was mentioned in the Section 2, the CRE has the following representation in terms of MRL

$$E(m(X)) = - \int \bar{F}(x) \log \bar{F}(x) dx = \mathcal{E}_1(X) \quad (8)$$

Also we mentioned that the CRE has the following covariance representation:

$$\mathcal{E}_1(X) = Cov(X, H_X(X)).$$

In this section we show that the CRE of a stand by system can be decomposed in terms of CRE of its components. Consider a system with the following structure. The system is constructed of  $n$  identical units which are connected sequentially. We assume that when unit number 1 with lifetime  $X_1$  fails unit number 2 with lifetime  $X_2$  starts operating automatically, and so on until the  $n$ th unit, with lifetime  $X_n$ , fails. We further assume that the units have independent lifetimes. Then the lifetime of the system is  $T = \sum_{i=1}^n X_i$ . Assume that  $\mu_i = E(X_i)$  is the mean time to failure of  $X_i$  and  $\mu = E(T) = \sum_{i=1}^n \mu_i$  is the mean time to failure of the system.

Before giving the main result let us consider the following lemma:

**Lemma 3.1.** *Let  $X_1, X_2, \dots, X_n$  be iid and assume that  $T = \sum_{i=1}^n X_i$ . Then*

$$Cov(T, H_T(T)) = \sum_{i=1}^n Cov(X_i, H_T(T)) \quad (9)$$

*Proof.* The proof follows from the covariance properties of sum of random variables.  $\square$

Now we have the following result.

**Theorem 2.** *The cumulative residual entropy of the system lifetime has the following decomposition:*

$$\mathcal{E}(T) = Cov(T, H_T(T)) = \sum_{i=1}^n \alpha(X_i, T) \mathcal{E}(X_i)$$

where

$$\alpha(X_i, T) = \frac{Cov(X_i, H_T(T))}{Cov(X_i, H_{X_i}(X_i))}$$

and  $\mathcal{E}(X_i)$  is the entropy of the components.

*Proof.* We have

$$\begin{aligned} \mathcal{E}(T) = Cov(T, H_T(T)) &= \sum_{i=1}^n Cov(X_i, H_T(T)) \\ &= \sum_{i=1}^n \frac{Cov(X_i, H_T(T))}{Cov(X_i, H_{X_i}(X_i))} Cov(X_i, H_{X_i}(X_i)) \\ &= \sum_{i=1}^n \alpha(X_i, T) \mathcal{E}(X_i) \end{aligned}$$

$\square$

**Corollary 3.1.** It is interesting to note that  $\alpha(X_i, T)$  is always between zero and one. This is so because in  $\alpha(X_i, T)$ ,  $H_T(T)$  is trivially an increasing function of  $X_i$  and hence  $0 \leq \text{Cov}(X_i, H_T(T))$ . This in turn implies that  $0 \leq \alpha(X_i, T)$ . On the other hand it is easy to show that  $\text{Cov}(X_i, H_T(T)) \leq \text{Cov}(X_i, H_{X_i}(X_i))$ . Hence we have  $\alpha(X_i, T) \leq 1$ . This result shows that the cumulative residual entropy can be decomposed as a mixture of the cumulative residual entropy of its components. We also conclude from this result that

$$\mathcal{E}_1(T) \leq \sum_{i=1}^n \mathcal{E}_1(X_i),$$

That is, the uncertainty of the system is less than the sum of uncertainties of its components in the sense of cumulative residual entropy. In particular when the  $X_i$ 's are iid we have  $\mathcal{E}_1(T) \leq n\mathcal{E}_1(X_1)$ .

**Corollary 3.2.** As a result we can also conclude equivalently that for the system described above

$$E(m_T(T)) \leq \sum_{i=1}^n E(m_{X_i}(X_i)),$$

where  $m_T$  and  $m_{X_i}$  are the MRL's of the system and the components, respectively. In particular when  $X_i$ 's are iid we have  $E(m_T(T)) \leq nE(m_{X_1}(X_1))$ .

For the dual measure of uncertainty  $\mathcal{E}_2(X)$  we have  $\mathcal{E}_2(X) = E(k(X))$ , where  $k(t)$  is MIT of  $X$ . Also, as we mentioned  $\mathcal{E}_2(X)$  has a covariance representation as follows

$$\mathcal{E}_2(X) = \text{Cov}(X, \phi_X(X)),$$

where  $\phi_X(X) = \log F(X)$ . Hence based on what we have discussed above, the cumulative past entropy of the system described above can be written as

$$\begin{aligned} \mathcal{E}_2(T) = \text{Cov}(T, \phi_T(T)) &= \sum_{i=1}^n \text{Cov}(X_i, \phi_T(T)) \\ &= \sum_{i=1}^n \beta(X_i, T) \mathcal{E}_2(X_i) \end{aligned}$$

where

$$\beta(X_i, T) = \frac{\text{Cov}(X_i, \phi_T(T))}{\text{Cov}(X_i, \phi_{X_i}(X_i))}$$

From this we conclude, based on the fact that  $0 \leq \beta(X_i, T)$ ,

$$\mathcal{E}_2(T) \leq \sum_{i=1}^n \mathcal{E}_2(X_i),$$

which is also equivalent to say that

$$E(k_T(T)) \leq \sum_{i=1}^n E(k_{X_i}(X_i)),$$

where  $k_T$  and  $k_{X_i}$  are the MIT's of the system and the components, respectively.

## References

- [1] Asadi, M & Zohrevand, Y (2007), ‘On the dynamic cumulative residual entropy’, *Journal of Statistical Planning and Inference*, **137**(6), 1931–1941.
- [2] Asadi, M, Ebrahimi, N, & Soofi, E (2017), ‘Connections of Gini, Fisher, and Shannon by Bayes risk under proportional hazards’, *J. Applied Probability*, **54**, 1027–1050.
- [3] Di Crescenzo, A, & Longobardi M (2009). ‘On cumulative entropies’, *Journal of Statistical Planning and Inference*, **139**(12), 4072–87.
- [4] Rao, M (2005), ‘More on a new concept of entropy and information’, *Journal of Theoretical Probability*, **18**(4), 967–981.
- [5] Rao, M, Chen, Y, Vemuri, BC, & Wang, F (2004), ‘Cumulative residual entropy: a new measure of information’, *Information Theory, IEEE Transactions on*, **50**(6), 1220–1228.
- [6] Schechtman, E, & Yitzhaki, S. (1999), ‘On the proper bounds of the Gini correlation’, *Economics letters*, **63**(2), 133–138.
- [7] Schechtman, E, & Yitzhaki, S (1987), ‘A measure of association based On Gini’s mean difference’, *Communications in statistics-Theory and Methods*, **16**(1), 207–231.
- [8] Wang, F, & Vemuri, BC, (2007), ‘Non-rigid multi-modal image registration using cross-cumulative residual entropy’, *International journal of computer vision*, **74**(2), 201–215.
- [9] Wang, Y, Hossain, AM, & Zimmer, WJ (2003), ‘Monotone log-odds rate distributions in reliability analysis’, *Communications in statistics-Theory and Methods*, 2227–2244.
- [10] Yitzhaki, S (2003), ‘Gini’s mean difference: A superior measure of variability for non-normal distributions’, *Metron*, **61**(2), 285–316.
- [11] Yitzhaki, S & Schechtman, E (2013), ‘*The Gini Methodology: A Primer on a Statistical Methodology*’, Springer: New York.
- [12] Yitzhaki S, Wodon QT (2003), ‘Mobility, inequality, and horizontal equity’, Available at SSRN 419322.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## Reliability analysis of $k$ -out-of- $n:F$ system with soft and hard failures

Nezakati, E. <sup>1</sup> Razmkhah, M. <sup>2</sup>

<sup>1,2</sup> Department of Statistics, Ferdowsi University of Mashhad

### Abstract

In this paper, degradation performance of a  $k$ -out-of- $n:F$  system with soft and hard failures is considered. Four different cases are introduced for intensity function of hard failures. The expected number of soft and hard failures and the reliability function are derived in these cases. Also, a comparison study is done for a 2-out-of-4: $F$  system to compare reliability and expected number of hard failures for different cases of intensity function.

**Keywords:** Degradation, Soft failure, Hard failure, Intensity function, Reliability, Expected number of failures.

## 1 Introduction

Systems have become more and more reliable today, although the demand for highly reliable products also continues to increase. A promising way for reliability modeling for highly reliable systems is to make use of degradation signals that reflect the health conditions of a product. In the past few decades, performance degradation data has played an important role in reliability analysis for high reliability products. For initial studies, refer to [1]. Failures of a complex system can be categorized as soft or hard. As a hard failure causes the system to stop functioning. Soft failures are often hidden and are not self-announcing. The system can continue to operate in the presence of soft failures, although its performance or production capacity may be reduced. Soft failure often occurs

---

<sup>1</sup>nezakati-e@mail.um.ac.ir

due to degradation performance of component. For non-parametric, semi-parametric and parametric estimation of reliability characteristics base degradation, one can refer to [2], [3] and [4], respectively.

Although there are numerous publications about the reliability of k-out-of-n systems, the studies on degradation performance of these systems are limited. [6] investigated reliability function of a  $k$ -out-of- $n$  system that its components exposure to degradation and shock loads. Furthermore the expected number of soft and hard failures for a  $k$ -out-of- $n:F$  system have an important role in maintenance and cost analysis. The main goal of this paper is to investigate the reliability and the expected number of soft and hard failures of a  $k$ -out-of- $n:F$  system based on degradation performance of system. As known in the literature of reliability, a  $k$ -out-of- $n:F$  system is an active redundant system of  $n$  components, in which a minimum of  $n - k + 1$  components must be operational for the system to work. For more details, refer to [5].

The rest of the paper is organized as follows. Section 2 focuses on preliminaries and description of soft and hard failures. In Section 3, the number of soft and hard failures are introduced and a closed form is derived for their expectation. The reliability function of a  $k$ -out-of- $n:F$  system based on soft and hard failures is investigated in Section 4. Also, four different cases of intensity function is introduced for hard failures of this system. In Section 5, a comparison study is done between four cases of intensity functions for 2-out-of-4: $F$  system, reliabilities, expected number of soft and hard failures are compared. Finally, some conclusions are stated in Section 6

## 2 Preliminaries

Consider a  $k$ -out-of- $n:F$  system with two failure modes, soft and hard. A soft failure occurs when the degradation performance of the component exceeds a certain threshold and a hard failure occurs when the component stop working. Denote the failure time of soft and hard failures by  $T_i^0$  and  $T_i^1$ , respectively. Let us consider degradation paths of  $n$  components as  $X_1(t), \dots, X_n(t)$  and define the level of degradation or performance such that when degradation path exceeds the specified threshold, say  $d_i$ , a soft failure occurs. So, the soft failure time of the  $i$ th component, is defined as the time when the degradation path reaches or exceeds the critical threshold  $d_i$ , i.e.,

$$T_i^0 = \inf\{t > 0; \quad X_i(t) \geq d_i\}; \quad i = 1, \dots, n. \quad (1)$$

As previously mentioned, the other type of failure is due to the hard (traumatic) failure. Denote the hazard rate function of hard failure for the  $i$ th component by  $\lambda_i(t)$ . Since the hard failure is influenced by degradation, the hazard rate function depends on  $X_i(t)$  and  $t$ . We assume that given  $X_i(t) = x_i$ , the conditional hazard rate of  $T_i^1$  has a multiplicative form. That is,

$$\lambda_i(t | X_i(t) = x_i) = \lambda_0(t, \theta)\lambda(x_i); \quad i = 1, \dots, n, \quad (2)$$

where  $\lambda_0(\cdot)$  and  $\lambda(\cdot)$  are the baseline hazard and intensity functions, respectively. From model (2), it is observed that the conditional hazard rate function depends on degradation only through the intensity function. Similar to [7], we assume that the baseline hazard

rate function is of the form  $\lambda_0(t, \theta) = (1+t)^\theta$ . So, the conditional hazard rate function of the  $i$ th component is as follows

$$\lambda_i(t | X_i(t) = x_i) = (1+t)^\theta \lambda(x_i); \quad i = 1, \dots, n. \quad (3)$$

In order for the stochastic deterioration process to be monotonic, we can best consider it as a gamma process [8]. In words, a gamma process is a stochastic process with independent, non-negative increments having a gamma distribution. Hence, we assume that  $X_i(t)$  obeys a Gamma process with shape parameter  $a_i(t)$  and scale parameter  $b_i$ , denoted by  $X_i(t) \sim \Gamma(a_i(t), b_i)$ , for  $i = 1, \dots, n$ . Empirical studies show that the shape parameter of Gamma process at time  $t$  can often be described by a power law model. Hence, for positive real constants  $a_i$  and  $\gamma$ , we consider shape parameters as  $a(t) = a_i t^\gamma$ , ( $i = 1, \dots, n$ ).

### 3 Estimation of soft and hard failure numbers

A crucial step in determining maintenance policies and asset life costs is the computation of the expected number of failures for an asset. Denote the number of soft and hard failures up to time  $t$  by  $N_s(t) = \sum_{i=1}^n I(X_i(t) \geq d_i)$  and  $N_h(t) = \sum_{i=1}^n I(T_i^1 \leq t)$ , respectively, where  $I(\cdot)$  stands for indicator function. To compute the expected number of soft and hard failures, we must first compute reliability of  $T_i^0$  and  $T_i^1$ . Using (1), the reliability of  $T_i^0$  is as follows

$$R_{T_i^0}(t) = P(T_i^0 > t) = F_{X_i(t)}(d_i), \quad (4)$$

where  $F_{X_i(t)}(\cdot)$  is the distribution function of gamma process  $X_i(t)$ . Now, to compute the reliability function of  $T_i^1$ , by conditioning on  $X_i(t) = x_i$ , we have

$$\begin{aligned} R_{T_i^1}(t) &= P(T_i^1 > t) \\ &= \int_0^\infty P(T_i^1 > t | X_i(t) = x_i) f_{X_i(t)}(x_i) dx_i \\ &= \int_0^\infty \exp\left\{-\int_0^\infty \lambda_i(s | X_i(s) = x_i) ds\right\} f_{X_i(t)}(x_i) dx_i \\ &= \int_0^\infty \exp\left\{-\lambda(x_i) \int_0^\infty (1+s)^{\theta+1} ds\right\} f_{X_i(t)}(x_i) dx_i \quad (5) \\ &= \int_0^\infty \exp\left\{-\frac{\lambda(x_i)}{\theta+1}((1+t)^{\theta+1} - 1)\right\} f_{X_i(t)}(x_i) dx_i, \quad (6) \end{aligned}$$

where (5) is resulted using (3).

Now, to obtain  $E(N_s(t))$  and  $E(N_h(t))$  we have

$$E(N_s(t)) = \sum_{i=1}^n P(X_i(t) \geq d_i) = n - \sum_{i=1}^n R_{T_i^0}(t) \quad (7)$$

and

$$E(N_h(t)) = \sum_{i=1}^n P(T_i^1 \leq t) = n - \sum_{i=1}^n R_{T_i^1}(t), \quad (8)$$

where, (7) and (8) are computed using (4) and (6), respectively. So, expected number of soft and hard failures can obtain using (7) and (8), respectively. In the next section, we study the reliability function of  $k$ -out-of- $n:F$  system by considering some different cases of intensity function of hard failures.

## 4 Reliability function of $F$ -system

Consider a  $k$ -out-of- $n:F$  system with failure times of  $T_1, \dots, T_n$ , such that  $T_i = \min\{T_i^0, T_i^1\}$ . Denote the corresponding order statistics by  $T_{1:n} < \dots < T_{n:n}$ . Then the failure time of the system is given by  $T_{k:n}$ . To determine the reliability function of such a system, let us first present reliability function of the  $i$ th component in the following theorem.

**Theorem 4.1.** *Assume that the soft and hard failure times are  $T_i^0$  and  $T_i^1$ , respectively. Reliability function of the  $i$ th component at time  $t$  is as follows*

$$R_{T_i}(t) = \int_0^{d_i} \exp\left\{\frac{1}{\theta+1}(1 - (1+t)^{\theta+1})\lambda(x_i)\right\} \frac{1}{b_i^{a_i t^\gamma} \Gamma(a_i t^\gamma)} x_i^{a_i t^\gamma - 1} \exp\left\{-\frac{x_i}{b_i}\right\} dx_i. \quad (9)$$

*Proof.* Note that by conditioning on  $X_i(t) = x_i$ , the random variables  $T_i^0$  and  $T_i^1$  are conditionally independent. So, based on the definition of  $T_i$ , we have

$$\begin{aligned} R_{T_i}(t) &= P(\min\{T_i^1, T_i^2\} > t) \\ &= \int_0^\infty P(T_i^0 > t \mid X_i(t) = x_i) P(T_i^1 > t \mid X_i(t) = x_i) f_{X_i(t)}(x_i) dx_i, \end{aligned} \quad (10)$$

where  $f_{X_i(t)}(\cdot)$  stands for the pdf of  $X_i(t)$ . From (1), the event  $\{T_i^0 > t\}$  is equivalent to  $\{X_i(t) < d_i\}$ . Hence, we have

$$P(T_i^0 > t \mid X_i(t) = x_i) = \begin{cases} 1, & x_i < d_i \\ 0, & x_i > d_i \end{cases}. \quad (11)$$

By substituting (11) in (10), we have

$$\begin{aligned} R_{T_i}(t) &= \int_0^{d_i} P(T_i^1 > t \mid X_i(t) = x_i) f_{X_i(t)}(x_i) dx_i \\ &= \int_0^{d_i} \exp\left\{\frac{1}{\theta+1}(1 - (1+t)^{\theta+1})\lambda(x_i)\right\} f_{X_i(t)}(x_i) dx_i, \end{aligned} \quad (12)$$

where, (12) is obtained similar (6). So, the proof is completed.  $\square$

Note that to evaluate the reliability function in (9), it is needed to determined the functional form of intensity function,  $\lambda(x_i)$ . In Table 1, four different cases of  $\lambda(x_i)$  are suggested and corresponding reliabilities of the  $i$ th component are computed.

**Table 1.** Four cases of intensity function  $\lambda(x_i)$  and corresponding reliabilities.

	$\lambda(x_i)$	$R_{T_i}(t)$
Case I	$\exp\{\beta_i x_i\}$	$\int_0^{d_i} \exp\left\{\frac{\exp\{\beta_i x_i\}}{\theta+1}(1-(1+t)^{\theta+1})\right\} \frac{1}{b_i^{a_i t^\gamma} \Gamma(a_i t^\gamma)} x_i^{a_i t^\gamma - 1} \exp\left\{-\frac{x_i}{b_i}\right\} dx_i$
Case II	$\beta_i x_i$	$\int_0^{d_i} \exp\left\{\frac{\beta_i x_i}{\theta+1}(1-(1+t)^{\theta+1})\right\} \frac{1}{b_i^{a_i t^\gamma} \Gamma(a_i t^\gamma)} x_i^{a_i t^\gamma - 1} \exp\left\{-\frac{x_i}{b_i}\right\} dx_i,$
Case III	$x_i^{\beta_i}$	$\int_0^{d_i} \exp\left\{\frac{x_i^{\beta_i}}{\theta+1}(1-(1+t)^{\theta+1})\right\} \frac{1}{b_i^{a_i t^\gamma} \Gamma(a_i t^\gamma)} x_i^{a_i t^\gamma - 1} \exp\left\{-\frac{x_i}{b_i}\right\} dx_i,$
Case IV	$\beta_i$	$\exp\left\{\frac{\beta_i}{\theta+1}(1-(1+t)^{\theta+1})\right\} F_{X_i(t)}(d_i)$

The reliability function of a  $k$ -out-of- $n:F$  system can be obtained based on [9], p. 96, with some changes, which is as follows

$$R_{k:n}(t) = \sum_{i=0}^{k-1} \sum_{S_i} \left\{ \prod_{\ell=1}^i (1 - R_{T_{j_\ell}}(t)) \prod_{\ell=i+1}^n R_{T_{j_\ell}}(t) \right\}, \quad (13)$$

where  $R_{T_{j_\ell}}(t)$  is as defined in (9) and summation index  $S_i$  extends over all permutations  $(j_1, \dots, j_n)$  of integers  $\{1, \dots, n\}$  such that  $\{j_1 < \dots < j_i\}$  and  $\{j_{i+1} < \dots < j_n\}$ . Note that one can compute the reliability of a  $k$ -out-of- $n:F$  system for all cases of Table 1, by substituting  $R_{T_i}(t)$  of this table in (13). In the next section we compare the reliability and the expected number of soft and hard failures for cases in Table 1.

## 5 Comparison study

In this section, we want to compare the reliability and expected number of soft and hard failures for cases presented in Table 1. To this end, we consider a four components system. For simplicity the interpretation, we consider shape parameter of Gamma process  $X_i(t)$  as a quadratic form, such that  $X_i(t) = a_i t^2$  ( $i = 1, \dots, 4$ ). Also, the parameters of degradation parameters  $(a_i, b_i)$  of Gamma process  $X_i(t)$  are considered to be  $(a_1, b_1) = (0.03, 2)$ ,  $(a_2, b_2) = (0.03, 1)$ ,  $(a_3, b_3) = (0.02, 1)$  and  $(a_4, b_4) = (0.01, 1)$ . The parameter of baseline hazard function  $\lambda_0(\cdot)$  is fixed as  $\theta = 1$  and the parameters of intensity function  $\lambda(x_i)$  are considered to be  $\beta_1 = 1.3, \beta_2 = 1.2, \beta_3 = 1.5$  and  $\beta_4 = 1$  for  $i = 1, \dots, 4$ . Moreover, we assume that the thresholds for degradation of the components are  $d_1 = d_2 = d_3 = d_4 = 2$ . Since, intensity function does not affect on the number of soft failures, therefore,  $E(N_s(t))$  is the same for four different cases presented in Table 1. The expected number of soft failures are computed using (7) and the results are given in Table 2.

The reliabilities and expected number of hard failures for all cases in Table 1 are computed using (13) and (8) based on given parameters and the results are shown in Figures 1 and 2, respectively. From Figure 1, it is seen that

- Reliability function of cases I and IV are lower than cases II and III.
- Reliability of case IV is the lowest among all cases. It says that the system without degradation performance, has low reliability.

- Reliability of case III is the greatest one among all cases. So, intensity function  $\lambda(x_i) = x_i^{\beta_i}$  is the most suitable intensity function for this structure.

Also, from Figure 2 it is observed that

- Expected number of hard failures of cases I and IV are greater than cases II and III.
- Expected number of hard failures of case IV is the greatest one among all cases. It means that the system without degradation performance, has high expected number of hard failures.
- Expected number of hard failures of case III is the lowest one among all cases.

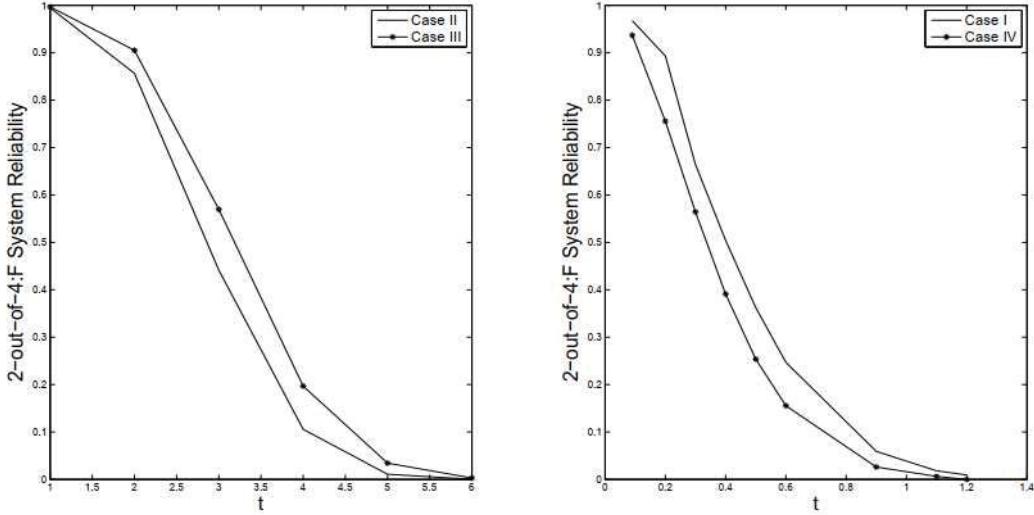


Figure 1: Reliability function for different cases of intensity function.

From the values of Table 2 and Figure 2, it is observed that the expected number of soft failures is lower than the expected number of hard failures. Therefore, to maintain the system, more attention should be paid to hard failures.

**Table 2.** Expected Number of Soft Failures.

$t$	1	2	3	4	5	6
$E(N_s(t))$	0.0021	0.0096	0.0259	0.0572	0.1129	0.2045

## 6 Conclusion

A  $k$ -out-of- $n:F$  system was considered and its degradation performance was investigated based on soft and hard failures. Also, reliability function and expected number of soft

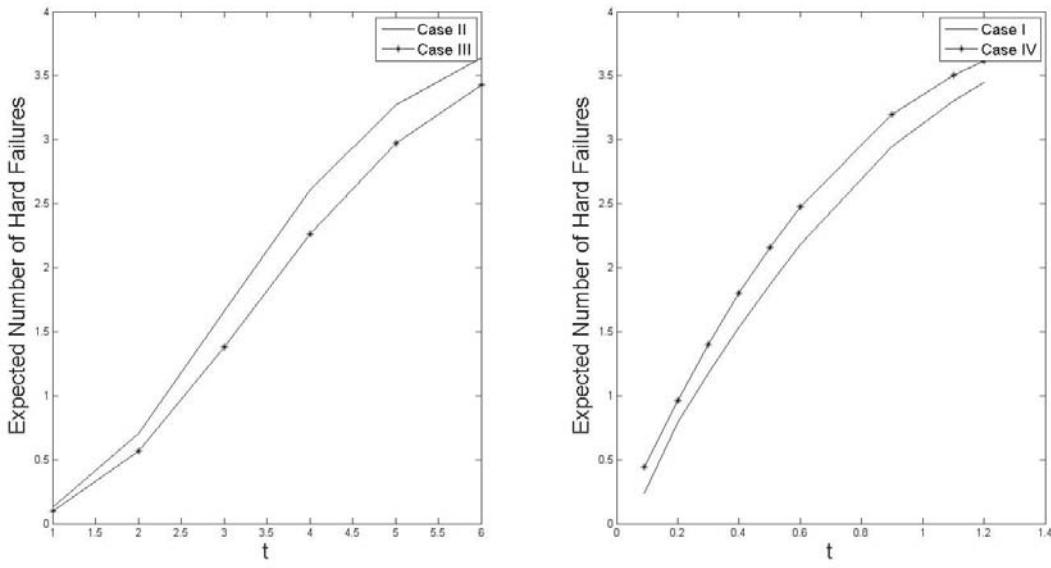


Figure 2: Expected number of hard failures for different cases of intensity function.

and hard failures were derived. Four different cases were considered for intensity function of hard failures and reliability and expected number of soft and hard failures compared for these cases for 2-out-of-4: $F$  system. It was observed that power low model have the greatest reliability and the lowest expected number of failures and is more suitable than the other cases.

## References

- [1] Meeker, W.Q. and Hamada, M. (1995), Statistical tools for the rapid development and evaluation of high-reliability products, *IEEE Transactions on Reliability*, **44**(2), 187-198.
- [2] Bagdonavicius, V., Bikellis, A., and Kazakevicius, V. (2004), Statistical and failure time data with multiple failure modes. *Lifetime Data Analysis*, **10**, 65-81.
- [3] Bagdonavicius, V., Haghghi, F., and Nikulin, M.S. (2005), Statistical analysis of degradation and failure time data using the general path model. *Communication in Statistics-Theory and Methods*, **34**, 1771-1793.
- [4] Haghghi, F., Nazeri rad, N., and Nooraee. N. (2008), On the parametric degradation model with multiple failure modes. *Proceedings of Second International Conference on Accelerated Life Testing in Reliability and Quality Control. IMB, University Victor Segalen Bordeaux 2, Bordeaux, France*.
- [5] Lawless, J. F. (2003), *Statistical models and methods for lifetime data*, Second Edition, Wiley Series in Probability and Statistics.

- [6] Song, S., Chatwattanasiri, N., Coit, D. W., Feng, Q. and Wattanapongsakorn, N., (2012), Reliability analysis for k-out-of-n systems subject to multiple dependent competing failure processes, *Proceeding of International conference on Quality, Reliability, Risk and Maintenance Safety Engineering*, 36-42.
- [7] Nikulin, M. and Wu, H.D.I. (2016). *The Cox model and its applications*, Springer, Chapter 9.
- [8] Noortwijk, J.M.V. (2009), A survey of the application of gamma processes in maintenance. *Reliability Engineering and System Safety*, **94**(1), 2-21.
- [9] David, H. A. and Nagaraja, H. N. (2003), *Order Statistics*, Third Edition, Wiley Series in Probability and Statistics.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## On a goodness of fit test for normality based on Lin-Wong divergence with type-I censored data

Pakgohar, A.<sup>1</sup>      Habibirad, A.<sup>2</sup>      Yousefzadeh, F.<sup>3</sup>

<sup>1,2</sup> Department of Statistics, Ferdowsi University of Mashhad

<sup>3</sup> Department of Statistics, Faculty of Mathematics and Statistics, University of Birjand

### Abstract

Normality testing in goodness of fit test field has long been an interesting issue in statistical inferences. In this paper, we use the Lin-Wong divergence as our proposed measure of distance between distributions and compare its performance with that of other well known distance measures using the empirical distribution function and Shannon entropy concept such as Kullback-Leibler, Cramer von Misses, Anderson-Darling and Kolmogorov-Smirnov. Finally, the use of the proposed test is shown in a illustrative example.

**Keywords:** Anderson-Darling, Goodness of fit test, Lin Wong divergence, Normal distribution, Type I censored scheme.

## 1 Introduction

A divergence measure shows that how one probability distribution diverges from an alternative probability distribution. In other words, it is an instrument for measuring the similarity between two probability distribution patterns. Thus the divergence family unifies commonly used goodness of fit (GoF) test.

The Lin-Wong (LW) divergence measure which was introduced by Lin and Wong [7] yields

---

<sup>1</sup>a\_pakgohar@pnu.ac.ir

<sup>2</sup>ahabibi@um.ac.ir

<sup>3</sup>fyousefzadeh@birjand.ac.ir

a quasi-distance value related to well known Kullback-Leibler (KL) divergence measure which was introduced by Kullback and Leibler [6] and named skew divergence [3]. The LW measure is defined as

$$D_{\text{LW}}(f\|f_0) = \int_{-\infty}^{\infty} f(x) \ln \frac{2f(x)}{f(x) + f_0(x)} dx,$$

in which  $f(x)$  and  $f_0(x)$  represent the probability density function (PDF) for the data in practical and estimated cases, respectively.

The PDF for Type-I censored data represented by a random variable  $\min(X, C)$  with  $C$  as the censor point (a fixed value) would be as the following:

$$f_C(x) = \begin{cases} f(x) & \text{if } x < C \\ \bar{F}(C) & \text{if } x \geq C, \end{cases} \quad (1)$$

where,  $\bar{F}(x) = 1 - F(x)$ , represent survival functions. Here, we construct a GoF test based on LW for normality which can be observed in Type-I censored scheme. This can be often seen in cases which there is not a large enough time window to make observations over all the events while we are not totally inattentive toward events which happen after the censor point. An example can be when we conduct monitoring or follow-up studies on chronic or epidemiological illnesses such as long term cancers or addiction to narcotics.

The LW divergence measure on Type-I censored data is defined as

$$D_{\text{LW}}^{C_1}(f\|f_0) = \int_{-\infty}^C f(x) \ln \frac{2f(x)}{f(x) + f_0(x)} dx + \bar{F}(C) \ln \frac{2\bar{F}(C)}{\bar{F}(C) + \bar{F}_0(C)}. \quad (2)$$

Equivalently,

$$D_{\text{LW}}^{C_1}(f\|f_0) = -H^{C_1}(f) - \int_{-\infty}^C f(x) \ln \frac{f(x) + f_0(x)}{2} dx - \bar{F}(C) \ln \frac{\bar{F}(C) + \bar{F}_0(C)}{2}, \quad (3)$$

where  $H^{C_1}(f) = -\int_{-\infty}^C f(x) \ln f(x) dx - \bar{F}(C) \ln \bar{F}(C)$ , is called as Shannon entropy on Type-I censored data.

The GoF hypothesis test is expressed here as the following:  $H_0 : f(x; \theta) = f_0(x; \theta)$ , while the alternative hypothesis can be written as  $H_1 : f(x; \theta) \neq f_0(x; \theta)$ .

GoF tests are categorized in different ways. The statistical instruments based on empirical distribution function (EDF) and the methods based on Shannon entropy are two well known examples of various types of those tests. The most common measures used usually along with EDF method in tests are: Anderson-Darling (AD), Cramer von Misses (CM) and Kolmogorov-Smirnov (KS). Pettitt and Stephens [9] examined the AD, CM and KS measures when applied to censored data.

Assuming  $u_{(i)} = F_n(x_i; \hat{\theta})$ , and  $u_{(n_0+1)} = F_n(C; \hat{\theta})$ , the revised forms of AD and CM statistics on Type-I censored data are as the following:

$$\begin{aligned} AD = & -\frac{1}{n} \sum_{i=1}^{n_0+1} (2i-1) \{ \ln u_{(i)} - \ln(1 - u_{(i)}) \} - 2 \sum_{i=1}^{n_0+1} \ln(1 - u_{(i)}) \\ & - \frac{1}{n} \{ (n_0 - n + 1)^2 \ln(1 - u_{(n_0+1)}) - (n_0 + 1)^2 \ln u_{(n_0+1)} + n^2, u_{(n_0+1)} \}. \end{aligned} \quad (4)$$

$$CM = \sum_{i=1}^{n_0+1} \left( u_{(i)} - \frac{2i-1}{2n} \right)^2 + \frac{n_0+1}{12n^2} + \frac{n}{3} \left( u_{(n_0+1)} - \frac{n_0+1}{n} \right)^3. \quad (5)$$

Where  $n_0$  is the number of data value less than the fixed value censor point  $C$  within the relevant sample.

Also, the revised form of KS statistic applied on Type-I censored data is defined as the following according to [5]:

$$KS = \max_{1 \leq i \leq n_0} \left\{ \frac{i}{n} - u_{(i)}, u_{(i)} - \frac{i-1}{n}, u_{(n_0+1)} - \frac{n_0}{n} \right\}, \quad (6)$$

Recently many GoF tests have been reported based on normal probability distribution. For instance, Vasicek in (1976), Arizonoand and Ohta (1989), Esteban et al. (2001), Balakrishnan et al. (2007), and Alizadeh (2010) Zamanzade and Arghami (2012) , Bitaraf et al (2017) as well Evren and Ustaoglu (2017) have tested for normal distribution using some distance measures based on Shannon entropy. Furthermore, procedures of GoF tests for exponential and normal probability distributions can be found which are conducted using censored data. Some cases are reported by Pettitt (1976), Park (2005), Balakrishnan et al. (2007), Lim and Park (2007), Habibi Rad et al. (2011), and Park and Shin (2014) as well as Baratpour and Habibi Rad (2016). However, relatively fewer studies have examined the Goodness of Fit tests when applied to Type-I censored data. Park (2013) and Pakyari (2015) can be mentioned as two examples.

This paper is organized as the following: In Section 2 we introduced LW divergence estimators based on Vasicek [10] and Bitaraf et al.[2]. Then, we express the convergence and stability characteristics of those proposed estimators as the new probe of GoF tests on Type-I censored data in order to normal distribution. The critical values and distinction power of AD, CM, KS, KL and two proposed estimators for LW distance measure are presented in section 3. In addition, the results obtained using Monte Carlo (MC) simulation methods along with those mentioned tests are analyzed over two sets of real data samples. Finally, some concluding remarks are presented in section 4.

## 2 Test statistics

We use the LW divergence measure given in (3) to facilitate the decision-making process when choosing the hypothesis for a GoF test. Obviously, the numerical value obtained for LW measure should be non-negative and it is clearly equal to zero if and only if  $f(x; \theta) = f_0(x; \theta)$  a.e.

Considering  $(X_{(i-m)}, X_{(i+m)})$  as an interval with a continuous non-negative PDF  $f(x; \theta)$ , we would have  $x_i \in (x_{(i-m)}, x_{(i+m)})$  in which

$$f(x_i) = \frac{F(x_{(i+m)}) - F(x_{(i-m)})}{x_{(i+m)} - x_{(i-m)}}.$$

Since the EDF converges to the CDF, we can write:

$$\frac{2m/n_0}{x_{(i+m)} - x_{(i-m)}} \xrightarrow{pr} f(x_i) \quad \text{as } n_0 \rightarrow \infty, \quad m \rightarrow \infty, \quad \frac{m}{n_0} \rightarrow \infty, \quad (7)$$

where the window size  $m$  is a positive integer smaller than  $\frac{n_0}{2}$ .

Hence, if we assume  $\hat{\theta}$  to be a convergent estimator for the parameter  $\theta$ , and assume that  $n_0 \rightarrow \infty$ ,  $m \rightarrow \infty$ , and  $\frac{m}{n_0} \rightarrow \infty$  so it can be inferred that:

$$\frac{1}{n} \sum_{i=1}^{n_0} \ln\left(1 + \frac{f_0(x; \hat{\theta})(x_{(i+m)} - x_{(i-m)})}{2m/n_0}\right) \xrightarrow{pr.} E_f\left(\ln\left(1 + \frac{f_0(X; \theta)}{f(X; \theta)}\right)\right). \quad (8)$$

Characteristics of (7) have been studied by [10] in an attempt to find an estimator for entropy. On the other hand, using the definition  $T_i$  provided by [2], we can get  $T_i \xrightarrow{pr.} f(x_i)$ , in which  $T_i = \frac{T_{i_1} + T_{i_2}}{2}$ , and  $T_{i_j} = \frac{n_0}{W_j(m-j)}(X_{(i+m-j)} - X_{(i+m-j)})$ ,  $j = 1, 2$ , so that  $W_j$  satisfies the following relation:

$$W_j = \begin{cases} 1, & 1 \leq i \leq m - j, \\ 2, & m - j + 1 \leq i \leq n_0 - m + j, \\ 1, & n_0 - m + j + 1 \leq i \leq n_0. \end{cases}$$

So, as  $n_0 \rightarrow \infty$ ,  $m \rightarrow \infty$ ,  $\frac{m}{n_0} \rightarrow 0$ , we would have

$$\frac{1}{n} \sum_{i=1}^{n_0} \ln\left(1 + f_0(x; \hat{\theta})T_i\right) \xrightarrow{pr.} E_f\left(\ln\left(1 + \frac{f_0(X; \theta)}{f(X; \theta)}\right)\right). \quad (9)$$

Letting  $n_0 \geq 4$ , and  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n_0)}$  as the order statistics satisfies the condition  $X_{(i)} = X_{(1)}$ , if  $i < 1$ , and  $X_{(i)} = X_{(n_0)}$ , if  $i > n_0$ . Then we can use the Vasicek estimator for the entropy and formulate an estimator based on Shannon entropy for Type-I censored data as the following [10]:

$$H_V^{C1} = \frac{1}{n} \sum_{i=1}^{n_0} \ln \frac{n_0}{2m} (X_{(i+m)} - X_{(i-m)}) - \frac{n - n_0}{n} \ln \frac{n - n_0}{n}. \quad (10)$$

Besides, by taking  $\frac{i-1}{n_0} < p \leq \frac{i}{n_0}$ ,  $i = m + 1, m + 2, \dots, n_0 - m$ , the estimator proposed by Bitaraf et al. for Shannon entropy is defined as the following [2]:

$$H_B^{C1} = \frac{1}{n} \sum_{i=1}^{n_0} \ln T_i - \frac{n - n_0}{n} \ln \frac{n - n_0}{n}. \quad (11)$$

Therefore, the Vasicek estimator for LW divergence measure when applied on Type-I censored data would be given as:

$$LW_V^{C1} = -\frac{1}{n} \sum_{i=1}^{n_0} \ln\left(\frac{1}{2} + \frac{n(X_{(i+m)} - X_{(i-m)})f_0(x; \hat{\theta})}{4m}\right) - \frac{n - n_0}{n} \ln\left(\frac{n - n_0 + n\bar{F}_0(C; \hat{\theta})}{2n(n - n_0)}\right), \quad (12)$$

where  $\bar{F}_0(C; \hat{\theta})$  is the estimated survival function.

The revised form of the KL divergence measure on Type-I censored data is given as:

$$KL_V^{C1} = -\frac{1}{n} \sum_{i=1}^{n_0} \ln\left(\frac{n(X_{(i+m)} - X_{(i-m)})f_0(x; \hat{\theta})}{2m}\right) - \frac{n - n_0}{n} \ln\left(\frac{n\bar{F}_0(C; \hat{\theta})}{n - n_0}\right). \quad (13)$$

Also, we can introduce another estimator for LW from (11) as:

$$LW_B^{C1} = -\frac{1}{n} \sum_{i=1}^{n_0} \ln\left(\frac{1}{2} + \frac{f_0(x; \hat{\theta})T_i}{2}\right) - \frac{n-n_0}{n} \ln\left(\frac{n-n_0+n\bar{F}_0(C; \hat{\theta})}{2n(n-n_0)}\right). \quad (14)$$

Here, we will show that above mentioned tests produce consistent results.

Considering (7) and the definition of Shannon entropy, we would have,

$$\frac{1}{n} \sum_{i=1}^{n_0} \ln \frac{n_0}{2m} (X_{i+m} - X_{i-m}) \xrightarrow{pr} - \int_0^C f(x) \ln f(x) dx.$$

Furthermore, if  $n_0 \rightarrow \infty$ , then  $\frac{n-n_0}{n} \ln \frac{n-n_0}{n} \xrightarrow{pr} \bar{F}(C) \ln \bar{F}(C)$ . So we can write:

$$H_V^{C1} \xrightarrow{pr} H_f^{C1}(X) \quad \text{as} \quad n_0 \rightarrow \infty, m \rightarrow \infty, \frac{m}{n_0} \rightarrow 0.$$

Also, from the relation (8), the actual observed distribution function (EDF) converges to CDF:

$$LW_V^{C1} \xrightarrow{pr} D_{LW}^{C1}(f||f_0) \quad \text{as} \quad n_0 \rightarrow \infty, m \rightarrow \infty, \frac{m}{n_0} \rightarrow 0.$$

Similarly, given the convergence of the expression  $T_i$  which is proved by Bitaraf et al. (2017) we can write [2]:

$$LW_B^{C1} \xrightarrow{pr} D_{LW}^{C1}(f||f_0) \quad \text{as} \quad n_0 \rightarrow \infty, m \rightarrow \infty, \frac{m}{n_0} \rightarrow 0.$$

### 3 Simulation study

In this section, we present the GoF tests and critical values within results; meanwhile we study the power of those tests to recognize normal distribution through test hypothesis. We assess the performance of the proposed tests using simulated Type-I censored data in sample sizes of 10 and 20, as well as the significance levels of  $\alpha = 0.10$  and  $0.05$ . The results of analysis are provided here. We also studied the characteristics of divergence measures within the structure of Shannon entropy. For example LW divergence measure based on Vasicek and Bitaraf estimators besides, KL based on Vasicek estimator in comparison to AD, CM and KS. Since lifetime data distributions are within the alternatives, we would be able to use non-negative data within the sample set and get to a normal standard distribution. Therefore, the censor point  $C$  can be chosen as equivalent to quantiles rather than the median point. In this case, the censor point  $C$  refers to the quantiles 0.7 and 0.8. We determine the critical points using MC simulation with 10000 replicates. All simulations were carried out in R software package. We use the formula  $m = [\sqrt{n} + 0.5]$ , in order to chose the value of  $m$ . This formula was used by Wieczorkowski and Grzegorzewski in 1999.

To facilitate the comparison between distinction powers of the tests, we selected various alternatives such as Beta (B), Standard Cauchy (SC), Gamma (G), Gumbel (Gum), Laplace (Lap), Log-Normal (LN), Logistic (Log), T student (T) and Uniform (U) distributions which can be classified into four groups shown in Table 1. Those mentioned data

Table 1: Alternative distributions used in each power study

Non-negative random variable	
Group I: $X \in (0, 1)$	Group II: $X \in (0, \infty)$
U(0,1)	G(2)
B(2,2)	G(0.5)
B(2,0.5)	LN(0,0.5)
B(0.5, 0.5)	LN(0,1)
Real-valued random variable	
Group III: Symmetric distributions	Group IV: Asymmetric distributions
SC	Gum(2)
T(3)	Gum(0.5)
Log(0,1)	Log(0.5)
Lap(0,1)	Log(2)

distributions have been used by [4], [1] and [2].

From Tables 2 to 5,  $LW_B^{C1}$  statistic is the preferred one when compared to other examined statistics. It seems to have better performance in groups I and II than groups III and IV. When we compared the results in groups III and IV, the  $LW_B^{C1}$  statistic has a much superior power for group IV respect to group III. Our results show that it is better to use AD statistic for group III; while the statistics such as  $KL_V^{C1}$ ,  $LW_V^{C1}$  and  $LW_B^{C1}$  which are based on Shannon entropy and covered by the present paper should be preferred with group IV. For group II of distributions,  $LW_B^{C1}$  statistic seems to outperform the other statistics. As we expected before, the results of tables 4 and 5 are almost the same as tables 2 and 3. Also by comparing tables 2 and 3 to tables 4 and 5 it is clear that the same results are obtained. It can be observed that our proposed test is preferred at  $\alpha = 0.10$ , over the case of  $\alpha = 0.05$ . Besides obviously, when  $n$  number increases  $LW_B^{C1}$  test has more power. Also for group I,  $LW_B^{C1}$  statistic performs quite better than other statistics.

 Table 2: Power comparisons of  $\alpha = 0.05$  and  $F(C) = 0.7$  for normal distribution versus alternatives in Table 1

	$n = 10$						$n = 20$					
	$LW_B^{C1}$	$LW_V^{C1}$	$KL_V^{C1}$	KS	CM	AD	$LW_B^{C1}$	$LW_V^{C1}$	$KL_V^{C1}$	KS	CM	AD
Group I												
U(0,1)	<b>0.1412</b>	0.0947	0.1119	0.0434	0.0314	0.0321	<b>0.2323</b>	0.1528	0.1602	0.0603	0.0485	0.0668
B(2,2)	<b>0.0909</b>	0.0554	0.0666	0.0245	0.0173	0.0164	<b>0.1230</b>	0.0717	0.0629	0.0257	0.0174	0.0209
B(2,0.5)	<b>0.0697</b>	0.0193	0.0416	0.0006	0.0000	0.0000	<b>0.1314</b>	0.0015	0.0154	0.0000	0.0000	0.0000
B(0.5,0.5)	0.2732	0.2160	<b>0.2822</b>	0.1113	0.0974	0.1092	<b>0.5618</b>	0.4178	0.5245	0.2107	0.1919	0.2823
Group II												
G(2)	<b>0.0870</b>	0.0260	0.0549	0.0015	0.0000	0.0000	<b>0.1802</b>	0.0006	0.0241	0.0000	0.0000	0.0000
G(0.5)	0.3144	0.2864	<b>0.3351</b>	0.1528	0.1454	0.1551	<b>0.6175</b>	0.5441	0.6042	0.3087	0.3170	0.4030
LN(0,0.5)	<b>0.0742</b>	0.0187	0.0419	0.0009	0.0000	0.0000	<b>0.1429</b>	0.0009	0.0192	0.0000	0.0000	0.0001
LN(0,1)	<b>0.1434</b>	0.0494	0.0870	0.0055	0.0017	0.0017	<b>0.2405</b>	0.0350	0.0601	0.0046	0.0013	0.0028
Group III												
SC	0.0328	0.0261	0.1726	<b>0.3015</b>	0.2715	0.2876	0.0290	0.0090	0.4377	0.5643	0.565	<b>0.5997</b>
t(3)	0.0473	0.0400	0.0782	0.1323	0.1224	<b>0.1346</b>	0.0236	0.0142	0.1463	0.2474	0.2640	<b>0.2936</b>
Log(0,1)	0.0612	0.0568	0.0714	0.0950	0.0944	<b>0.1061</b>	0.0298	0.019	0.0846	0.1488	0.168	<b>0.1864</b>
Lap(0,1)	<b>0.0593</b>	0.0419	0.0511	0.0458	0.0368	0.0399	0.041	0.0253	0.0335	0.0464	0.0453	<b>0.0511</b>
Group IV												
Gum(2)	<b>0.1006</b>	0.0429	0.0579	0.0170	0.0092	0.0094	<b>0.1248</b>	0.0522	0.0497	0.0169	0.0061	0.0088
Gum(0.5)	0.1250	<b>0.1348</b>	0.1273	0.0696	0.0653	0.0651	<b>0.1276</b>	<b>0.1501</b>	0.1305	0.0835	0.0903	0.0867
Log(0.5)	<b>0.6253</b>	0.5839	0.6206	0.6029	0.5901	0.5960	0.6568	0.6383	<b>0.6473</b>	0.6401	0.6304	0.6321
Log(2)	0.0857	0.0943	<b>0.0946</b>	0.0668	0.0733	0.0803	0.0355	0.0370	0.0538	0.0825	<b>0.0984</b>	<b>0.0984</b>

Table 3: Power comparisons of  $\alpha = 0.05$  and  $F(C) = 0.8$  for normal distribution versus alternatives in Table 1

	$n = 10$						$n = 20$					
	$LW_B^{C1}$	$LW_V^{C1}$	$KL_V^{C1}$	KS	CM	AD	$LW_B^{C1}$	$LW_V^{C1}$	$KL_V^{C1}$	KS	CM	AD
Group I												
U(0,1)	0.2251	0.2348	<b>0.2381</b>	0.0721	0.0855	0.0898	0.2054	0.1690	<b>0.2350</b>	0.1109	0.1505	0.1765
B(2,2)	0.5135	<b>0.5228</b>	0.5213	0.0719	0.0919	0.1078	0.3132	0.2924	<b>0.3234</b>	0.0835	0.1219	0.147
B(2,0.5)	<b>0.0558</b>	0.0123	0.022	0.0062	0.0015	0.0015	<b>0.0647</b>	0.0096	0.0113	0.007	0.0015	0.002
B(0.5,0.5)	0.2002	0.1953	<b>0.2348</b>	0.1239	0.1664	0.1708	0.4898	0.3955	<b>0.5432</b>	0.2575	0.3136	0.3868
Group II												
G(2)	<b>0.1160</b>	0.0184	0.0451	0.0002	0.0001	0.0001	<b>0.2381</b>	0.0072	0.0283	0.0007	0.0001	0.0001
G(0.5)	0.3926	0.3830	<b>0.4249</b>	0.2269	0.2501	0.2497	0.7530	0.7265	<b>0.7797</b>	0.4429	0.5255	0.5876
LN(0,0.5)	<b>0.1250</b>	0.0325	0.0576	0.0048	0.0014	0.0014	<b>0.1742</b>	0.0354	0.0368	0.0052	0.0006	0.0011
LN(0,1)	<b>0.2206</b>	0.0838	0.1189	0.0189	0.0088	0.0074	<b>0.3258</b>	0.1424	0.1330	0.0324	0.0157	0.0215
Group III												
SC	0.0671	0.0507	0.2118	<b>0.2974</b>	0.2789	0.2874	0.0604	0.0237	0.4809	0.5826	0.5865	<b>0.6116</b>
t(3)	0.0936	0.0836	0.1289	0.1321	0.1328	<b>0.1390</b>	0.0559	0.046	0.057	0.0679	0.0621	<b>0.0721</b>
Logc(0,1)	0.1339	0.1275	<b>0.1424</b>	0.0846	0.0916	0.0987	0.0464	0.0388	0.0890	0.1494	0.1674	<b>0.1833</b>
Lap(0,1)	<b>0.0798</b>	0.0618	0.0663	0.0378	0.0355	0.0357	0.0436	0.0299	0.0340	<b>0.0482</b>	0.0407	0.0462
Group IV												
Gum(2)	<b>0.1159</b>	0.0385	0.0542	0.0121	0.0062	0.0057	<b>0.1497</b>	0.0660	0.0490	0.0176	0.0071	0.0074
Gum(0.5)	0.2592	<b>0.2773</b>	0.2705	0.0745	0.0808	0.0782	0.1781	<b>0.2167</b>	0.1826	0.1140	0.1332	0.1294
Log(0.5)	<b>0.6319</b>	0.5843	0.6182	0.5940	0.5813	0.5847	<b>0.6292</b>	0.6135	0.6158	0.6123	0.6086	0.6096
Log(2)	0.2248	<b>0.2319</b>	0.2260	0.0670	0.0659	0.0701	0.08	0.0796	0.0899	0.0754	0.0910	<b>0.1001</b>

Table 4: Power comparisons of  $\alpha = 0.10$  and  $F(C) = 0.7$  for normal distribution versus alternatives in Table 1

	$n = 10$						$n = 20$					
	$LW_B^{C1}$	$LW_V^{C1}$	$KL_V^{C1}$	KS	CM	AD	$LW_B^{C1}$	$LW_V^{C1}$	$KL_V^{C1}$	KS	CM	AD
Group I												
U(0,1)	<b>0.2432</b>	0.1772	0.1893	0.0897	0.0667	0.0695	<b>0.3903</b>	0.2643	0.2607	0.1200	0.0963	0.1262
B(2,2)	<b>0.1737</b>	0.1111	0.1169	0.0599	0.0457	0.0442	<b>0.2211</b>	0.1372	0.1138	0.0638	0.0406	0.0469
B(2,0.5)	<b>0.1337</b>	0.0572	0.0724	0.0840	0.0001	0.0011	<b>0.2528</b>	0.0090	0.0338	0.0002	0.0000	0.0000
B(0.5,0.5)	<b>0.4099</b>	0.3309	0.3834	0.1800	0.148	0.1780	<b>0.7174</b>	0.5563	0.6302	0.3056	0.2947	0.3920
Group II												
G(2)	<b>0.1771</b>	0.0752	0.0986	0.0110	0.0000	0.0010	<b>0.3098</b>	0.0136	0.0458	0.0002	0.0000	0.0000
G(0.5)	<b>0.4424</b>	0.4114	0.4422	0.2460	0.2312	0.2540	<b>0.7623</b>	0.6792	0.7162	0.4220	0.4291	0.5209
LN(0,0.5)	<b>0.1415</b>	0.0586	0.0766	0.0096	0.0000	0.0013	<b>0.2557</b>	0.0098	0.0342	0.0000	0.0000	0.0000
LN(0,1)	<b>0.2368</b>	0.1152	0.1431	0.232	0.0360	0.0054	<b>0.3987</b>	0.0754	0.0993	0.0710	0.0160	0.0041
Group III												
SC	0.0652	0.0417	0.2387	<b>0.3593</b>	0.3427	0.3573	0.0678	0.0169	0.4924	0.6257	0.6241	<b>0.6515</b>
t(3)	0.0807	0.0673	0.1297	0.1897	0.1856	<b>0.1970</b>	0.0484	0.0291	0.2058	0.3312	0.3332	<b>0.3646</b>
Log(0,1)	0.0967	0.0870	0.1223	0.1537	0.1615	<b>0.1684</b>	0.0629	0.0420	0.1397	0.2320	0.2494	<b>0.2682</b>
Lap(0,1)	<b>0.1009</b>	0.0743	0.0877	0.0856	0.0772	0.0813	0.0885	0.0576	0.0694	<b>0.0929</b>	0.0818	0.0921
Group IV												
Gum(2)	<b>0.1752</b>	0.0874	0.1014	0.0371	0.0202	0.0209	<b>0.2210</b>	0.1003	0.0818	0.0347	0.0181	0.0219
Gum(0.5)	0.1965	<b>0.2196</b>	0.2088	0.1249	0.1294	0.1239	0.1997	<b>0.2473</b>	0.2026	0.1559	0.1672	0.1560
Log(0.5)	<b>0.6566</b>	0.6204	0.6419	0.6255	0.6088	0.6164	<b>0.6592</b>	0.6284	0.6373	0.6288	0.6149	0.6166
Log(2)	0.1222	0.1443	<b>0.1456</b>	0.1205	0.1327	0.1364	0.0753	0.0870	0.1072	0.1289	0.1477	<b>0.1530</b>

### 3.1 Illustrative real examples

A real data set is analyzed in this section. We present the proposed methods in application in order to analyze the results. Example 3.1 represents the times of breakdown of insulation fluid samples (in minutes) tested at 34 kVs ([8], Table 1.1, page 105).

**Example 3.1.** Suppose that a decision had been made to terminate the life-testing after 12.06 minute. So 5 out of 19 observations would have been censored while 14 data points remained for study (see Table 6). Note that 12.06, datum is equivalent to the quantile

Table 5: Power comparisons of  $\alpha = 0.10$  and  $F(C) = 0.8$  for normal distribution versus alternatives in Table 1

	$n = 10$						$n = 20$					
	$LW_B^{C1}$	$LW_V^{C1}$	$KL_V^{C1}$	KS	CM	AD	$LW_B^{C1}$	$LW_V^{C1}$	$KL_V^{C1}$	KS	CM	AD
Group I												
U(0,1)	0.2821	0.3055	<b>0.3133</b>	0.1454	0.1698	0.1808	0.3224	0.2763	<b>0.3393</b>	0.2089	0.2545	0.2905
B(2,2)	0.5501	0.5674	<b>0.5684</b>	0.1448	0.1769	0.2070	0.3753	0.3441	<b>0.3836</b>	0.1580	0.2111	0.2492
B(2,0.5)	<b>0.1125</b>	0.0304	0.0392	0.0176	0.0072	0.0077	<b>0.1112</b>	0.0188	0.0186	0.0156	0.0045	0.0064
B(0.5,0.5)	0.3235	0.3040	<b>0.3506</b>	0.2326	0.2502	0.2698	<b>0.6474</b>	0.5593	0.6548	0.3810	0.4381	0.5155
Group II												
G(2)	<b>0.2244</b>	0.0592	0.0853	0.0060	0.0002	0.0002	<b>0.3631</b>	0.0260	0.0512	0.0013	0.0000	0.0003
G(0.5)	0.5428	0.5295	<b>0.5645</b>	0.3350	0.3593	0.3691	0.8597	0.8443	<b>0.8600</b>	0.5915	0.6593	0.7238
LN(0,0.5)	<b>0.2241</b>	0.0712	0.0963	0.0138	0.0029	0.0026	<b>0.3001</b>	0.0756	0.0632	0.0140	0.0035	0.0050
LN(0,1)	<b>0.3497</b>	0.1563	0.1868	0.0490	0.0223	0.0205	<b>0.4718</b>	0.2310	0.2051	0.0729	0.0298	0.0440
Group III												
SC	0.1094	0.0743	0.2732	<b>0.3753</b>	0.3525	0.3605	0.1078	0.0393	0.5219	0.6444	0.6344	<b>0.6586</b>
t(3)	0.1342	0.1122	0.1728	0.2126	0.2065	<b>0.2145</b>	0.0642	0.0459	0.2279	0.3328	0.3368	<b>0.3614</b>
Log(0,1)	0.1690	0.1563	<b>0.1775</b>	0.1541	0.1571	0.1631	0.0788	0.0685	0.1470	0.2292	0.2400	<b>0.2554</b>
Lap(0,1)	<b>0.1348</b>	0.0932	0.0995	0.0773	0.0667	0.0692	0.0874	0.0611	0.0729	<b>0.096</b>	0.0773	0.0830
Group IV												
Gum(2)	<b>0.2048</b>	0.0847	0.0946	0.0363	0.0207	0.0190	<b>0.2387</b>	0.1143	0.0882	0.0462	0.0170	0.0194
Gum(0.5)	0.3237	<b>0.3568</b>	0.3381	0.1537	0.1663	0.1650	0.2759	<b>0.3298</b>	0.2834	0.2114	0.2327	0.2341
Log(0.5)	<b>0.6686</b>	0.6156	0.6424	0.6199	0.5993	0.6026	<b>0.6619</b>	0.6325	0.6330	0.6299	0.6194	0.6206
Log(2)	0.2611	0.2756	<b>0.2817</b>	0.1316	0.1367	0.1422	0.1126	0.1202	0.1291	0.1449	0.1589	<b>0.1678</b>

0.7 in the sample set and censored observations are marked by an asterisk. In this case,  $n = 19$ ,  $n_0 = 14$  and  $C = 12.06$  (equivalently the  $F(12.06) = 0.7$ ). We are interested to assess whether the times to breakdown follow a normal distribution. Table 7 shows the values of the test statistics and the corresponding p-values for the proposed tests on Type-I censored data.

Table 6: Times to breakdown in minutes of an insulating fluid at 34 kV voltage level in Example 3.1.

0.19	0.78	0.96	1.31	2.78	3.16	4.15	4.67	4.85	6.5
7.35	8.01	8.27	12.06	<del>31.75</del>	<del>32.52</del>	<del>33.91</del>	<del>36.71</del>	<del>72.89</del>	

Table 7: Test statistics values and the p-values for normal distribution in Example 3.1.

The null hypothesis of normality						
Test	$LW_B^{C1}$	$LW_V^{C1}$	$KL_V^{C1}$	KS	CM	AD
Value (Example 1)	0.0353	0.1112	0.3344	0.2820	1.0655	5.667
P-value	0.0048	0.0006	0.0018	0.0428	0.0047	0.0063

As it can be seen in Table 7, the p-values obtained based on above mentioned tests are in support of the null hypothesis of normal distribution. The values of the statistics and corresponding p-values in Table 7 are quite large, strongly reject the alternative hypothesis. The  $LW_V^{C1}$  is stronger than  $LW_B^{C1}$  through the corresponding p-values.

## 4 Conclusions

In this paper, we have developed some GoF test for the normal distribution based on Type-I censored samples. The LW divergence measure which is proposed as the preferred

information measure in this study can be evaluated using the comparison of various distance measures such as AD, CM, KS and KL.

We have selected the alternatives from four different families based on random variables. The test statistics representing the capability of the tests for measuring the distance between the subject and reference distributions are compared here. Results suggest that GoF test for normal distribution would have more powerful when the distance is evaluated using the LW divergence measure. So that the LW as an estimator can lead to better results in measuring the magnitude of difference between two distribution patterns with higher precision. Therefore, we recommend the LW measure to be used with Type-I censored data.

## References

- [1] Alizadeh Noughabi, H. (2010). A new estimator of entropy and its application in testing normality. *Journal of Statistical Computation and Simulation*. 80(10): 1151-62.
- [2] Bitaraf, M., Rezaei, M., Yousefzadeh, F. (2017). Test for normality based on two new estimators of entropy. *Journal of Statistical Computation and Simulation*. 87(2): 280-294.
- [3] Deza, M.M., Deza, E. (2009). *Encyclopedia of distance*. Springer Berlin Heidelberg.
- [4] Esteban M, Castellanos M, Morales D, Vajda I. (2001). Monte Carlo comparison of four normality tests using different entropy estimates. *Communication in Statistics-Simulation and Computation*. 30(4): 761-785.
- [5] Koziol, J. A., Byar, D. P. (1975). Percentage points of the asymptotic distributions of one and two sample KS statistics for truncated or censored data. *Technometrics*. 17(4): 507-510.
- [6] Kullback, S., Leibler, R.A. (1951). On information and sufficiency. *The annals of mathematical statistics*. 22(1):79-86.
- [7] Lin, J., Wong, S.K.M. (1990). A New directed divergence measure and its characterization. *International Journal Of General System*. 17(1):73-81.
- [8] Nelson, W. (1982). *Applied Life Data Analysis*. JohnWiley and Sons, NewYork.
- [9] Pettitt, A. N., Stephens, M. A. (1976). Modified Cramer-von Mises statistics for censored data. *Biometrika*. 63(2): 291-298.
- [10] Vasicek, O. (1976). A test for normality based on sample entropy. *Journal of Research Statistical Society*. 38 (Serial. B): 54-59.



# Bayes estimation of mean lifetime in a proportional hazard rate model under progressive censoring

Razmkhah, M.<sup>1</sup> Khatib, B.<sup>2</sup>

<sup>1</sup> Department of Statistics, Ferdowsi University of Mashhad

<sup>2</sup> Department of Mathematics and Statistics, University of Neyshabur

## Abstract

Suppose that the lifetimes of some units are independent but not identically distributed random variables, such that they come from a proportional hazard rate model; moreover, suppose that units are placed under progressively Type-II censoring scheme. In this paper, one-parameter exponential distribution is considered as the baseline distribution; the Bayes estimation of the mean time to failure for the baseline distribution is derived under a Balanced loss function. Toward this end, a conjugate prior is considered and the corresponding posterior distribution is determined. The performance of the proposed procedure is investigated via a real data set.

**Keywords:** Balanced loss function, Conjugate prior, Maximum likelihood estimation, Progressive censoring scheme.

## 1 Introduction

Let  $X_1, \dots, X_n$  be the lifetimes of  $n$  independent experimental units, such that  $X_i$ ,  $1 \leq i \leq n$ , has the survival function

$$\bar{F}_i(x; \boldsymbol{\theta}) = [\bar{G}(x; \boldsymbol{\theta})]^{\lambda_i}, \quad (1)$$

<sup>1</sup>razmkhah\_m@um.ac.ir

<sup>2</sup>khatib\_b@neyshabur.ac.ir

where  $\boldsymbol{\theta}' = (\theta_1, \dots, \theta_t)$  is a common vector of parameters;  $\lambda_1, \dots, \lambda_n$  are known positive constants and  $\bar{G} = 1 - G$  is the survival function of the baseline distribution. This family is well-known as proportional hazard rate model (PHRM) with proportionality rates  $\lambda_1, \dots, \lambda_n$ ; see, for example [9]. This family includes several well-known lifetime distributions such as exponential, Weibull, Pareto, Burr Type-XII and so on. For example, suppose that  $n$  units which have been made by different companies and the corresponding lifetimes come from different distributions are independently placed on a life test. Also, suppose that there exists a hierarchical relation among these distributions through relation (1). If one consider the first distribution as the baseline distribution, that is  $\lambda_1 = 1$ , then  $\lambda_2, \dots, \lambda_n$  represent the proportionality rates of the other distributions with respect to the first (baseline) distribution.

Censored sampling arises in a life test whenever the experimenter does not observe the lifetimes of all experimental units. The model of progressive Type-II censoring is of importance in the field of reliability and life testing. In this censoring scheme,  $n$  units are simultaneously placed on a lifetime test and when the  $i$ th failure time occurs,  $R_i$  surviving units are randomly censored from the experiment,  $1 \leq i \leq m$ . Thus, if  $m$  failure times are observed, then  $R_1 + \dots + R_m$  units are censored; here,  $\mathbf{R} = (R_1, \dots, R_m)$  denotes the progressive censoring plan. In the special case of  $R_1 = \dots = R_{m-1} = 0$  and  $R_m = n - m$ , the progressive censoring scheme coincides with the Type-II censoring scheme. Statistical inferences based on progressively Type-II censored order statistics in the case of independent and identically distributed (IID) random variables have been extensively investigated by several authors. Balakrishnan et al. [4] studied the point and interval estimation for both location and scale parameters of the two-parameter exponential distribution based on progressively Type-II censored samples. Burkschat et al. [5] investigated the optimal plans in the model of progressive Type-II censoring for a location-scale family of distributions. For more details in this direction, see [1, 3]. This censoring scheme was generalized by [2] to the case of independent and non-identical distributed (INID) random variables. They developed the basic distribution theory for order statistics in this case. Also, Fischer et al. [7] studied a mixture representation for the joint distribution function of progressively Type-II censored order statistics from heterogeneous distributions and illustrated the applications of this representation to stochastic orderings and inequalities. Cramer and Lenz [6] and Mao and Hu [10] investigated the positive association and the stochastic properties of these statistics, respectively. Recently, Rezapour et al. [13] investigated some more properties of progressively Type-II censored order statistics in the INID case. Razmkhah and Simriz [12] studied statistical inferences based on INID progressively Type-II censored order statistics.

Here, the Bayesian point estimation of the unknown parameter  $\theta$  is concerned. With this in mind, the following Balanced loss function is used:

$$L_{\rho, w, T_0}(\theta, T) = w\rho(T_0, T) + (1 - w)\rho(\theta, T), \quad (2)$$

where  $\rho(\theta, T)$  is an arbitrary loss function, while  $T_0$  is a chosen a priori target estimator of the interested parameter, such as MLE. So, the mentioned loss function reflects a desire of closeness of the Bayes estimator  $T$  to both of the target estimator  $T_0$  and the unknown parameter  $\theta$  with the relative importance of these criteria governed by the choice of  $w \in [0, 1]$ . For more details, see [8].

## 2 Preliminaries and model description

Let  $X_1, \dots, X_n$  be the lifetimes of  $n$  units which are independently and simultaneously placed on a test for which  $X_r$  comes from cumulative distribution function (cdf)  $F_r(x; \boldsymbol{\theta})$  with corresponding pdf  $f_r(x; \boldsymbol{\theta})$ ,  $1 \leq r \leq n$ , where  $\boldsymbol{\theta}' = (\theta_1, \dots, \theta_t)$  is the common vector of parameters of the various distributions. Moreover, let  $\mathbf{R} = (R_1, \dots, R_m)$  be the progressive censoring plan with  $n = m + \sum_{i=1}^m R_i$ . For brevity, we denote by  $\gamma_j = n - \sum_{i=1}^{j-1} R_i - j + 1$  the number of units remaining in the experiment before the  $j$ th failure time.

Now, let us first define the following random variable to identify the failed units on a lifetime test

$$\Delta_i^{(j)} = \begin{cases} 1, & \text{if the lifetime of the } i\text{th unit coincides with the } j\text{th failure time,} \\ 0, & \text{otherwise,} \end{cases}$$

where for each  $j = 1, \dots, m$ ,  $\sum_{i=1}^n \Delta_i^{(j)} = 1$ . Also, after the  $j$ th failure time,  $R_j$  of surviving units are removed from the experiment. Therefore, we use the following random variable to specify the units removed from the test

$$H_i^{(j)} = \begin{cases} 1, & \text{if the } i\text{th unit is removed from the test after the } j\text{th failure time,} \\ 0, & \text{otherwise,} \end{cases}$$

such that for each  $j = 1, \dots, m$ ,  $\sum_{i=1}^n H_i^{(j)} = R_j$ . Notice that for a fixed  $j$ , say  $j_0$ , if  $\Delta_i^{(j_0)} = 1$  or  $H_i^{(j_0)} = 1$ , then for other values of  $j$  ( $j \neq j_0$ ),  $\Delta_i^{(j)} = 0$  and  $H_i^{(j)} = 0$ .

According to the above random variables, it is reasonable to use the following random vector to make inferences about the parameters of interest

$$\mathbf{X} = \{X_{j:m:n}^{\mathbf{R}}, \Delta_i^{(j)}, H_i^{(j)}, 1 \leq j \leq m, 1 \leq i \leq n\}.$$

Razmkhah and Simriz [12] showed that given  $\boldsymbol{\theta}$ , the joint pdf of  $\mathbf{x}$ , the observed value of  $\mathbf{X}$  in (3), can be presented as

$$f(\mathbf{x}|\boldsymbol{\theta}) = \left( \prod_{j=1}^m \gamma_j \right) \prod_{j=1}^m \left\{ \prod_{i=1}^n [f_i(x_j; \boldsymbol{\theta})]^{\delta_i^{(j)}} [\bar{F}_i(x_j; \boldsymbol{\theta})]^{\eta_i^{(j)}} \right\}, \quad (3)$$

where  $x_j$ ,  $\delta_i^{(j)}$  and  $\eta_i^{(j)}$  are the observed values of  $X_{j:m:n}^{\mathbf{R}}$ ,  $\Delta_i^{(j)}$  and  $H_i^{(j)}$ , respectively. Notice that the constant  $\gamma_j$  represents the number of ways in which the  $j$ th progressively Type-II censored order statistic may occur.

In the next section, the Bayes estimation of the interested parameter under the Balanced loss function (2) is investigated.

## 3 Main results

Let lifetimes  $X_1, \dots, X_n$  be independent random variables for which  $X_i$  ( $1 \leq i \leq n$ ) comes from the model (1) with the baseline one-parameter exponential distribution with the cdf

$$G(x; \theta) = 1 - e^{-\theta x}, \quad (4)$$

where  $\theta$  is a real valued parameter. Therefore, (3) may be rewritten as

$$f(\mathbf{x}|\boldsymbol{\theta}) = \left( \prod_{j=1}^m \gamma_j \right) \theta^m e^{-\theta \sum_{j=1}^m \sum_{i=1}^n \lambda_i (\delta_i^{(j)} + \eta_i^{(j)}) D(x_j)}. \quad (5)$$

Note that the mean lifetime of the baseline distribution is  $\theta^{-1}$ . It is not difficult to show that the MLE of  $\theta^{-1}$  is

$$T_0(\mathbf{x}) = \frac{1}{m} \sum_{j=1}^m \sum_{i=1}^n \lambda_i (\Delta_i^{(j)} + H_i^{(j)}) D(X_{j:m:n}^{\mathbf{R}}). \quad (6)$$

On the other hand, to obtain the Bayes estimator, let us consider the conjugate prior distribution. Toward this end, we assume  $\theta$  obeys  $\Gamma(\alpha, \beta)$  distribution with the following pdf

$$\pi(\theta) = \frac{1}{\Gamma(\alpha)} \beta^\alpha \theta^{\alpha-1} e^{\beta\theta}, \quad \theta > 0. \quad (7)$$

Since  $\sum_{j=1}^m \sum_{i=1}^n \delta_i^{(j)} = m$ , using (3) and (7), the posterior density function of  $\theta$  is derived as

$$\pi(\theta|\mathbf{x}) \propto \theta^{\alpha+m-1} e^{-\theta(\beta + \sum_{j=1}^m \sum_{i=1}^n \lambda_i (\delta_i^{(j)} + \eta_i^{(j)}) D(x_j))}, \quad \theta > 0. \quad (8)$$

Assuming squared error loss (SEL) function  $\rho(\theta, T) = (T - \theta)^2$ , the Bayes estimator for the mean lifetime  $\frac{1}{\theta}$  is given by

$$T_\pi(\mathbf{X}) = \frac{1}{\alpha + m - 1} \left( \beta + \sum_{j=1}^m \sum_{i=1}^n \lambda_i (\Delta_i^{(j)} + H_i^{(j)}) D(X_{j:m:n}^{\mathbf{R}}) \right). \quad (9)$$

Hence, the Bayes estimator under the Balanced loss function in (2) is given by

$$T = w T_0(\mathbf{X}) + (1 - w) T_\pi(\mathbf{X}), \quad (10)$$

where  $T_0(\mathbf{X})$  and  $T_\pi(\mathbf{X})$  are as defined in (6) and (9), respectively.

## 4 Real example

To illustrate the performance of the proposed procedure in this paper, we use a real data set which consists of the time (in hours) of successive failures of the air conditioning system in ten Boeing 720 jet aircrafts; see, [11] for a detailed description of the data set. He tested and accepted the hypothesis that the successive failure times are iid exponential for each aircraft, but with different failure rates. Therefore, we assume that the corresponding failure times for the  $i$ th aircraft come from the cdf  $F_i(x; \sigma) = 1 - e^{-\lambda_i \theta x}$ , which coincides with a proportional hazard rate family in (1). Since in the assumptions of our model,  $\lambda_i$ 's are known parameters, we consider some arbitrary values for  $\lambda_1, \dots, \lambda_{10}$  as presented in Table 1. Moreover, three observations related to the  $i$ th aircraft, denoted by  $X_{3i-2}, X_{3i-1}$  and  $X_{3i}$  ( $1 \leq i \leq 10$ ), have the same distribution as  $F_i(x, \sigma)$ . In fact, we use a sample of

size thirty of failure times for which the first three of them come from cdf  $F_1(x; \theta)$ , the second three of them come from cdf  $F_2(x; \theta)$  and eventually the last three of them come from cdf  $F_{10}(x; \theta)$ . Notice that the common parameter in these distributions is  $\theta$  which is of interest; moreover, there exist ten known parameters  $\lambda_1, \dots, \lambda_{10}$  which construct the different distributions related together via relation in (1). Summary descriptions are reported in Table 1.

**Table 1.** Summary descriptions of Boeing 720 jet aircraft data.

$i$	1	2	3	4	5	6	7	8	9	10
$X_{3i-2}$	194	413	90	74	55	23	97	50	359	50
$X_{3i-1}$	15	14	10	57	320	261	51	44	9	254
$X_{3i}$	41	58	60	48	56	87	11	102	12	5
$\lambda_i$	1	0.9017	0.9836	0.6803	0.6230	1.3770	1.0656	1.2787	0.4098	0.7705

Using the data in Table 1 and by using the progressive censoring plan  $\mathbf{R}^* = (5, 5, 5, 5, 5)$ , the first five progressively Type-II censored order statistics have been extracted. The results are tabulated in Table 2. From the entries of this table, the values of  $\Delta_i^{(j)}$  and  $H_i^{(j)}$  in the data set  $\mathbf{B}$  may also be specified.

**Table 2.** Progressively Type-II censored order statistics extracted from the data in Table 1.

$i$	$X_{i:m:n}^{\mathbf{R}^*}$	$x_i$	Censored units after the $i$ th failure time
1	$X_{30}$	5	$X_2 \ X_{10} \ X_{12} \ X_{23} \ X_{26}$
2	$X_8$	10	$X_5 \ X_6 \ X_{15} \ X_{18} \ X_{22}$
3	$X_{21}$	11	$X_3 \ X_9 \ X_{16} \ X_{27} \ X_{28}$
4	$X_{20}$	51	$X_{11} \ X_{13} \ X_{19} \ X_{25} \ X_{29}$
5	$X_7$	90	$X_1 \ X_4 \ X_{14} \ X_{17} \ X_{24}$

Using (6) and the data in Table 2, the observed value of the MLE of  $\theta^{-1}$  on the basis of the data set  $\mathbf{X}$  is given by

$$\begin{aligned}
T_0(\mathbf{x}) &= \frac{1}{5} \left\{ x_1 (\lambda_2 + \lambda_{10} + \lambda_{12} + \lambda_{23} + \lambda_{26} + \lambda_{30}) \right. \\
&\quad + x_2 (\lambda_5 + \lambda_6 + \lambda_8 + \lambda_{15} + \lambda_{18} + \lambda_{22}) \\
&\quad + x_3 (\lambda_3 + \lambda_9 + \lambda_{16} + \lambda_{21} + \lambda_{27} + \lambda_{28}) \\
&\quad + x_4 (\lambda_{11} + \lambda_{13} + \lambda_{19} + \lambda_{20} + \lambda_{25} + \lambda_{29}) \\
&\quad \left. + x_5 (\lambda_1 + \lambda_4 + \lambda_7 + \lambda_{14} + \lambda_{17} + \lambda_{24}) \right\} \\
&= 187.3066. \tag{11}
\end{aligned}$$

Moreover, using (9) and (10), the Bayes estimator is

$$T = w \times 187.3066 + (1 - w) \times \frac{\beta + 936.533}{\alpha + 4}.$$

Taking the known parameters of prior distribution as  $\alpha = 1.8$  and  $\beta = 0.01$ , we get

$$T = w \times 187.3066 + (1 - w) \times 161.4729,$$

where  $w$  is an arbitrary real valued in  $[0, 1]$ , which may control the role of classic and Bayes estimation. In fact, By the above scenario, any real value in the interval  $[161.4729, 187.3066]$  is a Bayes estimation for the baseline mean time to failure of the air conditioning system in the Boeing 720 jet aircrafts.

## 5 Conclusion

The problem of Bayes estimation was investigated under progressive censoring scheme in the case of INID random lifetimes. To model such lifetimes, we use a proportional hazard rate model with known proportionality rates. Moreover, it was assumed that the baseline distribution is exponential with unknown parameter  $\theta$ . In spite of MLE, the Bayes estimator was obtained under a Balanced loss function and conjugate prior distribution. The results may extended to the case of other baseline distribution for the PHRM as well as other loss functions or other prior distributions. Also, a sensitivity analysis may be performed to study the effect of varying prior parameters and various censoring schemes  $\mathbf{R} = (R_1, \dots, R_m)$ .

## References

- [1] Balakrishnan, N. and Aggarwala, R. (2000), *Progressive Censoring: Theory, Methods and Applications*, Birkhauser, Boston.
- [2] Balakrishnan, N. and Cramer, E. (2008), Progressive censoring from heterogeneous distributions with applications to robustness, *Annals of the Institute of Statistical Mathematics*, **60**, 151–171.
- [3] Balakrishnan, N. and Cramer, E. (2014), *The Art of Progressive Censoring: Applications to Reliability and Quality*, Birkhauser, New York.
- [4] Balakrishnan, N., Cramer, E., Kamps, U. and Schenk, N. (2001), Progressive Type-II censored order statistics from exponential distributions, *Statistics*, **35**, 537–556.
- [5] Burkschat, M., Cramer, E. and Kamps, U. (2006), On optimal schemes in progressive censoring, *Statistics & Probability Letters*, **76**, 1032–1036.
- [6] Cramer, E. and Lenz, U. (2010), Association of progressively Type-II censored order statistics, *Journal of Statistical Planning and Inference*, **140**, 576–583.
- [7] Fischer, T., Balakrishnan, N. and Cramer, E. (2008), Mixture representation for order statistics from INID progressively censoring and its applications, *Journal of Multivariate Analysis*, **99**, 1999–2015.
- [8] Jafari Jozani, M., Marchand, E. and Parsian, A. (2012), Bayesian and Robust Bayesian analysis under a general class of balanced loss functions, *Statistical Papers*, **53**, 51–60.
- [9] Lawless, J. L. (2003), *Statistical Models and Methods for Lifetime Data*, 2nd ed. John Wiley, New York.

- [10] Mao, T. and Hu, T. (2010), Stochastic properties of INID progressive Type-II censored order statistics, *Journal of Multivariate Analysis*, **101**, 1493–1500.
- [11] Proschan, F. (1963), Theoretical explanation of observed decreasing failure rate, *Technometrics*, **5**, 375–383.
- [12] Razmkhah, M. and Simriz, S. (2017), Statistical inferences based on INID progressively Type-II censored order statistics, *Annals of the Institute of Statistical Mathematics*, DOI:10.1007/s10463-017-0598-9.
- [13] Rezapour, M., Alamatsaz, M.H., Balakrishnan, N. and Cramer, E. (2013), On properties of progressively Type-II censored order statistics arising from dependent and non-identical random variables *Statistical Methodology*, **10**, 58–71.



## Application of skip-free markov chains for degradation

Sajadi, F.A.<sup>1</sup>

Department of Statistics, University of Isfahan

### Abstract

In this paper we review some applications of stochastic processes in reliability theory. In particular, we focus attention on the so called skip-free Markov chains on  $\mathbb{N}$  for which upward jumps may be only of unit size, and there is no restriction on downward jumps. These processes can be used as models for the time-evolution of degradation. Under some sufficient conditions, there exists the usual stochastic ordering between hitting times of two skip-free Markov chains with the same state-space. This ordering leads to an application in the frame of degradation models.

**Keywords:** Skip-free Markov chains , degradation, stochastic ordering.

## 1 Introduction

Markov chains provide interesting stochastic models of various natural/artificial phenomena. In this paper, we consider a special class of finite-space Markov chains in discrete time, for which upward jumps may be only of unit size, and there is no restriction on downward jumps. Then we will discuss, how these kind of chains, can be interpreted in terms of evolution in discrete time of levels of degradation for appropriate reliability structures.

**Definition 1.1.** A Markov chain  $(X_n)_{n \geq 0}$ , with the state space  $E_k = \{0, 1, \dots, k\}$  is called skip-free to the right if its transition probabilities satisfy the following conditions:

<sup>1</sup>f.sajadi@sci.ui.ac.ir

$$\begin{aligned} p_{i,i+1} &> 0 & \text{for } i = 0, 1, \dots, k-1, \\ p_{i,j} &= 0 & \text{for } i = 0, 1, \dots, k-2, j \geq i+2. \end{aligned} \quad (1)$$

where  $p_{i,j} := P(X_n = j | X_{n-1} = i)$ .

Assuming that 0 is the initial state of the chain and  $k$  is an absorbing state. We denote by  $T_h$  the stopping times

$$T_h := \inf \{n \in \mathbb{N} : X_n \geq h\}, h = 1, 2, \dots, k.$$

$T_h$  is thus the waiting time until the chain reaches or exceeds the level  $h$ . If the Markov chain starts from the state zero,  $T_h$  is the hitting time of the state  $h$ . For results concerning the probability distributions of  $T_h$ , under different assumptions, we refer to [1], [2],[3] and references cited therein. Besides the computation of the probability distribution of  $T_h$ , typical problems related with the analysis of  $T_h$  are the computation of the expected value  $\mathbb{E}(T_h | X_0 = 0)$  and stochastic comparisons between  $T_h$  and  $\bar{T}_h$  for two different chains  $(X_n)_{n \geq 0}$  and  $(\bar{X}_n)_{n \geq 0}$  belonging to the class of skip-free Markov chains, [4],[5]. For an irreducible skip-free Markov chain on the non-negative integers with absorbing state  $k$ , under some conditions, the hitting (absorbing) time of state  $k$ , when departing from state 0 is distributed as the sum of  $k$  independent geometric (or exponential) random variables. For the skip-free chain, authors in [6], first proved the result in the continuous time situation. By using the duality, a stochastic proof to both discrete and continuous time cases, is given in [3]. Among different notions of stochastic orders that might be considered for the  $\mathbb{N}$ -valued random variables  $T_h$  and  $\bar{T}_h$ , in this paper we use the usual stochastic order  $T_h \preceq_{st} \bar{T}_h$ , that is defined in terms of tail behavior of the distributions , [8]. More precisely, if  $Y$  and  $\bar{Y}$  are discrete random variables taking on values in  $\mathbb{N}$ , then we have the following. Let  $p_i = \mathbb{P}(Y = i)$  and  $q_i = \mathbb{P}(\bar{Y} = i)$ ,  $i \in \mathbb{N}$ . We say  $Y \preceq_{st} \bar{Y}$  if, and only if,

$$\sum_{j=i}^{\infty} p_j \leq \sum_{j=i}^{\infty} q_j, i \in \mathbb{N}.$$

Suppose a device  $D$  starts working at time 0, and some physical quality of  $D$  that can be seen as a measure of degradation is checked at any discrete time  $n$ ,  $n = 1, 2, \dots$ . Let  $X_n$  be the age of device  $D$  at time  $n$ . Assume that the increase of age is equal to one unit. Without taking any maintenance action, we have  $X_n = n$ , that is, the age of  $D$  at time  $n$  is equal to  $n$ . Now, suppose in order to reduce the value of degradation and to make  $D$  as good as new, a maintenance policy is implemented over  $D$ , at times  $n = 1, 2, \dots$ . Define the reduction of age at time  $n$  by a non-negative random variable,  $R_n$ . Therefore at time  $n = 1$  we have  $X_1 = X_0 + 1 - R_1$ , at time  $n = 2$ ,  $X_2 = 2 - (R_1 + R_2)$  and so on. Hence by setting  $X_0 = 0$ , we have the following recursion formula for the age of device  $D$  at time  $n$ :

$$X_n = X_{n-1} + 1 - R_n, n = 1, 2, \dots$$

Observe that, conditionally on  $X_{n-1} = h - 1$ , random variable  $R_n$  is a  $\{0, 1, \dots, h\}$ -valued random variable with a probability distribution  $Q_h$  which only depending on the value  $h$ . Therefore we obtain that  $(X_n)_{n \geq 0}$  is a time-homogeneous, skip-free, Markov chain, whose transition matrix is determined by the distributions  $Q_1, Q_2, \dots, Q_k$ .

In section 2, first we state the Theorem 2 in [5] which is an extension of Theorem 4.1 in [7]. These theorems, give some sufficient conditions for the stochastic ordering between hitting times, in a fixed state, for two Markov chains. Then to see an application in the frame of degradation models, through an example, we show that how to decide what is preferable between two different repair policies.

## 2 Stochastic comparisons between $T_h$ and $\acute{T}_h$

Let  $C_k$  be the class of transition matrices on the state space  $E_k$  satisfying (1). Consider two Markov chains  $(X_n)_{n \geq 0}$  and  $(\acute{X}_n)_{n \geq 0}$  in  $C_k$ , with transition matrices  $P = (p_{i,j})_{i,j \in E_k}$ ,  $\acute{P} = (\acute{p}_{i,j})_{i,j \in E_k}$ , and initial distributions  $\pi_0 = (\pi_0(i))_{i \in E_k}$ ,  $\acute{\pi}_0 = (\acute{\pi}_0(i))_{i \in E_k}$ , respectively. We furthermore consider  $T_h$  and  $\acute{T}_h$  where  $T_h$  is defined for  $(X_n)_{n \geq 0}$  in (1) and  $\acute{T}_h$  is the analogue for  $(\acute{X}_n)_{n \geq 0}$ . For a transition matrix  $P$ , we will use the notation  $p_{i,.}^{(n)} = (p_{i,0}^{(n)}, p_{i,1}^{(n)}, \dots, p_{i,k}^{(n)})$ , where  $p_{i,j}^{(n)}$  denotes the transition probability from  $i$  to  $j$  in  $n$  steps.

**Theorem 2.1.** *Let  $P = (p_{i,j})_{i,j \in E_k}$ ,  $\acute{P} = (\acute{p}_{i,j})_{i,j \in E_k}$  be two transition matrices in  $C_k$ .*

- Under the conditions  $\acute{p}_{i,.} \preceq_{st} p_{i,.}$ , for each  $0 \leq i \leq k-1$ ,  $\acute{\pi}_0 \preceq_{st} \pi_0$ , one has the stochastic comparison

$$T_h \preceq_{st} \acute{T}_h,$$

for  $h = 1, \dots, k$ , (Theorem 4.1, [7]).

- Assume that, for any  $i = 0, 1, \dots, k-1$ , there exists  $m(i)$  such that (a)  $i+m(i) \leq k$ ; and (b)  $\acute{p}_{i,.}^{(m(i))} \preceq_{st} p_{i,.}^{(m(i))}$ . Moreover suppose that the initial measures are stochastically ordered  $\acute{\pi}_0 \preceq_{st} \pi_0$ . Then

$$T_k \preceq_{st} \acute{T}_k, \quad (2)$$

(Theorem 2, [5]).

Note that in Theorem 4.1 in [7], the case where  $m(i) = 1$  for  $i = 1, 2, \dots, k-1$  is considered. Also, the result there, permits to compare  $T_h$  with  $\acute{T}_h$ , for any  $0 \leq h \leq k$ .

**Example 2.1.** Let  $E_3 = \{0, 1, 2, 3\}$  and  $X_n$  be the age of some device at time  $n$ ,  $n = 0, 1, 2, \dots$ . Consider two repair policies for this device through two following sequences of  $(Q_h)$  and  $(\acute{Q}_h)$  which denote the conditional distribution of  $R_n$  given  $X_{n-1} = h-1$  and the conditional distribution of  $\acute{R}_n$  given  $\acute{X}_{n-1} = h-1$ , respectively:

$$(Q_1(0), Q_1(1)) = (0.4, 0.6),$$

$$(Q_2(0), Q_2(1), Q_2(2)) = (0.5, 0.5, 0),$$

$$(Q_3(0), Q_3(1), Q_3(2), Q_3(3)) = (0.5, 0, 0, 0.5),$$

$$(\acute{Q}_1(0), \acute{Q}_1(1)) = (0.5, 0.5),$$

$$(\acute{Q}_2(0), \acute{Q}_2(1), \acute{Q}_2(2)) = (0.1, 0, 0.9),$$

$$(\dot{Q}_3(0), \dot{Q}_3(1), \dot{Q}_3(2), \dot{Q}_3(3)) = (0.7, 0, 0, 0.3).$$

Observe that  $\dot{Q}_1$  is stochastically smaller than  $Q_1$ . Also  $\dot{Q}_3 \preceq_{st} Q_3$ . Since the transition matrix of these Markov chains  $((X_n)$  and  $(\dot{X}_n))$ , is directly characterized by the sequence  $(Q_h)_{h=1,2,3}$  and  $(\dot{Q}_h)_{h=1,2,3}$  respectively, therefore for this device, by Theorem 2.1,  $T_1 \preceq_{st} \dot{T}_1$  and  $T_3 \preceq_{st} \dot{T}_3$ .

**Example 2.2.** Machines replacement at pre-established age, [9]:

Let  $D_1, D_2, \dots$  are identical devices designed to work in one and the same position in a system  $S$ . We start with  $D_1$  at time 0. For  $r = 1, 2, \dots$ , if  $D_r$  exits its perfection-state or it reaches its age  $k$ ,  $D_{r+1}$  is installed in the place of  $D_r$ . Define  $\mathbb{N}$ -valued i.i.d random variables,  $N_1, N_2, \dots$  where  $N_r$  represents the number of times in which  $D_r$  is able to work in the perfection-state. Define the probability  $q_h$ , to be the conditional probability of  $N_1 = h + 1$  given  $N_1 > h$ , for  $h = 0, 1, \dots$ . Consider, moreover, the number of times in which  $D_r$  remains in operations by the random variable  $Y_r = \min(N_r, k)$ . Then the instant of  $m$ -th replacement, will be  $\sum_{r=1}^m Y_r$  and the discrete-time counting process  $R_n = \sup \{m \mid \sum_{r=1}^m Y_r \leq n\}$ , counts the total number of replacements within time  $n$ , for  $n = 1, 2, \dots$ . By setting  $X_0 = 0$ , the age of the device working at time  $n$ , represented by  $X_n = n - \sum_{r=1}^{R_n} Y_r$ . Observe that  $(X_n)_{n \geq 0}$  is an homogeneous Markov chain belonging to the class  $C_k$ . For this Markov chain,  $p_{k,0} = 1$  and positive transition probabilities are given as follows:

$$p_{h,h+1} = q_h, p_{h,0} = 1 - q_h, h = 0, 1, 2, \dots, k - 1.$$

Therefore for instant  $k = 3$ , the transition probability matrix  $P$  is,

$$P = \begin{bmatrix} 1 - q_0 & q_0 & 0 & 0 \\ 1 - q_1 & 0 & q_1 & 0 \\ 1 - q_2 & 0 & 0 & q_2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In this example, the first-passage time  $T_k$  in the state  $k$  is the number of times until maturation of age  $k$  for the working device, that is until the first replacement, due to planned maintenance rather than to a failure. In this respect, any pair of matrices  $P$ ,  $\dot{P}$  for which one can prove  $T_k \preceq_{st} \dot{T}_k$  leads to an application in the frame of Machines replacement models.

### 3 Related work

As we seen in the previous section, the results on stochastic comparisons between hitting times for skip-free Markov chains, depend on transition probability matrix,  $P$ . On the other hand, the distribution and then the expected value of hitting time of these kind of Markov chains, depends on the nonzero eigenvalues of matrix  $P$ . Currently, we are working on finding relationship between the eigenvalues of  $P$  and the conditions under which the results of stochastic ordering for the different hitting times, hold.

## References

- [1] Diaconis, P. and Miclo, L.(2009), On times to quasi-stationarity for birth and death processes. *J. Theoret. Probab.*, **22**(3), 558-586.
- [2] Fill, J. A. (2009), On hitting times and fastest strong stationary times for skip-free and more general chains. *J. Theoret. Probab.*, **22**(3),587-600.
- [3] Fill, J. A. (2009), The passage time distribution for a birth-and-death chain: strong stationary duality gives a first stochastic proof. *J. Theoret. Probab.*, **22**(3), 543-557.
- [4] Zhou,K. (2013), Hitting time distribution for skip-free Markov chains: a simple proof. *Statist. Probab. Lett.*, **83**(7), 1782-1786.
- [5] De Santis, E. and Spizzichino, F. (2016), Some Sufficient Conditions for Stochastic Comparisons Between Hitting Times for Skip-free Markov Chains. *Methodol Comput Appl Probab.* , **18** ,1021-1034.
- [6] Brown, M. and Shao, Y.S.(1987), Identifying coefficients in the spectral representation for first passage time distributions. *Probab. Engrg. Inform. Sci*, **1**, 69-74.
- [7] Irle, A. and Gani, J. (2001), The detection of words and an ordering for Markov chains. *J. Appl. Probab.*, **38A**, 66-77.
- [8] Shaked, M. and Shanthikumar, J. G.(2007), *Stochastic orders*. Springer Series in Statistics. Springer, New York.
- [9] De Santis, E. and Spizzichino, F. (2013), Random evolution of degradation and occurrences of words in random sequences of letters. In: Frenkel IB, Karagrigoriou A, Lisnianski A, Kleyner AV (eds). Wiley, New York.



## On the reliability modeling of weighted $k$ -out-of- $n$ systems with randomly chosen components

Salehi, M.<sup>1</sup>      Shishebor, Z.<sup>2</sup>      Asadi, M.<sup>3</sup>

<sup>1,2</sup> Department of Statistics, Faculty of science, Shiraz University

<sup>3</sup> Department of Statistics, University of Isfahan

### Abstract

The weighted  $k$ -out-of- $n$  (in briefly  $k/n$ ) systems are among the most important kind of redundancy structures. We consider a weighted  $k/n$  system with dependent components where the system is built from two classes  $\mathfrak{C}_X$  and  $\mathfrak{C}_Y$  of components that are categorized according to their weights and reliabilities. It is assumed that a random number  $M$  of the components are chosen from  $\mathfrak{C}_X$  with distribution  $F_X$  and  $n - M$  components are from the set  $\mathfrak{C}_Y$  with distribution function  $F_Y$ . We further assume that the structure of dependency of the components can be modeled by a copula function. The reliability of the system, at any time  $t$ , is expressed as a mixture of reliability of weighted  $k/n$  systems with fixed number of the components of types  $\mathfrak{C}_X$  and  $\mathfrak{C}_Y$  in terms of the probability mass function  $M$ . Some stochastic orderings are made between two different weighted  $k/n$  systems. It is shown that when the random mechanism of the chosen components for two systems are ordered in usual stochastic (*st*) order then, under some conditions, the lifetimes of the two systems are also ordered in *st* order. We also compare the lifetimes of two different systems in the sense of stochastic precedence concept. The results are examined by several illustrative examples under different conditions.

**Keywords:** Reliability, Weighted  $k$ -out-of- $n$  system, Copulas, Stochastic order, Stochastic precedence.

<sup>1</sup>salehi\_marziye@shirazu.ac.ir

<sup>2</sup>shisheb@shirazu.ac.ir

<sup>3</sup>m.asadi@sci.ui.ac.ir

# 1 Introduction

$k/n$  systems are important redundancy structures in reliability engineering. A  $n$ -component system is said to be a  $k/n$  system if it operates as long as at least  $k$  components out of the  $n$  components operate. In these kind of systems all components have an equal portion to the performance of the entire system. Hence the number of working components specify the operation of the system. The stochastic and aging properties of  $k/n$  systems have been extensively investigated in the reliability literature (see, for example, Kuo and Zuo [5]). Recently  $k/n$  systems are extended to weighted  $k/n$  systems where the weight associated with each component can be considered as load/capacity of that component. If the performance of the  $k/n$  system is characterized by the total weights of operating components, then the system is said to have a weighted  $k/n$  structure. In other words, a weighted  $k/n$  system has an operational level of at least  $k$  if the total weight of working components is  $k$  or more. As a specific example of a weighted  $k/n$  system from real life applications refer to Samaniego and Shaked [9]. The reliability assessment and stochastic properties of weighted  $k/n$  systems under the assumption of independence between the components lifetimes, are studied by several authors (see, among others, Chen and Yang [2], Rushdi [8], Wu and Chen [12]). For an extensive survey on applications of the systems with weighted components see Samaniego and Shaked [9]. Recently, attempts have been made to study the properties of the weighted  $k/n$  systems consisting of dependent components. Copula functions are one of the popular applied methods for modeling dependence among lifetimes of the components. To the study of copula theory, we refer the reader to Nelsen [7]. Analysis of the systems reliability using copula functions can be found in Tang et al. [11].

According to the definition of a weighted  $k/n$  system, the system lies among the structures that are made of leastwise two different kinds of components (for more description and reliability properties of such systems, we refer the reader to Navarro et al. [6]). Recently Eryilmaz [3, 4], has considered a weighted  $k/n$  system including two kinds of components. Such a system is assumed to be composed of  $n$  components in which a fixed number  $m$  out of  $n$  is from a class  $\mathfrak{C}_X$  of identically distributed components and the rest of  $n - m$  components are from a class  $\mathfrak{C}_Y$  includes of identically distributed components. The cited author has obtained some results on the stochastic and aging properties of the lifetimes of weighted  $k/n$  systems based on various copulas.

The goal of the present paper is to investigate the reliability and stochastic properties of weighted  $k/n$  systems consist of a random number of components that the components are from two different types.

# 2 Preliminaries

In this section we bring some notions related with main results. Throughout the paper, for continuous random variable (r.v)  $X$ , let  $\bar{F}$ ,  $f$  and  $\lambda_F = f/\bar{F}$  be survival function, density function and failure rate function, respectively. For continuous random variable  $Y$ , functions  $\bar{G}$ ,  $g$  and  $\lambda_G = g/\bar{G}$  are similarly defined. We say that a random variable  $X$  is smaller than the random variable  $Y$  in the

[(i)]usual stochastic order (denoted by  $X \leq_{st} Y$ ) if for all  $t$ ,  $\bar{F}(t) \leq \bar{G}(t)$  ; hazard rate order (denoted by  $X \leq_{hr} Y$ ) if for all  $t$ ,  $\lambda_F(t) \geq \lambda_G(t)$ ; likelihood ratio order (denoted by  $X \leq_{lr} Y$ ) if  $g(t)/f(t)$  is an increasing function of  $t$ .

It is well known that the following implications hold between these orderings:

$$X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{st} Y.$$

The concepts of partial orderings are extended to the multivariate case. For  $n$ -dimensional vector of random variables  $\mathbf{X}$ , let  $\bar{F}$  and  $f$  be survival function (i.e.,  $\bar{F}(\mathbf{t}) = P(\mathbf{X} > \mathbf{t})$ ) and density function, respectively. For  $n$ -dimensional vector of random variables  $\mathbf{Y}$ , functions  $\bar{G}$  and  $g$  are similarly defined. We denote  $\min\{\mathbf{t}_1, \mathbf{t}_2\} = (\min\{t_{11}, t_{12}\}, \dots, \min\{t_{n1}, t_{n2}\})$  and  $\max\{\mathbf{t}_1, \mathbf{t}_2\} = (\max\{t_{11}, t_{12}\}, \dots, \max\{t_{n1}, t_{n2}\})$ . We say that a random vector  $\mathbf{X}$  is smaller than the random vector  $\mathbf{Y}$  in the multivariate stochastic order if and only if

$$P(\mathbf{X} \in \mathfrak{U}) \leq P(\mathbf{Y} \in \mathfrak{U}),$$

for all sets  $\mathfrak{U} \subseteq R^n$  (where  $\mathfrak{U}$  is an upper set). Suppose that

$$\bar{F}(\mathbf{t}_1)\bar{G}(\mathbf{t}_2) \leq \bar{F}(\min\{\mathbf{t}_1, \mathbf{t}_2\})\bar{G}(\max\{\mathbf{t}_1, \mathbf{t}_2\}) \quad \forall \mathbf{t}_1, \mathbf{t}_2 \in \mathbb{R}^n.$$

Then we say that the random vector  $\mathbf{X}$  is smaller than the random vector  $\mathbf{Y}$  in the multivariate hazard rate order. Let

$$f(\mathbf{t}_1)g(\mathbf{t}_2) \leq f(\min\{\mathbf{t}_1, \mathbf{t}_2\})g(\max\{\mathbf{t}_1, \mathbf{t}_2\}) \quad \forall \mathbf{t}_1, \mathbf{t}_2 \in \mathbb{R}^n.$$

Then we say that the random vector  $\mathbf{X}$  is smaller than the random vector  $\mathbf{Y}$  in the multivariate likelihood ratio order.

In reliability theory, the notion of “stochastic precedence” (*sp*) has been introduced to compare the performance of two systems (see [1]). The concept of *sp* is defined as follows. Let  $X_1$  and  $X_2$  be two random variables. Then we say that  $X_1$  is stochastically precede  $X_2$  (written  $X_1 \leq_{sp} X_2$ ) if and only if  $P(X_1 \leq X_2) \geq 1/2$ . The random variables  $X_1$  and  $X_2$  are *sp*-equivalent if and only if they satisfy  $P(X_1 \leq X_2) = 1/2$ . For two independent random variables  $X$  and  $Y$  if  $X \leq_{st} Y$ , then  $X$  is better than  $Y$  in the sense of stochastic precedence (*sp*) [1].

### 3 System description

In the following, we consider a weighted  $k/n$  system which consists of  $n$  dependent components. These lifetimes belong to two distinct classes  $\mathfrak{C}_X = \{X_1, \dots, X_m\}$  and  $\mathfrak{C}_Y = \{Y_{m+1}, \dots, Y_n\}$  with sizes  $m$  and  $n - m$ , respectively. Let  $F_X$  and  $F_Y$  denote the distribution functions of the components in  $\mathfrak{C}_X$  and  $\mathfrak{C}_Y$ , respectively. We assume that the components lie in two situations at any time, either working or failed. When a component in  $\mathfrak{C}_X$  ( $\mathfrak{C}_Y$ ) lies in a working situation it has an operating level of  $w$  ( $w^*$ ) as the weight of the component. So, the system is in performance level  $k$  the total weights of operating components is  $k$  or above. In other words, components of the system are chosen from two distinct classes of components such that  $m$  of them have the weight

$w$  with lifetime distribution  $F_X$ , and the remaining  $n - m$  components have the weight  $w^*$  with lifetime distribution  $F_Y$ . If  $X_1, \dots, X_m, Y_{m+1}, \dots, Y_n$  denote the lifetimes of the system components in the two classes, then the total weight of the system is expressed by the stochastic process  $W_n(t)$  at time  $t \geq 0$

$$W_n(t) = \sum_{i \in \mathfrak{C}_X} w_i I(X_i > t) + \sum_{i \in \mathfrak{C}_Y} w_i^* I(Y_i > t)$$

and the lifetime of the ,  $T$ , is

$$T = \inf\{t : W_n(t) < k\}.$$

Hence the system reliability is

$$R(t) = P(T > t) = P(W_n(t) \geq k), \quad \forall t \geq 0.$$

Suppose that the dependence between  $X_1, \dots, X_m, Y_{m+1}, \dots, Y_n$  is modeled by the  $n$ -dimensional copula function  $C$ , i.e. the joint distribution function of  $X_1, \dots, X_m, Y_{m+1}, \dots, Y_n$  is displayed as

$$H(t_1, \dots, t_n) = C(F_X(t_1), \dots, F_X(t_m), F_Y(t_{m+1}), \dots, F_Y(t_n)).$$

For a system with two different kind of dependent components when dependency is modeled by copulas, the survival function  $R(t)$  is studied by Eryilmaz [4].

In what follows, we consider the case when  $n$ -copula is symmetric and the size of two classes  $\mathfrak{C}_X$  and  $\mathfrak{C}_Y$  are random. Let  $M$  be a random variable with support contained in  $\{0, 1, \dots, n\}$ . If  $m$  (the number of components from  $\mathfrak{C}_X$ ) is selected randomly according to the random variable  $M$ , then the system reliability function can be represented as

$$\begin{aligned} R_M(t) &= \sum_{m=0}^n \sum_{\substack{wy+w^*z \geq k \\ 0 \leq y \leq m \\ 0 \leq z \leq n-m}} \binom{m}{y} \binom{n-m}{z} \sum_{l=0}^y \sum_{s=0}^z (-1)^{(l+s)} \binom{y}{l} \binom{z}{s} \\ &\quad \times \underbrace{C(\underbrace{F_X(t), \dots, F_X(t)}_{(m-y+l)\text{times}}, \underbrace{F_Y(t), \dots, F_Y(t)}_{(n-m-z+s)\text{times}})}_{P(M=m)}. \end{aligned}$$

The last equality follows from (4) in Eryilmaz [4].

Let us look at the following example which rely on the notations introduced just now.

**Example 1.** Consider a weighted 3/4 system with  $w = 2$ ,  $w^* = 1$ ,  $F_X(t) = 1 - \exp\{-0.1t\}$ ,  $F_Y(t) = 1 - \exp\{-0.4t\}$ . Assume that the dependence structure among components is generated by FGM copula. If the random variables  $M_1$  and  $M_2$  follow

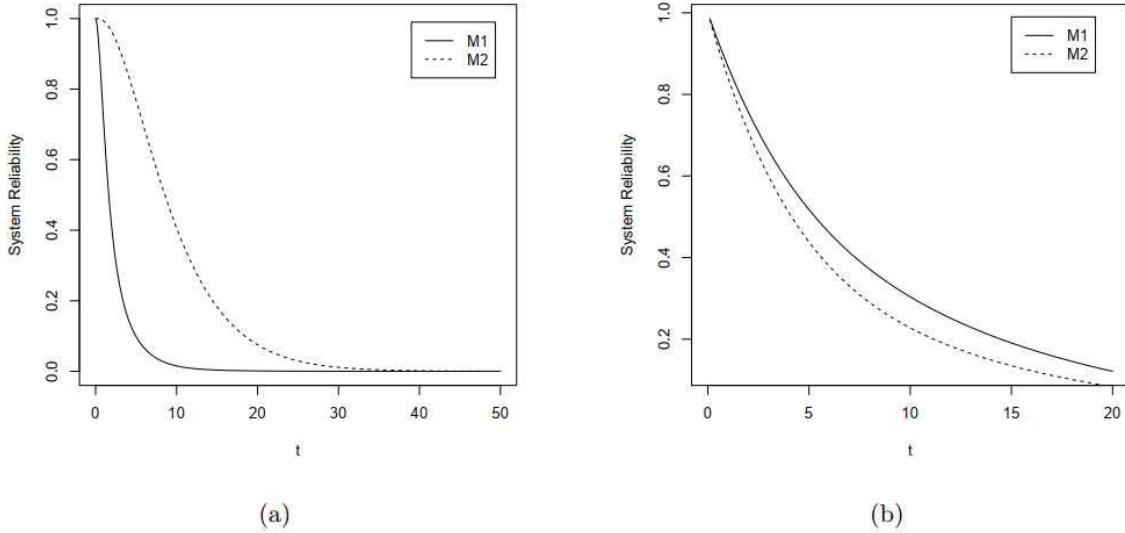


Figure 1: Systems reliability (a): related to Example 1, (b): related to Example 2.

Binomial distribution  $B(n, 0.1)$  and  $B(n, 0.9)$ , respectively, then for  $i = 1, 2$ ,

$$\begin{aligned}
R_{M_i}(t) &= \sum_{m=0}^4 P(2 \sum_{i \in \mathfrak{C}_X} I(X_i > t) + \sum_{i \in \mathfrak{C}_Y} I(Y_i > t) \geq 3) P(M_i = m) \\
&= \sum_{m=0}^4 R_m(t) P(M_i = m) \\
&= \sum_{m=0}^4 \sum_{\substack{2y+z \geq 3 \\ 0 \leq y \leq m \\ 0 \leq z \leq 4-m}} \binom{m}{y} \binom{4-m}{z} \sum_{l=0}^y \sum_{s=0}^z (-1)^{l+s} \binom{y}{l} \binom{z}{s} \\
&\quad (1 - \exp\{-0.1t\})^{m-y+l} (1 - \exp\{-0.4t\})^{4-m-z+s} \\
&\quad (1 + \theta(\exp\{-0.1t\})^{m-y+l} (\exp\{-0.4t\})^{4-m-z+s}) P(M_i = m)
\end{aligned}$$

Figure 1, (a), shows reliability results of these systems when  $\tau = 0.001$ .

## 4 Comparison between two systems based on usual stochastic ordering

In this section we consider two independent weighted  $k/n$  systems with lifetimes  $T_1$  and  $T_2$ . We aim to investigate conditions under which the reliability of the systems are ordered in two cases. In the first case, let  $T_i$  be the lifetime of the system built for the case when  $M_i$  components are chosen randomly from the class  $\mathfrak{C}_{X_i}$ ,  $i = 1, 2$ . The reliability of the two systems are compared in Theorem 2.4. In the second case, let  $T_i$  be the lifetime of the system built for the case when the distribution functions of the components in  $\mathfrak{C}_{X_i}$  and  $\mathfrak{C}_{Y_i}$  are in stochastic order for  $i = 1, 2$ . A comparison of systems reliability is shown in Theorem 2. To organize the proof of the theorems, we present the following lemmas.

**Lemma 4.1.** Let  $X_1, \dots, X_n$  be continuous random variables with symmetric copula  $C_{X_1, \dots, X_n}$ , marginal distribution functions  $F_{X_1}, \dots, F_{X_n}$ , and joint distribution function  $H_{X_1, \dots, X_n}$ . Let  $Z_i = I(X_i > t)$ ,  $i = 1, \dots, n$ . Then there exist a symmetric selection for the copula  $C_{Z_1, \dots, Z_n}$ .

**Lemma 4.2.** Assume that  $R_m$  represents the reliability of a weighted  $k/n$  system with two distinct classes  $\mathfrak{C}_X$  and  $\mathfrak{C}_Y$  where the size of the class  $\mathfrak{C}_X$  is  $m$ . Suppose that the components of the system are dependent with a symmetric copula. Let also  $w$  and  $w^*$  be the weights of the components and  $\bar{F}_X$  and  $\bar{F}_Y$  denote the reliability functions of the components in  $\mathfrak{C}_X$  and  $\mathfrak{C}_Y$ , respectively. If  $w \leq w^*$  and  $\bar{F}_X(t) \leq \bar{F}_Y(t)$ , for all  $t$ , then  $R_m \leq R_{m-1}$ , for all  $m = 0, \dots, n$ , with convention  $R_{-1}(t) = 0$ , where

$$R_m(t) = P \left( w \sum_{i=1}^m I(X_i > t) + w^* \sum_{i=1}^{n-m} I(Y_i > t) \geq k \right).$$

Now, we are ready to prove the main theorems.

**Theorem 1.** Suppose that  $T_i$ ,  $i = 1, 2$  represents the lifetime of a weighted  $k/n$  system where the system components are dependent with a symmetric copula and the components are chosen randomly from two distinct classes  $\mathfrak{C}_X$  and  $\mathfrak{C}_Y$ . Let also  $w$  and  $w^*$  be the weights of the components and  $\bar{F}_X$  and  $\bar{F}_Y$  denote the reliability functions of the components in  $\mathfrak{C}_X$  and  $\mathfrak{C}_Y$ . Assume that the random variables  $M_1$  and  $M_2$  are the numbers of selected components from the set  $\mathfrak{C}_X$  in the first and second system, respectively.

[(i)] If  $M_1 \leq_{st} M_2$ ,  $w \leq w^*$  and for all  $t$ ,  $\bar{F}_X(t) \leq \bar{F}_Y(t)$ , then  $T_1 \geq_{st} T_2$ . If  $M_1 \leq_{st} M_2$ ,  $w \geq w^*$  and for all  $t$ ,  $\bar{F}_X(t) \geq \bar{F}_Y(t)$ , then  $T_1 \leq_{st} T_2$ .

The following example gives an illustration of Theorem 2.4.

**Example 2.** Consider a weighted 2/2 system with  $w = 1$ ,  $w^* = 2$ ,  $F_X(t) = 1 - \exp\{-0.2t\}$ ,  $F_Y(t) = 1 - \exp\{-0.1t\}$  and assume that the dependence structure among components is generated by Clayton copula. We consider two different cases for random variable  $M$ . In the first one we assume that  $M = M_1$  takes its values randomly from  $\{0, 1, 2\}$  with uniform probabilities and in the second case  $M = M_2$  takes its values on the same set with probabilities  $\{1/6, 2/6, 3/6\}$ .

We leave it to the reader to verify that  $M_1 \leq_{st} M_2$  and  $\bar{F}_X(t) \leq \bar{F}_Y(t)$ . So the conditions of Theorem 2.4 are satisfied. Hence  $R_{M_1}(t) \geq R_{M_2}(t)$ . The plots of the reliability functions of the systems are depicted in Figure 1, (b), when  $\tau = 0.5$ . In the case that the Clayton copula evinces negative dependency, the obtained result is similar to the positive dependency.

In the following, we explore the conditions under which the stochastic ordering between the components of two systems imply the stochastic ordering between the lifetimes of the systems.

**Theorem 2.** Suppose that  $T_i$  represent the lifetime of a weighted  $k/n$  system with two classes  $\mathfrak{C}_{X_i}$  and  $\mathfrak{C}_{Y_i}$  where the size of the class  $\mathfrak{C}_{X_i}$  is random,  $i = 1, 2$ . Let  $M$  be a random variable with support contained in  $\{0, 1, \dots, n\}$ . Suppose the size of the class  $\mathfrak{C}_{X_i}$  is selected randomly according to the random variable  $M$ . Moreover, Let  $F_{X_i}$  and  $F_{Y_i}$  indicate the

distribution functions of the components in  $\mathfrak{C}_{X_i}$  and  $\mathfrak{C}_{Y_i}$ . We assume when a component in  $\mathfrak{C}_{X_i}$  and  $\mathfrak{C}_{Y_i}$  lies in a working situation it has an operating level of  $w$  ( $w^*$ ) as the weight of the component,  $i = 1, 2$ . Let the dependence structure among components built with a symmetric copula. If  $F_{X_1}(t) \leq F_{X_2}(t)$ ,  $F_{Y_1}(t) \leq F_{Y_2}(t)$ , for all  $t$ , then  $T_1 \geq_{st} T_2$ .

Let  $T_1$  and  $T_2$  be two independent lifetimes of weighted  $k/n$  systems. Note that since,

$$P(T_1 \geq T_2) = \int_0^\infty P(T_1 \geq t) dF_{T_2}(t),$$

then it is easy to see that the usual stochastic order implies the *sp* order. Hence if  $T_1$  and  $T_2$  are the lifetimes of the systems described in Theorems 2.4 (i) and 2, then we have  $P(T_1 \geq T_2) \geq 1/2$ .

In the following, we give two examples to measure the difference between  $P(T_1 \geq T_2)$  and  $1/2$  in two cases, the first is when the dependence structure is Clayton and the second is when the component lifetimes are independent.

**Example 3.** Let the weighted  $k/10$  system for  $k = 1, \dots, 11$  with  $w = 1$ ,  $w^* = 2$ ,  $F_{X_1}(t) = 1 - \exp\{-0.1t\}$ ,  $F_{Y_1}(t) = 1 - \exp\{-0.5t\}$ ,  $F_{X_2}(t) = 1 - \exp\{-0.2t\}$ ,  $F_{Y_2}(t) = 1 - \exp\{-0.6t\}$ . So the conditions of Theorem 2 are satisfied. Thus by Remark 4, we have  $P(T_1 \geq T_2) \geq 1/2$ . Now, assume that the dependence structure among components is generated by Clayton copula when  $\tau = 0.5$ . In the case that the random variable  $M$  follows Binomial distribution  $B(n, 0.7)$ , the value of  $P(T_1 \geq T_2)$  for different  $k$  are calculated. Given all the results together in Table 1.

Table 1: The values of  $P(T_1 \geq T_2)$  for system in Example 3.

k	1	2	3	4	5	6	7	8	9	10
Clayton	0.6775	0.6763	0.6759	0.6747	0.6703	0.6585	0.6359	0.6047	0.5749	0.5566

## 5 Conclusions

In this paper one of the most important kind of redundancy structures is considered , the weighted  $k/n$  systems. We assumed that the system is built from two classes of components that are categorized according to their weights and reliabilities. It is assumed that a number of components are selected randomly from the first type and the rest of components are from the second type. Further more, copula function is considered for modeling of the structure of dependency within the components. The dynamic reliability of the system, is expressed as a mixture of reliability of weighted  $k/n$  systems with fixed number of the components of two types in terms of the probability mass function of the random number. We have shown that when the random mechanism of the chosen components for two systems are ordered in stochastic (*st*) order then, under some conditions on the weights and distributions of components lifetime, the lifetimes of the two systems are also ordered in the sense of *st*. In addition, we have considered conditions under which, the lifetimes of two different systems are compared in the sense of stochastic precedence concept. We have presented some illustrative examples under different conditions for our results.

## References

- [1] Arcones, M.A., Kvam, P.H. and Samaniego, F.J. (2002). Nonparametric estimation of a distribution subject to a stochastic precedence constraint. *Journal of the American Statistical Association* 97: 170-182.
- [2] Chen, Y. and Yang, Q. (2005). Reliability of two-stage weighted k-out-of-n systems with components in common. *IEEE Transactions on Reliability* 54: 431-440.
- [3] Eryilmaz, S. and Sarikaya, K. (2013). Modeling and analysis of weighted-k-out-of-n:G system consisting of two different types of components. *Proceedings of the Institution of Mechanical Engineers Part O: Journal Risk and Reliability* 201: 1-7.
- [4] Eryilmaz, S. (2014). Multivariate copula based dynamic reliability modeling with application to weighted-k-out-of-n systems of dependent components. *Structural Safety* 51: 23-28.
- [5] Kuo, W. and Zuo, M.J. (2003). *Optimal Reliability Modeling: Principles and Applications*. New York: John Wiley and Sons.
- [6] Navarro, J., Pellerey, F. and Di Crescenzo, A. (2015). Orderings of coherent systems with randomized dependent components. *European Journal of Operational Research* 240: 127-139.
- [7] Nelsen, R.B. (2006). *An introduction to copulas*. New York: Springer.
- [8] Rushdi, A.M. (1990). Threshold systems and their reliability. *Microelectron Reliability* 30: 299-312.
- [9] Samaniego, F.J. and Shaked, M. (2008). Systems with weighted components. *Statistics and Probability Letters* 78: 815-823.
- [10] Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic orders*. Springer series in statistics. New York: Springer.
- [11] Tang, X.S., Li, D.Q., Zhou, C.B., Phoon, K.K. and Zhang, L.M. (2013). Impact of copulas for modeling bivariate distributions on system reliability. *Structural Safety* 44: 80-90.
- [12] Wu, J.S. and Chen, R.J. (1994). An algorithm for computing the reliability of a weighted-k-out-of-n system. *IEEE Transactions on Reliability* 43: 327-328.



## Bayesian and non-bayesian reliability analysis of the exponential family based on dual generalized order statistics

Salim, N.<sup>1</sup> Dordizadeh, H.<sup>2</sup>

<sup>1,2</sup> Department of Statistics, Lorestan University

### Abstract

In this paper, the Bayesian and the non-bayesian estimators for the parameters, the reliability and the hazard functions of the class of an exponential family are obtained based on dual generalized order statistics (*DGOS*). The Bayesian estimators are acquired under the symmetric (squared error loss) and the asymmetric (linex) loss functions. Finally, a simulation study is performed to compare different Bayesian estimators based on different loss functions with the classical estimators.

**Keywords:** Bayes analysis, DGOS, Exponential family, Loss function, MLE.

## 1 Introductions

The dual generalized order statistics (*DGOS*) was introduced by Kampas et al [5] and it was widely used by several authors. Let  $\{X_j; i \geq 1\}$  be a sequence of random variables (*RV*) from independent identically distribution (*i.i.d*) of a population with absolutely continuous cumulative distribution function (*cdf*)  $F$  and corresponding probability density function (*pdf*) of  $f$ ; then the joint *pdf* of the first  $n$ , *DGOS* can be written as

$$f^{X(1,n,m,k), \dots, X(n,n,m,k)}(x_1, \dots, x_n) = \prod_{j:1}^n \gamma_j \left\{ \prod_{j:1}^{n-1} (\bar{F}(x_j))^{\gamma_j - \gamma_{j+1} - 1} f(x_j) \right\} (\bar{F}(x_n))^{\gamma_{n-1}} f(x_n), \quad (1)$$

<sup>1</sup>salimstat@yahoo.com

<sup>2</sup>Shafaghemra@yahoo.com

on the cone  $F^{-1}(0) < x_1 < \dots < x_n < F^{-1}(1) \subseteq R^n$ ,  $\gamma_1 \dots \gamma_n > 0$  such that  $\gamma_n = k > 0$  and  $\gamma_r = k + n - r + M_r$  for all  $r : 1, \dots, n - 1$  where  $M_r = \sum_{j:r}^{n-1} m_j$  and  $\tilde{m} = (m_1, \dots, m_{n-1}) \in R^{n-1}$ .

If  $m_j = 0; j : 1, \dots, n - 1$  and  $\gamma_n = 1$  then (1) reduces to the ordinary order statistics, for  $m_j = -1; j : 1, \dots, n - 1$  and  $\gamma_n = k$  procures the k-records and for  $m_j = -1; j : 1, \dots, n - 1$  and  $\gamma_n = 1$  obtains the records values pdf, the sequential order statistics can be achieved if  $m_j = (n - j + 1)\alpha_j - (n - j)\alpha_{j+1} - 1; j : 1, \dots, n - 1, k = \alpha_n$  and  $\alpha_1, \dots, \alpha_n > 0$ . More details about DGOS and some applications, can be found in, for instance [1, 2, 3, 8, 9], recently Jaheen and Al Harbi[4] studied the Bayesian estimation of exponentiated weibull DGOS. [10] introduced a general form for some exponential distributions as following

$$F_X(x) = 1 - \exp\{-\beta Q(x; \theta)\}, \quad x, \beta, \theta > 0, \quad (2)$$

$$f_X(x) = \beta q(x; \theta) \exp\{-\beta Q(x; \theta)\}, \quad x > 0.$$

With attantion to introduced family in (2), the reliability or survial function , $R(t)$ , and hazard function , $H(t)$ , at a mission time  $t$  are respectivly, given by

$$R(t) = \exp\{-\beta Q(t; \theta)\},$$

and

$$H(t) = q(t; \theta).$$

Thus, considered distributional family in this paper have strictly decreasing reliability. Rest of the paper is organized as follows. In section 2 we have derived the maximum likelihood estimators (MLEs) of model parameters, the rilability and the hazard functions. Section 3 is devoted to give Bayesian estimators under the symmetric and the asymmetric loss functions.

## 2 Frequentist estimation

In analyzing the reliability data, the maximum likelihood (ml) and inference based on the likelihood play a fundamental role. One of the reasons for using ML is that this method has the ability to use in different types of data collection and that it is optimal for under-conditions under the circumstances. In other words, if the sample size is large enough, the ML method leads to estimates that are asymptotically of the least variance and will have a normal distribution. Given all these features, it would seem that this method can not be ignored. Therefore, we will devote the beginning of the discussion to this method.

Let  $X(1, n, m, k), \dots, X(n, n, m, k)$  be  $n$ , DGOS drawn from introduced model in (2). The likelihood function is given by

$$\begin{aligned}
L(\beta, \theta; \mathbf{x}) &= \beta^n \prod_{j:1}^n \gamma_j \left\{ \exp \left\{ -\beta \left( \sum_{j:1}^{n-1} \{ (\gamma_j - \gamma_{j+1} - 1) Q(x_j; \theta) - Q(x_j; \theta) \} \right) \right\} \right\} \\
&\quad \times \prod_{j:1}^n q(x_j; \theta) \exp \{ -\gamma_n \beta Q(x_n; \theta) \}.
\end{aligned} \tag{3}$$

We presen *MLE* based on *DGOS*, by using (3) the natural logarithm of (3) we obtain

$$\begin{aligned}
\ell(\beta, \theta, \mathbf{x}) &= \sum_{j:1}^n \ln \gamma_j + \sum_{j:1}^{n-1} [(\gamma_j - \gamma_{j+1} - 1) \ln \bar{F}(x_j) + \ln f(x_j)] \\
&\quad + (\gamma_n - 1) \ln \bar{F}(x_n) + \ln f(x_n) \\
&= \sum_{j:1}^n \ln \gamma_j + \sum_{j:1}^{n-1} \{ (\gamma_j - \gamma_{j+1} - 1) \beta Q(x_j; \theta) + \ln q(x_j; \theta) \\
&\quad - \beta Q(x_j; \theta) \} + (\gamma_n - 1) \beta Q(x_n; \theta) + \ln q(x_n; \theta) - \beta Q(x_n; \theta).
\end{aligned} \tag{4}$$

If both of the parameters  $\theta$  and  $\beta$  are unknown, their *MLEs*,  $\hat{\theta}_{ML}$  and  $\hat{\beta}_{ML}$  can be obtained by solving the following likelihood equations

$$\begin{aligned}
\frac{\partial \ell}{\partial \beta} &= \sum_{j:1}^{n-1} \{ (\gamma_j - \gamma_{j+1} - 1) Q(x_j; \theta) - Q(x_j; \theta) \} + (\gamma_n - 2) Q(x_n; \theta), \\
\frac{\partial \ell}{\partial \theta} &= \sum_{j:1}^{n-1} \{ (\gamma_j - \gamma_{j+1} - 1) \beta \frac{\partial}{\partial \theta} Q(x_j; \theta) + \frac{\frac{\partial}{\partial \theta} q(x_j; \theta)}{q(x_j; \theta)} \\
&\quad - \beta \frac{\partial}{\partial \theta} Q(x_j; \theta) \} + (\gamma_n - 2) \beta Q(x_n; \theta) + \frac{\frac{\partial}{\partial \theta} q(x_n; \theta)}{q(x_n; \theta)}.
\end{aligned}$$

Since the equations can not be solved analytically, we can use the well-known Newton–Raphson technique.

Using the stability property of *MLE* method, we can derive the *MLE* of the reliability,  $R(t)$ , and the hazard function,  $H(t)$ , at mission time  $t$  respectively as follows

$$\hat{R}(t) = \exp \{ -\hat{\beta} Q(t; \hat{\theta}) \}, \tag{5}$$

and

$$\hat{H}(t) = \hat{\beta} q(t; \hat{\theta}). \tag{6}$$

### 3 Bayesian infrence

This section is concerned with the problem of obtaining Bayesian estimators. For calculate the estimators, first its need to acquire the posterior distributions.

### 3.1 Posterior distributions

For case that both the  $\beta$  and  $\theta$  are unknown, we use Soland's method [6]. We assume that the parameter  $\theta$  has a discrete prior and  $\beta$  has a continuous conditional prior for given  $\theta$ . Means  $P(\theta = \theta_k) = \xi_k$   $k : 1, \dots, N$  such that  $\sum_{k:1}^N \xi_k = 1$  and for given  $\theta_k$ , conditional conjunction prior distribution of  $\beta$  as follows

$$\pi(\beta|\theta_k) = b_k e^{-b_k \beta}, \quad b_k > 0, \beta > 0, \quad (7)$$

where  $\xi_k$ ;  $k : 1, \dots, N$  are the hyper parameters. We use the Soliman's method for estimate  $b_k$  [7]. By using (5) and (7), the expectation of  $R(t)$  at mission time  $t$  can be expressed as

$$E R(t) = \int_0^\infty R(t) \pi(\beta|\theta_k) d\beta = \frac{b_k}{b_k + Q(t;\theta_k)}.$$

Let  $E R(t) = \hat{R}(t)$ , for a given mission time  $t$ , then  $\hat{b}_k$  are given by

$$\hat{b}_k = \frac{\hat{R}(t)}{1 - \hat{R}(t)} Q(t;\theta_k), \quad k : 1, 2, \dots, N. \quad (8)$$

Where  $\hat{R}(t)$  is MLE of  $R(t)$ .

The conditional posterior distribution of  $\beta$  for the given  $\theta$ ; from (3) and (7) can be expressed as

$$\begin{aligned} \pi^*(\beta|\theta_k, \mathbf{x}) &= \frac{\pi(\beta|\theta_k) L(\beta, \theta_k; \mathbf{x})}{\int_0^\infty \pi(\beta|\theta_k) L(\beta, \theta_k; \mathbf{x}) d\beta} \\ &= \frac{\beta^n \exp\{-\beta [\sum_{j:1}^{n-1} (\gamma_j - \gamma_{j+1}) Q(x_j; \theta_k) + (b_k + \gamma_n Q(x_n; \theta_k))] \}}{\int_0^\infty \beta^n \exp\{-\beta [\sum_{j:1}^{n-1} (\gamma_j - \gamma_{j+1}) Q(x_j; \theta_k) + (b_k + \gamma_n Q(x_n; \theta_k))] \} d\beta} \\ &= \frac{\beta^n \{ \sum_{j:1}^{n-1} (\gamma_j - \gamma_{j+1}) Q(x_j; \theta_k) + b_k + \gamma_n Q(x_n; \theta_k) \}^{n+1}}{\Gamma(n+1)} \\ &\quad \times \exp\{-\beta \{ \sum_{j:1}^{n-1} (\gamma_j - \gamma_{j+1}) Q(x_j; \theta_k) + b_k + \gamma_n Q(x_n; \theta_k) \} \}. \end{aligned} \quad (9)$$

Moreover, the distribution of  $\theta$  given  $\mathbf{x}$  can be written as

$$P_k = P(\theta = \theta_k | \mathbf{x}) = \int_0^\infty \pi^*(\beta, \theta_k | \mathbf{x}) d\beta. \quad (10)$$

From (3), (7) and (10), the joint posterior distribution of  $\beta$  and  $\theta$  is given by

$$\begin{aligned} \pi^*(\beta, \theta_k | \mathbf{x}) &= \frac{P(\theta = \theta_k) \pi(\beta|\theta_k) L(\beta, \theta_k; \mathbf{x})}{\sum_{k:1}^N \int_0^\infty P(\theta = \theta_k) \pi(\beta|\theta_k) L(\beta, \theta_k; \mathbf{x}) d\beta} \\ &= \frac{\xi_k b_k \beta^n \prod_{j:1}^n q(x_j; \theta_k) \exp\{-\beta C_k\}}{\sum_{k:1}^N \xi_k b_k \prod_{j:1}^n q(x_j; \theta_k) \Gamma(n+1) C_k^{-(n+1)}}. \end{aligned} \quad (11)$$

Finally, the marginal posterior mass function of  $\theta_k$  by integration of (11) consequence

$$P_k = \frac{\xi_k b_k \prod_{j:1}^n q(x_j; \theta_k) C_k^{-(n+1)}}{\sum_{k:1}^N \xi_k b_k \prod_{j:1}^n q(x_j; \theta_k) C_k^{-(n+1)}}. \quad (12)$$

Where  $C_k = \sum_{j:1}^{n-1} (\gamma_j - \gamma_{j+1}) Q(x_j; \theta_k) + 2 b_k + \gamma_n Q(x_n; \theta_k)$ .

### 3.2 Symmetric loss function

In Bayesian perspective, choice of loss function ( $LF$ ) is an important part in the estimation problems. In most of the Bayesian estimation problems, authors consider the underlying  $LF$  to be square error loss function ( $SELF$ ) to reason ease of computation. The Bayesian estimators for model parameters by using (11) and (12) can be obtained as

$$\hat{\beta}_{BS} = \int_0^\infty \beta \pi^*(\beta, \theta_k | \mathbf{x}) d\beta = \sum_{k:1}^N (n+1) P_k A_k^{-1}, \quad (13)$$

and

$$\hat{\theta}_{BS} = \sum_{k:1}^N P_k \theta_k. \quad (14)$$

Similarly, the Bayesian estimator for the reliability function with fixed  $t > 0$  can be obtained as

$$\hat{R}_{BS}(t) = \int_0^\infty R(t) \pi^*(\beta, \theta_k | \mathbf{x}) d\beta = \sum_{k:1}^N P_k \left( \frac{A_k}{A_k + Q(t; \theta_k)} \right)^{n+1}, \quad (15)$$

and for the hazard function as

$$\hat{H}_{BS}(t) = \sum_{k:1}^N P_k \frac{q(t; \theta_k)}{A_k} (n+1). \quad (16)$$

Where  $A_k = \sum_{j:1}^{n-1} (\gamma_j - \gamma_{j+1}) Q(x_j; \theta_k) + b_k + \gamma_n Q(x_n; \theta_k)$ , such that (13), (15) and (16) achieved by integrating of Gamma distribution.

### 3.3 Asymmetric loss function

The Linex loss function ( $LLF$ ) is given by

$$L(\delta, \omega(\theta)) = e^{a(\delta - \omega(\theta))} - a(\delta - \omega(\theta)) - 1, \quad a \neq 0,$$

and  $\delta$  is an estimator of  $\omega(\theta)$ . The behavior of the  $LLF$  changes with the choice of  $a$ .  $a > 0$  is used when overestimation is more serious than underestimation while if  $a < 0$

is used in the reverse situation. As  $a$  is nears zero, the  $LLF$  is look like  $SELF$ . The Bayesian estimator under  $LLF$  given by

$$\hat{\omega}_{BL}(\theta) = -\frac{1}{a} \ln[E_\omega(e^{-a\omega(\theta)})].$$

So by using (7),(9) and (10) the Bayesian estimator for model parameters under the  $LLF$  can be derived as

$$\hat{\beta}_{BL} = -\frac{1}{a} \ln\left[\sum_{k:1}^N P_k \left(\frac{A_k}{A_k + a}\right)^{n+1}\right], \quad (17)$$

and

$$\hat{\theta}_{BL} = -\frac{1}{a} \ln\left[\sum_{k:1}^N P_k e^{-a\theta_k}\right]. \quad (18)$$

Moreover, the Bayesian estimators for the reliability and the hazard functions under the  $LLF$  with fixed  $t > 0$  can be obtained as

$$\hat{R}_{BL}(t) = -\frac{1}{a} \ln\left[\sum_{k:1}^N \sum_{\ell:1}^{\infty} P_k \frac{(-a)^\ell}{\ell!} \left(\frac{A_k}{A_k + \ell Q(t; \theta_k)}\right)^{n+1}\right], \quad (19)$$

and

$$\hat{H}_{BL}(t) = -\frac{1}{a} \ln\left[\sum_{k:1}^N P_k \left(\frac{A_k}{A_k + a q(t; \theta_k)}\right)^{n+1}\right]. \quad (20)$$

Where  $A_k = \sum_{j:1}^{n-1} (\gamma_j - \gamma_{j+1}) Q(x_j; \theta_k) + b_k + \gamma_n Q(x_n; \theta_k)$ , such that (17), (19) and (20) achieved by integrating of Gamma distribution.

## 4 Simulation

We conducted a simulation experiment in following, to inquire the performance of the Bayesian and the non-Bayesian estimators.

1. Compute the  $MLE$  of  $\hat{\beta}, \hat{\theta}$  and obtain the  $\hat{R}(t)$ .
2. Compute the  $P_k$  by (8) and (12).
3. Compute the Bayesian estimators of  $\hat{\beta}, \hat{\theta}, \hat{R}(t)$  and  $\hat{H}(t)$  for a given mission time  $t > 0$ , under  $SELF$  using (13), (14), (15), (16) and under  $LLF$  by using (17), (18), (19), (20).
4. The above steps are repeated 1000 times, and we then obtain mean and mean squared error.

In all above cases the parameter of  $LLF$  chosen as  $a = 0.5$ . Since the weibull and rayleigh distributions are well-known in the reliability theory, we simulated their model parameters, the reliability and the hazard functions according to ML and the Bayesian estimators based on symmetric and asymmetric loss functions respectively in Table 5, 5 and 5.

## 5 Conclusion

Bayesian and Maximum Likelihood estimators of the model parameters, the reliability and the hazard functions are achieved. According to the results of Table 5, when lifetime increased, it can be seen that the reliability function slightly declined, while, there was a gradual growth in the risk function. However, we can observe to compare the Table 5 and 5, the behavior of  $BS$  and  $BL$  estimators of the reliability and the hazard functions significantly resembled. Overall, One of the another results of this paper is: the most of the well-definition lifetime distributions are contained by the introduced family.

Table 1. ML estimation of model parameters, the reliability and the hazard functions based  $DGOS$

Distribution	$n$	$m$	$k$	$\hat{\beta}$	$\hat{\theta}$	$\hat{R}$			$\hat{H}$		
						$t = 2$	$t = 3$	$t = 4$	$t = 2$	$t = 3$	$t = 4$
Weibull	6	4	2	0.0347	1.882	0.8799	0.7660	0.6244	0.1203	0.1721	0.2218
	8	6	4	0.0418	1.973	0.8486	0.6940	0.5250	0.1618	0.2401	0.3177
	10	8	6	0.0326	1.982	0.8919	0.7871	0.6676	0.1041	0.1451	0.1837
	12	10	8	0.0764	1.8893	0.7602	0.5544	0.3622	0.2589	0.3713	0.4796
Rayleigh	6	4	2	0.0564	0.779	0.9591	0.8709	0.7478	0.1183	0.1243	0.1803
	8	6	4	0.0670	0.8431	0.9561	0.8556	0.7461	0.0775	0.1445	0.2115
	10	8	6	0.0761	0.5617	0.9243	0.7976	0.6380	0.1093	0.1853	0.2613
	12	10	8	0.1035	0.7800	0.9246	0.7730	0.5828	0.1273	0.2308	0.3343

Table 2. Bayes estimation of model parameters, the reliability and the hazard functions according to  $SELF$  based  $DGOS$ . ( $a = 0.5$ )

Distribution	$N$	$n$	$m$	$k$	$\hat{\beta}_{BS}$	$\hat{\theta}_{BS}$	$\hat{R}_{BS}$		$\hat{H}_{BS}$	
							$t = 3$	$t = 6$	$t = 3$	$t = 6$
Weibull	10	6	4	2	0.1042	2.0361	0.4064	0.3093	0.6254	0.4270
	15	8	6	4	0.1119	0.16970	0.4909	0.4971	0.4602	0.1190
	20	10	8	6	0.1262	1.0002	0.6469	0.5244	0.1264	0.1093
	25	12	10	8	0.1057	1.0107	0.7263	0.5651	0.1146	0.0962
Rayleigh	10	6	4	2	0.1941	2.0004	0.9078	0.7016	0.1941	0.2226
	15	8	6	4	0.2994	1.4814	0.6632	0.5969	0.5404	0.2811
	20	10	8	6	0.3785	1.0000	0.4753	0.8636	0.7571	0.0975
	25	12	10	8	0.3429	0.9901	0.5080	0.8179	0.6859	0.1499

Table 3. Bayes estimation of model parameters, the reliability and the hazard functions according to  $LLF$  based  $DGOS$ . ( $a = 0.5$ )

Distribution	$N$	$n$	$m$	$k$	$\hat{\beta}_{BL}$	$\hat{\theta}_{BL}$	$\hat{R}_{BL}$		$\hat{H}_{BL}$	
							$t = 3$	$t = 6$	$t = 3$	$t = 6$
Weibull	10	6	4	2	0.1039	2.0084	0.9252	0.9746	0.7166	0.4089
	15	8	6	4	0.1106	1.6411	0.8898	0.9328	0.4450	0.1188
	20	10	8	6	0.1260	1.0001	0.9079	0.9115	0.1262	0.1091
	25	12	10	8	0.1056	1.0068	0.9210	0.8752	0.1117	0.0960
Rayleigh	10	6	4	2	0.1932	2.0002	0.8609	0.8490	0.1932	0.2159
	15	8	6	4	0.2903	1.4200	0.9143	0.8651	0.4940	0.2778
	20	10	8	6	0.3769	1.0000	0.9099	0.8942	0.7503	0.0980
	25	12	10	8	0.3419	1.0310	0.9203	0.9513	0.6815	0.1385

## References

- [1] Ahsanullah, M. (2004). A characterization of the uniform distribution by dual generalized order statistics. *Communications in Statistics–Theory and Methods*, 2921-2928.
- [2] Arslan, G. (2011). On a characterization of the uniform distribution by generalized order statistics. *Journal of computational and applied mathematics*, **235**(16), 4532-4536.
- [3] Barakat, H. M., El-Adll, M. E. (2009). Asymptotic theory of extreme dual generalized order statistics. *Statistics Probability Letters*, **79**(9), 1252-1259.
- [4] Jaheen, Z. F., Al Harbi, M. M. (2011). Bayesian estimation based on dual generalized order statistics from the exponentiated Weibull model. *Journal of Statistical Theory and Applications*, **10**(4), 591-602.
- [5] KAMPS, M. B. E. C. U., Cramer, E., Burkschat, M. (2003). Dual generalized order statistics. *Metron*, **61**(1), 13-26.
- [6] Soland, R. M. (1969). Bayesian analysis of the Weibull process with unknown scale and shape parameters. *IEEE Transactions on Reliability*, **18**(4), 181-184.
- [7] Soliman, A. A. (2002). Reliability estimation in a generalized life-model with application to the Burr-XII. *IEEE Transactions on Reliability*, **51**(3), 337-343.
- [8] Tavangar, M. (2011). Power function distribution characterized by dual generalized order statistics. *Journal of Iranian Statistical Society*, **10**(1), 13-27.
- [9] Tavangar, M., Asadi, M. (2008). On a characterization of generalized Pareto distribution based on generalized order statistics. *Communications in Statistics–Theory and Methods*, **37**(9), 1347-1352.
- [10] Wang, L., Shi, Y. (2013). Reliability analysis of a class of exponential distribution under record values. *Journal of Computational and Applied Mathematics*, **239**, 367-379.



# Optimal step stress test plan based on bivariate gamma degradation model

Shemehsavar , S. <sup>1</sup>      Mosaebi Omshi , E. <sup>2</sup>

<sup>1,2</sup> School of Mathematics, Statistics and Computer Science, University of Tehran

## Abstract

Modern products usually have complex structure, situation where the failure of the test unit is not related deterministically to an observable degradation process. Step Stress Accelerated Degradation Test(SSADT) is applicable for assessing the lifetime distribution of high reliable products when the number of test units is relatively few. In this paper, we introduce SSADT model when the degradation path is latent and follow a Gamma process, so a bivariate Gamma process is considered in which the first component to be latent degradation process and the second component represent a marker process. For such a pattern, the optimal hold time is obtained at which the stress level is changed . The optimization criterion is to minimize the approximate variance of the estimated MTTF of the lifetime distribution of the products at the stress changing time.

**Keywords:** bivariate Gamma process, Step Stress Accelerated Degradation Test, Fisher information matrix, Optimal design, Marker process.

## 1 Introduction

Due to the rapid improvement of the high technology, the products today become more reliable and the products' life gets longer. It might take a long time, such as several years, for a product to fail, which makes it difficult or even impossible to obtain the failure information under usage condition for such highly reliable products. In comparison to

<sup>1</sup>shemehsavar@khayam.ut.ac.ir

traditional life testing, even under accelerated conditions, the degradation modeling will yields more useful data sooner. For highly reliable parts or systems, where failures are rare or take along time to occur, this may be the only possible approach. The rate of performance degradation under use condition is so slow that it is difficult to make useful inferences about reliability in a reasonable amount of time. In such cases, products are often subjected to elevated stress level in order to accelerate the degradation test. Such an experiment is called an accelerated degradation test(ADT). Tang and Chang[7] modeled nondestructive accelerated degradation data as a collection of stochastic processes for which the parameters depend on the stress levels. Whitmore and Schenkelberg[8] modeled accelerated degradation process by a Brownian motion with a time scale transformation. Although ADT is an efficient life-test method, it is usually very expensive to conduct. For a newly developed product or an expensive product, it may not be possible to have so many test units on hand. Moreover, the selection of suitable levels of stress is not straightforward. In this situation, a constant-stress ADT is therefore not suitable. To handle these problem, step stress accelerated degradation tests(SSADT) were proposed by Tseng and Wen[5] in a case study of LEDs.

In a SSADT experiment, each product is first tested, subject to a pre-determined stress level for a specified length of time. If it does not fail, it is tested again at a higher stress level for another specified length of time. The stress on the specimen is thus increased step by step until an appropriate termination time is reached. Obviously, the advantage of the SSADT is that only a few test units are needed for conducting this life test. To conduct a SSADT efficiently, special attention needs to be paid to the above mentioned design variables (sample size, measurement frequency, and termination time). The problem of optimizing the test design have been extensively studied in recent years. Three commonly used optimization criteria are the minimum approximated variance(Avar) of the maximum likelihood estimation(MLE) of reliability, mean time to failure(MTTF) and the quantiles of the lifetime distribution. Liao & Tseng[3] formulated a SSADT model with a Wiener process, and discussed such an optimal design problem. Tseng et al

## 2 Model description

Two random variables  $X$  and  $Y$  are said to have Kibble's bivariate gamma distribution[2] with shape parameter  $\alpha$ , scale parameters  $\lambda_1, \lambda_2$  and correlation  $\rho$ , if their joint density function is considered as follows:

$$f(x, y|\alpha, \rho) = \frac{(\lambda_1 \lambda_2)^\alpha}{(1 - \rho)\Gamma(\alpha)} \left(\frac{xy}{\rho \lambda_1 \lambda_2}\right)^{\left(\frac{\alpha-1}{2}\right)} \exp\left(-\frac{\lambda_1 x + \lambda_2 y}{1 - \rho}\right) I_{\alpha-1}\left(\frac{2\sqrt{\lambda_1 \lambda_2 \rho xy}}{1 - \rho}\right), \quad (1)$$

where  $x, y > 0$ ,  $0 \leq \rho < 1$ ,  $\alpha, \lambda_1, \lambda_2 > 0$ , and  $I_\alpha(\cdot)$  is the modified Bessel function of the first kind of order  $\alpha$  defined as

$$I_\alpha(\nu) = \sum_{k=0}^{\infty} \frac{\left(\frac{\nu}{2}\right)^{2k+\alpha}}{\Gamma(k+\alpha+1)k!}, \quad \nu > 0$$

Consider a two-dimensional gamma process,  $\{(X(t), Y(t)), t \geq 0\}$  with initial value  $\{(X(0), Y(0))\} = \{(0, 0)\}$ , such that the vector  $(X(t), Y(t))$ , has Kibble's bivariate gamma

distribution with shape parameter  $\alpha t$ , scale parameters  $\lambda_1, \lambda_2$  and correlation  $\rho$ . Every item is represented by a path of bivariate Gamma process  $\{(X(t), Y(t)), t \geq 0\}$ . We consider the component  $\{X(t), t \geq 0\}$  to be the latent degradation process which represents the level of degradation of a item at time  $t$ . The component  $\{Y(t), t \geq 0\}$  represents a marker process and  $Y(t)$  has a Gamma distribution with shape parameter  $\alpha t$  and scale parameter  $\lambda_2$ . The item fails when the degradation process  $\{X(t), t \geq 0\}$  reaches a failure threshold  $c > 0$  for the first time and we denote this first-hitting time by a random variable  $U$ . Now, consider this bivariate process to model a typical simple SSADT problem. Under a simple SSADT, each item is first tested, subject to a stress level  $S_1$  for a specified length of time  $[0, \tau]$ , where the pre-specified value  $0 < \tau < T$  is the time when the experimenter changes the stress level  $S_1$  to the higher level  $S_2$ , which we call it the *stress changing time*. If the item does not fail, it is tested again at a higher stress level  $S_2$  for another specified length of time  $[\tau, T]$ . The stress level of the experiment is then as

$$S = \begin{cases} S_1 & 0 \leq t < \tau \\ S_2 & \tau \leq t < T, \end{cases}$$

for which  $(S_0 < S_1 < S_2)$ , where  $S_0$  is the used stress level normal conditions.

So we focuses on the situation where marker measurements are taken only at the failure times, and assume each item is observed during a fixed period  $[0, T]$ . For the aforementioned step stress degradation plan we have bivariate gamma degradation process under the stress  $S_j$ ,  $j = 0, 1, 2$ , so that

$$(X(t|S_j), Y(t|S_j)) \sim BKGD(\alpha_j, \lambda_1, \lambda_2, \rho).$$

We define  $X(t|S_j) := X^j$ .

The following Arrhenius reaction rate model can be used to model the relationship between  $\alpha_j$ , and the stress  $S_j$ ,

$$\alpha_j = \exp(a + \frac{b}{273 + S_j}), \quad j = 0, 1, 2$$

For example, we can assume that temperature is an accelerating variable.

First consider any sample path of the  $X^j$  component over a time interval  $[0, T]$  and partition this sample path at the arbitrary time points  $0 = t_0 < t_1 < \dots < t_k = T$ , where  $k$  is a natural number. Now let  $\Delta t_i = t_i - t_{i-1}$  and  $\Delta x_i = x_i - x_{i-1}$ , where  $x_i = x(t_i)$ , is the realized increment for component  $X^j$  during the time interval  $(t_{i-1}, t_i]$  for  $i = 1, \dots, k$ . We refer to  $P = \{\Delta x_1, \Delta x_2, \dots, \Delta x_k\}$  as a sample path partition.

If  $\Delta Y_i^j = Y_i^j - Y_{i-1}^j$ , where  $Y_i^j = Y^j(t_i)$ , denotes the increment for the  $Y^j$  component in the  $i$ th interval. So the conditional random variable  $Y^j(t)|P$  has Randomized Gamma Distribution(RGD),

$$Y^j(t)|P \sim RGD(\alpha_j t + \eta, \frac{\lambda_2}{1 - \rho}) \quad (2)$$

Therefore, the conditional distribution of  $Y^j(t)|P$  is the same as the conditional distribution of  $Y^j(t)|x(t)$ ,

$$Y^j(t)|x(t) \sim RGD(\alpha_j t + \eta, \frac{\lambda_2}{1 - \rho}), \quad (3)$$

where,  $\eta \sim Poisson(\frac{\rho\lambda_1x(t)}{1-\rho})$ . Thus, for a failing path, the pdf of  $Y^j(u)|x(u)$ , where  $U = u$ , follows the equation (2.3) with  $t$  replaced by  $u$  and  $x(u) = c$ ,

$$P_{1j}(y|u) = P(Y^j(u) = y|U = u) = \frac{\frac{\alpha_j u + 1}{2} y^{\frac{\alpha_j u - 1}{2}} \exp(-\frac{\rho\lambda_1 c + \lambda_2 y}{1-\rho})}{\lambda_1^{\frac{\alpha_j u - 1}{2}} c^{\frac{\alpha_j u - 1}{2}} (1-\rho)\rho^{\frac{\alpha_j u - 1}{2}}} I_{\alpha_j u - 1}(\frac{2\sqrt{\rho\lambda_1\lambda_2 c y}}{1-\rho}), \quad (4)$$

Now we obtain the distribution for the failure time  $U$  of a failing item:

Suppose that an item fails when this component first reaches a threshold  $c > 0$ . We denote this first passage time to the threshold by random variable  $U$ , which is

$$U = \inf\{t|X^j(t) \geq c\}$$

We assumed  $U$  to be the first-hitting time of the gamma process  $X^j(t)$  to the threshold  $c$ , where  $X^j(t)$  is distributed as a gamma distribution with shape parameter  $\alpha_j t$  and scale parameter  $\lambda_1$ . Since  $X^j(t)$  is strictly increasing in  $t$ , we have  $\max_{0 \leq u \leq t} X^j(u) = X^j(t)$ . Hence, the survival probability for an item beyond the fixed time  $t$  is

$$P(U > t) = P(\max_{0 \leq u \leq t} X^j(u) \leq c) = P(X^j(t) \leq c) = \int_0^c \frac{\lambda_1^{\alpha_j} x^{\alpha_j t - 1} \exp(-\lambda_1 x)}{\Gamma(\alpha_j t)} dx$$

Also the failure probability for an item by the fixed time  $t$  is

$$P(U \leq t) = \frac{\Gamma(\alpha_j t, c\lambda_1)}{\Gamma(\alpha_j t)},$$

then

$$P_{2j}(u) = \frac{d}{du} \frac{\Gamma(\alpha_j u, c\lambda_1)}{\Gamma(\alpha_j u)}$$

The pdf of  $U$  is so complicated and it is very difficult to compute in practice, the distribution of  $U$  is approximated closely by Birnbaum-Saunders distribution, Park and Padgett[4],(Birnbaum and Saunders[1]) with parameters  $\alpha^* = \frac{1}{\sqrt{c\lambda_1}}$  and  $\beta^* = \frac{c\lambda_1}{\alpha_j}$ .

$$P_{2j}(u) \approx \frac{1}{2\alpha^*\beta^*\sqrt{2\pi}} [(\frac{u}{\beta^*})^{-1/2} + (\frac{u}{\beta^*})^{-3/2}] \exp(-\frac{(\frac{u}{\beta^*} - 2 + \frac{\beta^*}{u})}{2\alpha^{*2}}), \quad 0 < t < \infty \quad (5)$$

Now we can obtain the pdf for failing item.

The joint pdf of  $Y^j(u)$  and  $U$  equals

$$P_{F_j}(y, u) = P_{1j}(y|u)P_{2j}(u), \quad j = 1, 2, \quad (6)$$

Suppose that  $n$  items are subjected to a simple SSADT over the observation period  $(0, T]$  under independent conditions. By assumption that  $n$  items have failed, there are two vectors of length of  $n$ , as follows:

- a vector of failing times  $\mathbf{u} = (u_1, \dots, u_n)$ ,
- a vector of values of the Y-component of the bivariate gamma process at failing times  $\mathbf{y} = (y_1, \dots, y_n)$ .

The latent degradation component  $X$  is observed for failed items and is equal to the

failure threshold  $c$ . We denote the sample failing observations by  $(y_i, u_i)$ ,  $i = 1, \dots, \nu_1$  for the stress level  $S_1$ , and  $(y_i, u_i)$ ,  $i = \nu_1 + 1, \dots, n$  for the stress level  $S_2$ , the sample log-likelihood is given by

$$\log L(\theta) = \sum_{i=1}^{\nu_1} \log P_{F_1}(y_i, u_i; \theta) + \sum_{i=\nu_1+1}^n \log P_{F_2}(y_i, u_i; \theta). \quad (7)$$

Under stress levels  $S_j$ ,  $j = 1, 2$ , we have

$$P_{F_j}(y_i, u_i; \theta) = \frac{\lambda_2(\gamma_{ji} + \lambda_1 c)}{2(1-\rho)u_i\sqrt{2\pi\gamma_{ji}}} (\beta_i)^{\frac{\gamma_{ji}-1}{2}} e^{(-[\frac{(\gamma_{ji}-\lambda_1 c)^2}{2\gamma_{ji}} + \frac{\rho\lambda_1 c + \lambda_2 y_i}{1-\rho}])} I_{\gamma_{ji}-1}(\frac{2\sqrt{\rho\lambda_1\lambda_2 c y_i}}{1-\rho}), \quad (8)$$

where,  $\theta = (a, b, \lambda_1, \lambda_2, \rho)$ , and

$$\gamma_{1i} = \alpha_1 t_i, \quad \gamma_{2i} = \alpha_1 \tau + \alpha_2(t_i - \tau), \quad \beta_i = \frac{\lambda_2 y_i}{\lambda_1 \rho c}. \quad (9)$$

### 3 The optimal design

The optimization design considered in this article is to find the optimal stress changing time  $0 < \tau^* < T$  which minimizes the approximate variance of the estimated MTTF of the lifetime distribution of the product in normal conditions.

By  $\delta$ -method, the approximate variance of the MTTF<sub>0</sub> is found to be

$$AVar(\widehat{MTTF}_0) = H' I^{-1}(\hat{\theta}) H,$$

where

$$H' = \left( -\left( \frac{c\lambda_1}{\alpha_0} + \frac{1}{2\alpha_0} \right), \frac{-1}{273 + S_0} \left( \frac{c\lambda_1}{\alpha_0} + \frac{1}{2\alpha_0} \right), \frac{-c}{\alpha_0}, 0, 0 \right),$$

$H'$  denotes the transpose of  $H$  and  $I(\theta)$  is the Fisher information matrix of  $\theta$ . In order to calculate the estimate of the fisher information matrix of data about  $\theta$ ,  $I(\theta) = I_{r,s}(\theta)$ . First let  $\theta = (a, b, \lambda_1, \lambda_2, \rho) = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)$ , for  $(r, s) \in \{1, \dots, 5\}$  then we have

$$\hat{I}_{r,s}(\hat{\theta}) = E\left(\frac{-\partial^2 \log L(\theta)}{\partial \theta_r \partial \theta_s}\right)|_{\theta}$$

Park and Padgett[4] proposed a simple approximate formula for the MTTF(Mean Time To Failure) of the product under typical use stress  $S_0$  as

$$MTTF_0 = E(U) = \frac{c\lambda_1}{\alpha_0} + \frac{1}{2\alpha_0}, \quad (10)$$

The estimated MTTF under  $S_0$ ,  $\widehat{MTTF}_0$ , can be obtained by substituting  $\hat{a}, \hat{b}, \hat{\lambda}_1$  and  $S_0$  into (3.1) and  $\alpha_0 = \exp(a + \frac{b}{273+S_0})$  directly.

Hence before performing the optimization procedure, one should estimate the parameter vector  $\theta$  using a lifetime data in normal conditions. An approach for such an estimation

is the ML estimation using the likelihood of the data under the use stress level  $S_0$ . The log-likelihood of parameters  $\theta_0 = (\alpha_0, \lambda_1, \lambda_2, \rho)$  is equal to

$$\log L(\theta_0) = \sum_{i=1}^n \log P_{F_0}(y_i, t_i; \theta_0), \quad (11)$$

The likelihood (3.2) can be used along with the equation  $\alpha_j = \exp(a + \frac{b}{273+S_j})$  to estimate the parameters  $\theta = (a, b, \lambda_1, \lambda_2, \rho)$  based on a lifetime data collected in normal conditions. Avar of the ML estimate of the reliability of the products in normal conditions. This Avar function is a function of the stress changing time  $\tau$  and the parameter vector  $\theta$ .

## 4 Case study

We use aluminum reduction cells data from Whitmore, Crowder and Lawless[9]. Aluminum is produced by electrolysis of molten alumina and cryolite in reduction cells, consisting principally of a carbon anode and a carbon-lined steel box which acts as the cell's cathode. These cells are subjected to several thermal, chemical and mechanical stresses. In this paper, their degradation is marked by physical distortion of the steel box from cracking of the carbon lining. Table(7.1) in Whitmore et al[9] article, they present marker and failure data for 17 reduction cells of a particular design that were operated to failure under uniform conditions in a canadian aluminum smelter. For these data the threshold is taken to be  $c = 1$ . Using these data, one can obtain the ML estimates of the parameters using the likelihood in Whitmore et al[9] as  $a = 5.211$ ,  $b = -3221.99$ ,  $\rho = 0.675$ ,  $\lambda_1 = 3.7026$  and  $\lambda_2 = 1.2151$ . So, we derive  $\alpha_0 = 0.0044167$ ,  $\alpha_1 = 0.011512$  and  $\alpha_2 = 0.025611$ .

We performed the optimization process as mentioned in the previous section using the optimization procedures of software R.2.14.1 and obtained the value,  $\tau^* = 5.899024$ .

## References

- [1] Birnbaum, Z.W. and Saunders, S.C.(1969), A new Family of Life Distribution, *Journal of Applied Probability*, **6**, 319-327.
- [2] Balakrishnan, N. and Lai, C.D. (2007), *Continuous Bivariate Distributions*, Springer, Second Edition .
- [3] Liao, C.M. and S. T. Tseng, S.T.(2006), Optimal design for step-stress accelerated degradation tests, *IEEE Trans. Reliability*, **55**, 59-66.
- [4] Park, C. and Padgett , W.J. (2005), Accelerated Degradation Models for Failure based on GeometricBrownian Motion and Gamma Processes, *Lifetime Data Analysis*, **11**, 511-527.
- [5] Tseng, S.T. and Wen, Z.C. (2000), Step-Stress Accelerated Degradation Analysis of High-Reliable Products, *Journal of Quality Technology*, **32**, 209216.

- [6] Tseng, S.T., Balakrishnan, N and Tsai, C.C (2009), Optimal Step-Stress Accelerated Degradation Test Plan for Gamma Degradation Process, *IEEE Transactions on Reliability* , **58**, 611618.
- [7] Tang, L.C. and Chang, D.S. (1995), Reliability prediction using nondestructive accelerateddegradation data: case study on power supplies, *IEEE Transactions on Reliability*, **44**, 562-566.
- [8] Whitmore, G.A. and Schenkelberg,F. (1997), Modelling Accelerated Degradation Data using Wiener Diffusion with a time Scale Transformation, *Lifetime Data Analysis*, **3**, 119.
- [9] Whitmore, G.A., Crowder, M.J. and Lawless, J.F. (1998), Failure Inference from a Marker Process based on Bivariate Wiener Model, *Lifetime Data Anal*, , .



Forth Seminar on  
**Reliability Theory and its Applications**

April 25-26, 2018



## A new class of bivariate failure time distributions in shock and competing risk models

Shoaee, Sh.<sup>1</sup>

Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University,  
Tehran

### Abstract

In this paper, we introduce a new class of bivariate distributions. This new class of bivariate distributions contains several bivariate lifetime models and is more flexible. We call this new class of distributions as the bivariate proportional hazard rate-geometric (BPHRG) models. We investigate various properties of the new class of distributions. We propose to use the EM algorithm to compute the maximum likelihood estimators of the unknown parameters, and it is computationally quite tractable. We also evaluate the performance of the EM algorithm on a real data set. Finally, we compare BPHRG models to BPHR models.

**Keywords:** Bivariate model, EM algorithm, Proportional hazard rate model, Pseudo likelihood function, Monte Carlo simulation.

## 1 Introduction

The modeling of a lifetime is an important problem in a variety of scientific and technological fields. In this area, many new univariate distributions have been studied in the statistical literature recently.

The new univariate distributions using compounding continuous with discrete distributions have been introduced and studied in the recent years. Marshall and Olkin [7] introduced a class of univariate distributions which can be obtained by minimum and

<sup>1</sup>Sh\_Shoaee@sbu.ac.ir

maximum of independent and identically distributed continuous random variables, where the sample size follows the Geometric distribution. In fact, their method induces an extra parameter to a model, hence affords more flexibility. They discussed the various properties in case of Exponential and Weibull distributions. Also, extensive work has been done on their method and it is extended to some other lifetime distributions. For more details, the readers can refer to Silva et al. [9], Ghitany et al. [5], [6], Pham and Li [8], Barreto-Souza [2] and the references cited therein.

Interestingly, not much work has been done on the bivariate distributions mainly due to its analytical intractability. Therefore, we introduce a new class of bivariate distributions by the method of Marshall and Olkin [7].

In this paper, we let  $Z$  be a continuous random variable with the cumulative distribution function (CDF)  $F_{PHRM}(z; \alpha, \lambda)$ . Many lifetime distributions such as Exponential, Pareto Type I, Lomax, Burr Type XII, Weibull, Gompertz and so on can be represented in the form of

$$F_{PHRM}(z; \theta, \alpha) = 1 - [\bar{F}_B(z, \alpha)]^\lambda, \quad -\infty \leq b_1 < z < b_2 \leq \infty, \quad \lambda > 0. \quad (1)$$

Where  $\bar{F}_B(\cdot) = 1 - F_B(\cdot)$  is the baseline survival function with  $F_B(b_1) = 0$  and  $F_B(b_2) = 1$ . This family of distributions is called proportional hazard rate model (PHRM) and was originally proposed by Cox [3].

The proportional hazard rate model (1) can be widely used in reliability applications because it has many different forms of its reliability and hazard functions. By choosing the proper form of  $F_B(z, \theta)$ , it can be applied in many areas of quality control, survival, insurance and failure time modeling.

Therefore, the main aim of this paper is to introduce a class of bivariate distributions by compounding Geometric distribution and proportional hazard rate model. This class contains the family of bivariate proportional hazard rate distributions (BPHR) and is called the bivariate proportional hazard rate-geometric (BPHRG) class of distributions. In fact, this method produces a new class of bivariate distributions which are analytically quite tractable. Also, the marginals and conditionals are univariate proportional hazard rate-geometric distributions (UPHRG). So, different properties of this new class have been investigated. They are very flexible models, and the joint probability distribution function can take different shapes.

The maximum likelihood estimators (MLEs) of the unknown parameters of the BPHRG class of distributions cannot be obtained in closed forms. The Newton-Raphson or Gauss-Newton type algorithm iterative procedure is needed to solve these non-linear equations. Moreover, the choice of initial values and the convergence of the iterative algorithm are important subject. To avoid these problems, we investigate it as a missing value problem, and propose to use the expectation maximization (EM) algorithm to compute the MLEs. The paper is organized as follows. In Section 2, a family of the bivariate proportional hazard rate distributions and the structure of bivariate proportional hazard rate-geometric model is introduced. The statistical inference of the unknown parameters is presented in Section 3. In Section 4, the results of the simulation and the analysis of real data set are provided. Finally, we conclude the paper in Section 5.

## 2 Model formulation

In this section, we will introduce a family of the bivariate proportional hazard rate distributions and the bivariate proportional hazard rate-geometric distributions. Also, this new classes can be applied in shock or in competing risks models.

### 2.1 Bivariate proportional hazard rate distribution

Suppose  $U_i \sim PHRM(\alpha, \lambda_i)$ , for  $i = 0, 1, 2$  and they are independent. Define  $X_1 = \min\{U_0, U_1\}$ , and  $X_2 = \min\{U_0, U_2\}$ , then the bivariate vector  $(X_1, X_2)$  is a family of the bivariate proportional hazard rate distributions with the parameters  $\alpha, \lambda_0, \lambda_1, \lambda_2$  and it will be denoted from now on as  $BPHR(\alpha, \lambda_0, \lambda_1, \lambda_2)$ .

If  $(X_1, X_2) \sim BPHR(\alpha, \lambda_0, \lambda_1, \lambda_2)$  then, for  $z = \max\{x_1, x_2\}$

$$\begin{aligned}\bar{F}_{X_1, X_2}(x_1, x_2) &= P(X_1 \geq x_1, X_2 \geq x_2) = P(U_1 \geq x_1, U_2 \geq x_2, U_0 \geq z) \\ &= \begin{cases} \bar{F}_{PHRM}(x_1, \alpha, \lambda_1 + \lambda_0)\bar{F}_{PHRM}(x_2, \alpha, \lambda_2) & \text{if } x_2 < x_1 \\ \bar{F}_{PHRM}(x_1, \alpha, \lambda_1)\bar{F}_{PHRM}(x_2, \alpha, \lambda_2 + \lambda_0) & \text{if } x_1 < x_2 \\ \bar{F}_{PHRM}(x, \alpha, \lambda_0 + \lambda_1 + \lambda_2, \lambda) & \text{if } x_1 = x_2 = x. \end{cases}\end{aligned}$$

The joint PDF

$$f_{X_1, X_2}(x_1, x_2) = \begin{cases} f_{PHRM}(x_1, \alpha, \lambda_1 + \lambda_0)f_{PHRM}(x_2, \alpha, \lambda_2) & \text{if } x_2 < x_1 \\ f_{PHRM}(x_1, \alpha, \lambda_1)f_{PHRM}(x_2, \alpha, \lambda_2 + \lambda_0) & \text{if } x_1 < x_2 \\ \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2}f_{PHRM}(x, \alpha, \lambda_0 + \lambda_1 + \lambda_2) & \text{if } x_1 = x_2 = x. \end{cases} \quad (2)$$

### 2.2 Bivariate proportional hazard rate-geometric distributions

Now, suppose  $\{(X_{1n}, X_{2n}); n = 1, 2, \dots\}$  is a sequence of i.i.d. non-negative bivariate random variables with common joint distribution function  $F_{\mathbf{X}}(., .)$  where  $\mathbf{X} = (X_1, X_2)$  and  $N$  is a Geometric random variable independent of  $\{(X_{1n}, X_{2n}), n = 1, 2, \dots\}$ . Consider the following bivariate random variable  $\mathbf{Y} = (Y_1, Y_2)$ , so that,  $Y_i = \min\{X_{i1}, \dots, X_{iN}\}$ , for  $i = 1, 2$ . Therefore, the joint survival function of  $(Y_1, Y_2)$  becomes:

$$\bar{G}(y_1, y_2) = P(Y_1 > y_1, Y_2 > y_2) = \frac{\theta \bar{F}(y_1, y_2)}{1 - (1 - \theta) \bar{F}(y_1, y_2)}. \quad (3)$$

Therefore, the random variable  $\mathbf{Y} = (Y_1, Y_2)$  is said to have a bivariate proportional hazard rate-geometric distribution with parameters  $\theta, \alpha, \lambda_0, \lambda_1, \lambda_2$ , if the distribution  $F$  in (3) is  $BPHR(\alpha, \lambda_0, \lambda_1, \lambda_2)$ . Therefore, the joint survival function of  $(Y_1, Y_2)$  becomes:

$$\bar{G}(y_1, y_2) = \begin{cases} \frac{\theta \bar{F}_{PHRM}(y_1, \alpha, \lambda_0 + \lambda_1)\bar{F}_{PHRM}(y_2, \alpha, \lambda_2)}{1 - (1 - \theta) \bar{F}_{PHRM}(y_1, \alpha, \lambda_0 + \lambda_1) \bar{F}_{PHRM}(y_2, \alpha, \lambda_2)} & \text{if } y_2 \leq y_1 \\ \frac{\theta \bar{F}_{PHRM}(y_1, \alpha, \lambda_1)\bar{F}_{PHRM}(y_2, \alpha, \lambda_0 + \lambda_2)}{1 - (1 - \theta) \bar{F}_{PHRM}(y_1, \alpha, \lambda_1) \bar{F}_{PHRM}(y_2, \alpha, \lambda_0 + \lambda_2)} & \text{if } y_1 < y_2. \end{cases} \quad (4)$$

It will be denoted by  $(Y_1, Y_2) \sim BPHRG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$ .

**Proposition 2.1.** *Let  $F_{\mathbf{Y}}(y_1, y_2)$  be the joint survival function of BPHRG distribution given in (4). Then*

1. Each  $Y_i$  has an univariate proportional hazard rate-geometric distribution (UPHRG) with parameters  $\alpha, \lambda_0 + \lambda_i$  and  $\theta$ .
2. The random variable  $Y = \min(Y_1, Y_2)$  has an UPHRG distribution with parameters  $\lambda_0 + \lambda_1 + \lambda_2, \alpha$  and  $\theta$ .
3.  $P(Y_1 < Y_2) = \frac{\lambda_1}{\lambda_0 + \lambda_1 + \lambda_2}$ .

**Theorem 2.1.** Let  $(Y_1, Y_2) \sim BPHRG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$ , then the joint PDF of  $(Y_1, Y_2)$  is:

$$g_{Y_1, Y_2}(y_1, y_2) = \begin{cases} g_1(y_1, y_2) & \text{if } y_2 < y_1 \\ g_2(y_1, y_2) & \text{if } y_1 < y_2 \\ g_0(y_1, y_2) & \text{if } y_1 = y_2 = y. \end{cases}$$

where

$$\begin{aligned} g_1(y_1, y_2) &= \frac{\theta f_{PHRM}(y_1, \alpha, \lambda_0 + \lambda_1) f_{PHRM}(y_2, \alpha, \lambda_2)}{[1 - (1 - \theta) \bar{F}_{PHRM}(y_1, \alpha, \lambda_0 + \lambda_1) \bar{F}_{PHRM}(y_2, \alpha, \lambda_2)]^3} \\ &\quad \times \{1 + (1 - \theta) \bar{F}_{PHRM}(y_1, \alpha, \lambda_0 + \lambda_1) \bar{F}_{PHRM}(y_2, \alpha, \lambda_2)\}. \\ g_2(y_1, y_2) &= \frac{\theta f_{PHRM}(y_1, \alpha, \lambda_1) f_{PHRM}(y_2, \alpha, \lambda_1 + \lambda_2)}{[1 - (1 - \theta) \bar{F}_{PHRM}(y_1, \alpha, \lambda_1) \bar{F}_{PHRM}(y_2, \alpha, \lambda_0 + \lambda_2)]^3}, \\ &\quad \times \{1 + (1 - \theta) \bar{F}_{PHRM}(y_1, \alpha, \lambda_1) \bar{F}_{PHRM}(y_2, \alpha, +\lambda_0 + \lambda_2)\}. \\ g_0(y) &= \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} \times \frac{\theta f_{PHRM}(y, \alpha, \lambda_0 + \lambda_1 + \lambda_2)}{[1 - (1 - \theta) \bar{F}_{PHRM}(y, \alpha, \lambda_0 + \lambda_1 + \lambda_2)]^2}. \end{aligned}$$

Using the conditional probability mass function of  $N$  given  $Y_1 = y_1$  and  $Y_2 = y_2$ , we can compute:

$$E(N|y_1, y_2) = \begin{cases} \frac{(1 - \xi_1(y_1, y_2, \theta, \gamma))^2 - 6(1 - \xi_1(y_1, y_2, \theta, \gamma)) + 6}{(1 - \xi_1(y_1, y_2, \theta, \gamma))^2} & \text{if } y_2 < y_1 \\ \frac{(1 - \xi_2(y_1, y_2, \theta, \gamma))^2 - 6(1 - \xi_2(y_1, y_2, \theta, \gamma)) + 6}{(1 - \xi_2(y_1, y_2, \theta, \gamma))^2} & \text{if } y_1 < y_2 \\ \frac{1 + \xi_0(y_1, y_2, \theta, \gamma)}{1 - \xi_0(y_1, y_2, \theta, \gamma)} & \text{if } y_1 = y_2 = y. \end{cases}$$

### 3 Estimation

In this section, we describe the problem of computing the MLEs of the unknown parameters of the BPHRG distributions using the EM algorithm.

#### 3.1 EM algorithm

Suppose  $\{(y_{11}, y_{21}), \dots, (y_{1m}, y_{2m})\}$  is a random sample from BPHRG with parameters  $\Theta = (\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$ . We define the following notation:

$I_0 = \{i : y_{1i} = y_{2i} = y_i\}$ ,  $I_1 = \{i : y_{1i} > y_{2i}\}$  and  $I_2 = \{i : y_{1i} < y_{2i}\}$ . Also,  $|I_0| = m_0$ ,  $|I_1| = m_1$ ,  $|I_2| = m_2$  and  $m = m_0 + m_1 + m_2$ . Therefore, the log-likelihood function can be written as:

$$\ell(\Theta) = \sum_{i \in I_0} \ln g_0(y_i) + \sum_{i \in I_1} \ln g_1(y_{1i}, y_{2i}) + \sum_{i \in I_2} \ln g_2(y_{1i}, y_{2i}), \quad (5)$$

where  $g_0$ ,  $g_1$  and  $g_2$  are defined in Theorem 2.1. We can obtain the MLEs of the parameters by maximizing  $\ell(\Theta)$  in (5) with respect to the unknown parameters. Clearly, it is difficult to compute the MLEs of the unknown parameters directly. We propose to use the EM algorithm and treat this as a missing value problem.

For given  $n$ , consider that independent random variables

$$\{U_i | N = n\} \sim GP(\alpha, n\lambda_i), \quad i = 0, 1, 2. \quad (6)$$

Also, It is well known that

$$\{Y_1 | N = n\} = \min\{U_0, U_1\} | N = n, \quad \text{and} \quad \{Y_2 | N = n\} = \min\{U_0, U_2\} | N = n.$$

Assumed that for the bivariate random vector  $(Y_1, Y_2)$ , there is an associated random vectors

$$(\Delta_1, \Delta_2) = \begin{cases} (0, 0) & \text{if } Y_1 = U_0, Y_2 = U_0 \\ (0, 1) & \text{if } Y_1 = U_0, Y_2 = U_2 \\ (1, 0) & \text{if } Y_1 = U_1, Y_2 = U_0 \\ (1, 1) & \text{if } Y_1 = U_1, Y_2 = U_2 \end{cases}. \quad (7)$$

Here  $Y_i$ 's are same as defined above. Therefore, a sample is obtained from  $(Y_1, Y_2, \Delta_1, \Delta_2, N)$  which is the complete observation. It is clear that, if we know  $(Y_1, Y_2)$ , the associated  $(\Delta_1, \Delta_2)$  may not always be known. We compute the pseudo log-likelihood function. The conditional 'pseudo' log-likelihood function is formed by conditioning on  $N$ , and then  $N$  is replaced by  $E(N|Y_1, Y_2)$ .

In the 'E' step, we kept the log-likelihood contribution of all the observations belonging to  $I_0$  intact, as in this case the corresponding  $(\Delta_1, \Delta_2)$  are known completely. The observations are treated as missing observations, if they belong to  $I_1$  or  $I_2$ .

If  $(y_1, y_2) \in I_1$ , the 'pseudo observation' is formed, by fractioning  $(y_1, y_2)$  to two partially complete 'pseudo observations' of the form  $(y_1, y_2, u_1(\Theta))$  and  $(y_1, y_2, u_2(\Theta))$ . The fractional mass  $u_1(\Theta)$  and  $u_2(\Theta)$  assigned to the 'pseudo observation' are the conditional probabilities that  $(\Delta_1, \Delta_2)$  takes values  $(0,1)$  or  $(1,1)$ , respectively, given that  $(Y_1, Y_2) \in I_1$ . Similarly, if  $(Y_1, Y_2) \in I_2$ , 'pseudo observations' are formed as  $(y_1, y_2, v_1(\Theta))$  and  $(y_1, y_2, v_2(\Theta))$ , where  $v_1(\Theta)$  and  $v_2(\Theta)$  are the conditional probabilities that  $(\Delta_1, \Delta_2)$  takes values  $(1,0)$  and  $(1,1)$  respectively. Therefore,

$$v_1(\Theta) = \frac{\lambda_0}{\lambda_0 + \lambda_2}, \quad v_2(\Theta) = \frac{\lambda_2}{\lambda_0 + \lambda_2}, \quad u_1(\Theta) = \frac{\lambda_0}{\lambda_0 + \lambda_1}, \quad u_2(\Theta) = \frac{\lambda_1}{\lambda_0 + \lambda_1}.$$

Therefore, we will use the following notations in the  $k$ -th step of the EM algorithm for the estimates of the parameters.

- $\Theta^{(k)} = (\alpha^{(k)}, \lambda_0^{(k)}, \lambda_1^{(k)}, \lambda_2^{(k)})$  is defined for the estimates of the parameters in the  $k$ -th step.
- $E(N|y_{1i}, y_{2i}, \Theta) = a_i$ , and  $E(N|y_{1i}, y_{2i}, \Theta^{(k)}) = a_i^{(k)}$ .
- $u_1(\Theta^{(k)}) = u_1^{(k)}$ ,  $u_2(\Theta^{(k)}) = u_2^{(k)}$ ,  $v_1(\Theta^{(k)}) = u_1^{(k)}$  and  $v_2(\Theta^{(k)}) = u_2^{(k)}$ .

Now we are ready to provide the EM algorithm.

**E-Step:** At the  $k$ -step of the EM algorithm, the 'pseudo' log-likelihood function without the additive constant can be written as follows:

$$\begin{aligned}
\ell_{pseudo}(\Theta) = & (m_0 + 2m_1 + 2m_2) \ln \lambda_0 + (m_2 + m_1 u_2^{(k)}) \ln \lambda_1 + (m_2 v_2^{(k)} + m_1) \ln \lambda_2 \\
& + \lambda_0 \left\{ \sum_{i \in I_0} a_i^{(k)} \ln \bar{F}_B(y_i, \alpha) + \sum_{i \in I_2} a_i^{(k)} \ln \bar{F}_B(y_{2i}, \alpha) + \sum_{i \in I_1} a_i^{(k)} \ln \bar{F}_B(y_{1i}, \alpha) \right\} \\
& + \lambda_1 \left\{ \sum_{i \in I_0} a_i^{(k)} \ln \bar{F}_B(y_i, \alpha) + \sum_{i \in I_2} a_i^{(k)} \ln \bar{F}_B(y_{1i}, \alpha) + \sum_{i \in I_1} a_i^{(k)} \ln \bar{F}_B(y_{1i}, \alpha) \right\} \\
& + \lambda_2 \left\{ \sum_{i \in I_0} a_i^{(k)} \ln \bar{F}_B(y_i, \alpha) + \sum_{i \in I_2} a_i^{(k)} \ln \bar{F}_B(y_{2i}, \alpha) + \sum_{i \in I_1} a_i^{(k)} \ln \bar{F}_B(y_{2i}, \alpha) \right\} \\
& + \sum_{i \in I_0} \ln \frac{f_B(y_i, \alpha)}{\bar{F}_B(y_i, \alpha)} + \sum_{i \in I_2} \ln \frac{f_B(y_{1i}, \alpha)}{\bar{F}_B(y_{1i}, \alpha)} + \sum_{i \in I_1} \ln \frac{f_B(y_{2i}, \alpha)}{\bar{F}_B(y_{2i}, \alpha)} \\
& + \sum_{i \in I_1} \ln \frac{f_B(y_{2i}, \alpha)}{\bar{F}_B(y_{2i}, \alpha)} + \sum_{i \in I_0} \ln \frac{f_B(y_{1i}, \alpha)}{\bar{F}_B(y_{1i}, \alpha)} + m \ln \frac{\theta}{1 - \theta} + \ln(1 - \theta) \sum_{i=1}^m a_i^{(k)}. \quad (8)
\end{aligned}$$

**M-Step:** The 'M'-step involves maximizing  $\ell_{pseudo}(\Theta)$  with respect to the unknown parameters.

Now, the following steps can be used to compute the MLEs via the EM algorithm:

## ALGORITHM

- Step 1: Take some initial value of  $\Theta$ , say  $\Theta^{(0)} = (\theta^{(0)}, \alpha^{(0)}, \lambda_0^{(0)}, \lambda_1^{(0)}, \lambda_2^{(0)})$ .
- Step 2: Compute  $a_i^{(0)} = E(N|y_{1i}, y_{2i}; \Theta^{(0)})$ .
- Step 3: Compute  $u_1, u_2, v_1$ , and  $v_2$ .
- Step 4: Find  $\hat{\alpha}$  and say  $\hat{\alpha}^{(1)}$ .
- Step 5: Compute  $\hat{\lambda}_i^{(1)} = \hat{\lambda}_i(\hat{\alpha}^{(1)})$ ,  $i = 0, 1, 2$ .
- Step 6: Find  $\hat{\theta}$ .
- Step 7: Replace  $\Theta^{(0)}$  by  $\Theta^{(1)} = (\theta^{(1)}, \alpha^{(1)}, \lambda_0^{(1)}, \lambda_1^{(1)}, \lambda_2^{(1)})$ , go back to step 1 and continue the process until convergence take place.

## 4 Data analysis and comparison study

For illustrative purposes, we have analyzed one data set to see how the proposed model and the EM algorithm works in practice. The data set has been obtained from Davis [4]. The data set represents the total cholesterol contents of 23 subjects measured after 4-th week and 20-th week of the start of the medicine. Here, we consider four sub-models of BPHRG distributions: BWG, BLG, BCHG and BGG. Before progressing further, we have made the following transformation  $(X - 165)/100$  to all the observations mainly for model fitting and computational purposes.

For illustrative purposes, First we plot the scaled TTT plots, see Aarset [1] for details, of the marginals in Figure 1 for real data. Since both are concave functions, it can be assumed that the hazard function of the marginals is increasing functions. Moreover, the correlation between the two marginals is positive.

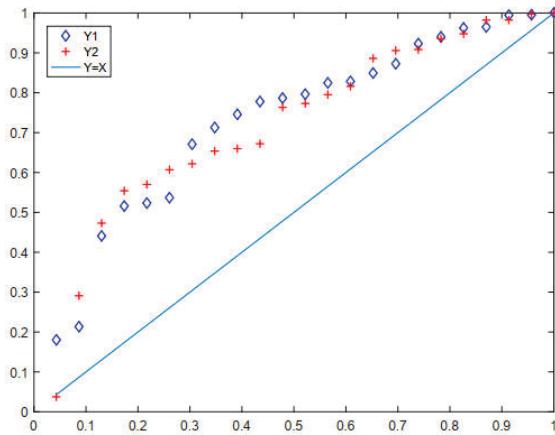


Figure 1: The scaled TTT plots of the marginals for data set.

Before going to analyze the data using four sub-models of BPHRG distributions, we fit the Weibull, Lomax, Chen and Gompertz distributions to  $Y_1$ ,  $Y_2$  and  $\min\{Y_1, Y_2\}$  separately. The MLE's of parameters, the corresponding Kolmogorov-Smirnov distances and the associated p-values are calculated and the results are presented in the Table 1. Based on the p-values the Weibull, Chen and Gampertz distributions cannot be rejected for the marginals and for the minimum also. We observe that the Lomax distribution is not suitable for this data set.

Now we will fit the BPHRG models. Therefore, three special cases of BPHRG distributions are considered: BWG, BCHG and BGG. Using the proposed EM algorithm, the MLE's and their corresponding log-likelihood values are calculated. For each fitted model, the Akaike Information Criterion (AIC) and the Bayesian information criterion (BIC) are calculated. The results are given in Table 2.

We also obtain the Kolmogorov-Smirnov (K-S) distances with the corresponding p-values between the fitted distribution and the empirical cdf for three random variables  $Y_1$ ,  $Y_2$  and  $\min(Y_1, Y_2)$ . The results are given in Table 3.

Finally, the likelihood ratio test (LRT) and the corresponding p-values are obtained for testing the BPHR models against the BPHRG models. On the other hand, our goal is

Table 1: The MLE's of parameters, the corresponding Kolmogorov-Smirnov (K-S) and the associated p-values.

	Model		$\alpha$	$\lambda$	K-S	P-value
Weibull	$Y_1$	2.3425	0.5315	0.1045	0.9313	
	$Y_2$	2.1126	0.7045	0.1045	0.9312	
	$\min\{Y_1, Y_2\}$	2.1131	0.8797	0.0993	0.9603	
	$Y_1$	289.0625	0.1760	0.5160	$1.8344 \times 10^{-6}$	
Lomax	$Y_2$	273.4375	0.1824	0.4717	$3.2314 \times 10^{-5}$	
	$\min\{Y_1, Y_2\}$	273.4275	0.1860	0.4738	$2.9085 \times 10^{-5}$	
	$Y_1$	1.2928	0.2784	0.1255	0.7994	
Cholesterol data set	$Y_2$	1.2860	0.3547	0.1508	0.6186	
	$\min\{Y_1, Y_2\}$	1.3524	0.4471	0.1347	0.7494	
	$Y_1$	1.6307	0.1169	0.1247	0.8058	
Chen	$Y_2$	1.6756	0.1370	0.1577	0.5631	
	$\min\{Y_1, Y_2\}$	1.8474	0.1397	0.1451	0.6650	
Gompertz						

Table 2: The MLE, the corresponding log-likelihood, AIC and BIC.

Data set	Model	$\hat{\alpha}$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\theta}$	$\log(\ell)$	AIC	BIC
Cholesterol data set	BWG	2.4921	0.0539	0.3428	0.4531	0.5183	-35.4855	80.9710	86.6485
	BCHG	1.4371	0.0234	0.1612	0.2237	0.4625	-32.3000	74.6000	80.2775
	BGG	2.0084	0.0060	0.0417	0.0580	0.6587	-32.2302	74.4604	80.1379

Table 3: The Kolmogorov-Smirnov (K-S) distances and the corresponding p-values for three random variables  $Y_1, Y_2$  and  $\min(Y_1, Y_2)$ .

Data set	Model	$Y_1$		$Y_2$		$\min\{Y_1, Y_2\}$	
		K-S	p-value	K-S	p-value	K-S	p-value
Cholesterol data set	BWG	0.1195	0.8432	0.0978	0.9652	0.1604	0.5421
		0.0886	0.9831	0.1074	0.9281	0.1954	0.3025
		0.1460	0.6336	0.1722	0.4528	0.1220	0.8432

to test the null hypothesis  $H_0 : BPHR$  against the alternative hypothesis  $H_1 : BPHRG$ . For this purpose, three special cases of BPHR distributions are considered: BWE, BCH and BG. The statistics and the corresponding p-values are given in Table 4. Hence, for any usual significance level, we reject proposed models in  $H_0$  (BPHR) in favor of the alternative models (BPHRG).

## 5 Conclusions

In this paper, we have proposed a new class of bivariate distributions. This class of bivariate distributions can be obtained by compounding geometric distribution and proportional hazard rate model. This class is contained the family of bivariate proportional hazard rate distributions (BPHR) and proportional hazard rate-geometric distributions (PHRG). Therefore, this new class is contained several bivariate lifetime models and is more flexible.

The estimation of unknown parameters by the method maximum likelihood is acquired. However, it is not directly easy to solve the associated log likelihood equations. There-

Table 4: The log-likelihood, AIC, BIC, LRT and the corresponding p-values for different models.

Data set	Test	Models					
		BWG	BWE	BCHG	BCH	BGG	BG
	AIC	80.9710	100.4322	74.6000	96.5322	74.4604	96.4608
	BIC	86.6485	105.4322	80.2775	101.0742	80.1379	101.1828
Cholesterol data set	$\log(\ell)$	-35.4855	-46.4451	-32.3000	-44.2661	-32.2302	-44.3204
	LRT		21.9192		23.9322		24.1802
	p-value		$2.8437 \times 10^{-6}$		$9.9789 \times 10^{-7}$		$8.7730 \times 10^{-7}$

fore, the corresponding maximum likelihood estimate is numerically calculated from the associated nonlinear equation by the EM algorithm. Therefore, we suggeste to use the EM algorithm to compute the MLE's of the unknown parameters, and it is observed that the proposed EM algorithm works quite well in practice. As shown, the proposed models work quite well for data analysis purposes. Finally, we compare BPHRG models to BPHR models.

## References

- [1] Aarset, M. (1987), How to identify a bathtub hazard rate, *IEEE Transactions on Reliability*, **36**, 106108.
- [2] Barreto-Souza, W. (2012), Bivariate gamma-geometric law and its induce levy process, *Journal of Multivariate Analysis*, **109**, 130145.
- [3] Cox, D. (1997), Regression models and life tables (with discussion), *Journal of the Royal Statistical Society. Series B (Methodological)*, **34**, 187220.
- [4] Davis, C. (2002), *Statistical methods for the analysis of repeated measurements*, Springer-Verlag, New York.
- [5] Ghitany, M. and Al-Hussaini, E. and Al-Jarallah, R. (2005), Marshall-Olkin extended weibull distribution and its application to censored data, *Journal of Applied Statistics*, **32**, 10251034.
- [6] Ghitany, M. and Al-Awadhi, F. and Alkhalfan, L. (2007), Marshall-Olkin extended lomax distribution and its application to censored data, *Communications in Statistics - Theory and Methods*, **36**, 18551866.
- [7] Marshall,A. and Olkin, I. (1997), A new method of adding a parameter to a family of distributions with application to the exponential and weibull families, *Biometrika*, **84**, 641652.
- [8] Pham, H. and Lai, C. (2007), On recent generalizations of the weibull distribution, *IEEE Transactions on Reliability*, **56** 454458.
- [9] Silva, R. and Bourguignon, M. and Dias, C. and Cordeiro, G. (2013), The compound class of extended weibull power series distributions, *Computational Statistics and Data Analysis*, **58**, 352367.



## On the behavior of the system failure rate

Tavangar, M.<sup>1</sup>

Department of Statistics, University of Isfahan

### Abstract

In this talk, we consider a coherent (or mixed) system with identical components where it is assumed that the lifetimes of the components are independent and have a common distribution function  $F$ . We are interested in the study of the behavior of the failure rate and reversed failure rate of the system or its component. Sufficient conditions are provided under which the system failure rate is increasing and the system reversed failure is decreasing. Some characterization results based on the generalized reversed failure rate function are also obtained.

**Keywords:** reliability, aging concepts, failure rate, reversed failure rate, order statistics.

## 1 Introduction

Evaluation of reliability of technical systems is of crucial importance in many application areas and has been the subject of a large literature. An important class of systems in this area is the class of coherent systems. The coherent system is a system consisting of  $n$  components with the properties of increasing structure function, and each of its components is relevant. See Barlow and Proschan (1975) for a basic study of coherent systems. In recent years, the concept of *system signature*, introduced by Samaniego (1985), has become a popular tool to assess the reliability of systems consisting of components with random failure times that are independent and identically distributed (i.i.d.), although this can be relaxed to assuming exchangeability. For a coherent system with lifetime  $T$

<sup>1</sup>m.tavangar@stat.ui.ac.ir

whose component lifetimes  $X_1, X_2, \dots, X_n$  are i.i.d. random variables with continuous distribution function  $F$ , the signature vector of the system is defined as a probability vector  $\mathbf{s} = (s_1, s_2, \dots, s_n)$  with

$$s_i = P\{T = X_{i:n}\}, \quad i = 1, 2, \dots, n,$$

where  $X_{i:n}$  is the  $i$ th order statistic among  $X_1, X_2, \dots, X_n$  and  $\sum_{i=1}^n s_i = 1$ . The probability  $s_i$  is in fact the probability that the  $i$ th component failure (with lifetime  $X_{i:n}$ ) causes the failure of the system. A survey of the various applications of system signature is given in Samaniego (2007). The concept of signature provides an interesting representation of the system reliability in terms of reliability of  $k$ -out-of- $n$  systems (a system with  $n$  components has  $k$ -out-of- $n$  structure if it operates as long as at least  $k$  of its components operate,  $k \leq n$ ). As the signature vector  $\mathbf{s}$  does not depend on the common distribution  $F$  of the components, the reliability function of the system lifetime  $T$  can be represented as

$$P\{T > t\} = \sum_{i=1}^n s_i P\{X_{i:n} > t\}. \quad (1)$$

It is known that the system reliability can also be expressed as a generalized mixture (i.e., a mixture with possibly negative weights) of the reliability of series and parallel systems. The weights (referred to as minimal and maximal signatures) can be obtained from the system signature and vice versa.

Boland and Samaniego (2004) propose the utilization of mixed systems which are stochastic mixtures of coherent systems of a given size. Using expression (1), it is self-evident that if the lifetimes are exchangeable, then any mixed system can be written as a mixture of  $X_{1:n}, X_{2:n}, \dots, X_{n:n}$ . The vector of the coefficients in this representation is called the signature vector of order  $n$  of the mixed system. Conversely, given a vector  $(s_1, s_2, \dots, s_n)$  of non-negative numbers such that  $s_1 + s_2 + \dots + s_n = 1$ , we can define the mixed system  $T$  which takes the value  $X_{i:n}$  with probability  $s_i$  for  $i = 1, 2, \dots, n$ .

In reliability theory and survival analysis to study the lifetime properties of a component (or any other living organism) there are several measures such as the hazard rate function and the reversed hazard rate function. Let  $T$  be a lifetime random variable with hazard rate  $h(t)$  and reversed hazard rate  $r(t)$ , respectively. Then, for  $t > 0$ , these measures are defined as follows:

$$h(t) = \frac{f(t)}{\bar{F}(t)}, \quad \bar{F}(t) > 0,$$

$$r(t) = \frac{f(t)}{F(t)}, \quad F(t) > 0,$$

where  $f(t)$ ,  $F(t)$  and  $\bar{F}(t)$  are the density function, the distribution function and the survival function of  $T$ , respectively. We refer the reader Shaked and Shanthikumar (2007) for some results regarding these measures.

In this talk, we present some results regarding the behavior of  $h(t)$  and  $r(t)$  for coherent and mixed systems.

## 2 System failure rate and reversed failure rate

In this section, some results on the behaviour of the failure rate and the reversed failure rate functions of coherent systems are provided. It is known that if the component lifetimes of a  $k$ -out-of- $n$  system are i.i.d. according to an IFR (increasing failure rate) distribution, then the distribution of the system lifetime is IFR (see Barlow and Proschan, 1975). Samaniego (2007) prove the result for a mixed system under a condition on the structure of the system. In the first result of the paper, we give a more simpler sufficient condition for a coherent (or mixed) system with IFR (DFR (decreasing failure rate)) component lifetimes to be IFR (DFR). In order to prove the result, the following lemma due to Karlin (1968) is needed. Let  $A$ ,  $B$  and  $C$  be subsets of the real line; and let  $L(x, z)$  be SR<sub>2</sub> for  $x \in A$ ,  $z \in B$ , and  $M(z, y)$  be SR<sub>2</sub> for  $z \in B$ ,  $y \in C$ . Then, for any  $\sigma$ -finite measure  $\mu$ ,  $K(x, y) = \int_B L(x, z)M(z, y)d\mu(z)$  is also SR<sub>2</sub> for  $x \in A$  and  $y \in C$ , and  $\varepsilon_i(K) = \varepsilon_i(L)\varepsilon_i(M)$  for  $i = 1, 2$ , where  $\varepsilon_i(K) = \varepsilon_i$  denotes the constant sign of the  $i$ th-order determinants.

**Theorem 2.1.** *Let  $\mathbf{s} = (s_1, s_2, \dots, s_n)$  be the signature of a coherent or mixed system based on  $n$  i.i.d. components with common lifetime distribution  $F$ . If  $F$  is IFR (DFR) and  $(n-i)s_{i+1}/\sum_{j=i+1}^n s_j$  (whenever defined) is non-decreasing (non-increasing) in  $i$ , then the system's lifetime is also IFR (DFR).*

**Proof.** For an  $n$ -component mixed system based on components with i.i.d. lifetimes and with common distribution  $F$ , the system failure rate can be written as (Samaniego, 2007)

$$h_T(t) = \frac{\sum_{i=0}^{n-1} (n-i)s_{i+1} \binom{n}{i} u_t^i}{\sum_{i=0}^{n-1} \left( \sum_{j=i+1}^n s_j \right) \binom{n}{i} u_t^i} h(t),$$

where  $u_t = F(t)/\bar{F}(t)$  represents the odds of failure versus survival. To prove the result, we need to show that the function  $\xi$ , given by  $\xi(x, k) = \sum_{i=0}^{n-1} c_{i,k} \binom{n}{i} x^i$  is TP<sub>2</sub> (RR<sub>2</sub>) in  $(x, k) \in [0, \infty) \times \{0, 1\}$ , where

$$c_{i,k} = \begin{cases} \sum_{j=i+1}^n s_j & k = 0 \\ (n-i)s_{i+1}, & k = 1 \end{cases}$$

It is easy to see that  $x^i$  is TP<sub>2</sub> in  $(x, i) \in [0, \infty) \times \{0, 1, \dots, n-1\}$ . Under the assumption of the theorem,  $c_{i,k}$  is TP<sub>2</sub> (RR<sub>2</sub>) in  $(i, k) \in \{0, 1, \dots, n-1\} \times \{0, 1\}$ . The required result then follows from Lemma 2. ■

**Remark 2.1.** It follows from the proof of Theorem 2.1 that if  $F$  is *initially* IFR and  $(n-i)s_{i+1}/\sum_{j=i+1}^n s_j$  is non-decreasing in  $i$ , then the system's lifetime is also *initially* IFR. Also it can be easily deduced that in the case where  $F$  is DFR and  $\mathbf{s}$  is a DFR discrete probability vector (that is,  $s_{i+1}/\sum_{j=i+1}^n s_j$  is non-increasing in  $i$ ), the system's lifetime is also DFR.

An analog of Theorem 2.1 about the preservation of decreasing reversed failure rate class under the formation of coherent system is as follows. The proof is similar to the proof of Theorem 2.1.

Table 1: Signatures of order 4 for all coherent systems with 1–4 components

$N$	$T = \phi(X_1, X_2, X_3, X_4)$	$\mathbf{s}$	$(n - i)s_{i+1} / \sum_{j=i+1}^n s_j$
1	$X_{1:1} = X_1$	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	constant
2	$X_{1:2} = \min(X_1, X_2)$	$(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}, 0)$	constant
3	$X_{2:2} = \max(X_1, X_2)$	$(0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2})$	increasing
4	$X_{1:3} = \min(X_1, X_2, X_3)$	$(\frac{3}{4}, \frac{1}{4}, 0, 0)$	constant
5	$\min(X_1, \max(X_2, X_3))$	$(\frac{1}{4}, \frac{5}{12}, \frac{1}{3}, 0)$	increasing
6	$X_{2:3}$	$(0, \frac{1}{2}, \frac{1}{2}, 0)$	increasing
7	$\max(X_1, \min(X_2, X_3))$	$(0, \frac{1}{3}, \frac{5}{12}, \frac{1}{4})$	$\times$
8	$X_{3:3} = \max(X_1, X_2, X_3)$	$(0, 0, \frac{1}{4}, \frac{3}{4})$	increasing
9	$X_{1:4} = \min(X_1, X_2, X_3, X_4)$	$(1, 0, 0, 0)$	constant
10	$\max(\min(X_1, X_2, X_3), \min(X_2, X_3, X_4))$	$(\frac{1}{2}, \frac{1}{2}, 0, 0)$	increasing
11	$\min(X_{2:3}, X_4)$	$(\frac{1}{4}, \frac{3}{4}, 0, 0)$	increasing
12	$\min(X_1, \max(X_2, X_3), \max(X_3, X_4))$	$(\frac{1}{4}, \frac{1}{12}, \frac{1}{6}, 0)$	$\times$
13	$\min(X_1, \max(X_2, X_3, X_4))$	$(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}, 0)$	increasing
14	$X_{2:4}$	$(0, 1, 0, 0)$	increasing
15	$\max(\min(X_1, X_2), \min(X_1, X_3, X_4), \min(X_2, X_3, X_4))$	$(0, \frac{5}{6}, \frac{1}{6}, 0)$	$\times$
16	$\max(\min(X_1, X_2), \min(X_3, X_4))$	$(0, \frac{2}{3}, \frac{1}{3}, 0)$	increasing
17	$\max(\min(X_1, X_2), \min(X_1, X_3), \min(X_2, X_3, X_4))$	$(0, \frac{2}{3}, \frac{1}{3}, 0)$	increasing
18	$\max(\min(X_1, X_2), \min(X_2, X_3), \min(X_3, X_4))$	$(0, \frac{1}{2}, \frac{1}{2}, 0)$	increasing
19	$\min(\max(X_1, X_2), \max(X_2, X_3), \max(X_3, X_4))$	$(0, \frac{1}{2}, \frac{1}{2}, 0)$	increasing
20	$\min(\max(X_1, X_2), \max(X_1, X_3), \max(X_2, X_3, X_4))$	$(0, \frac{1}{3}, \frac{2}{3}, 0)$	increasing
21	$\min(\max(X_1, X_2), \max(X_3, X_4))$	$(0, \frac{1}{3}, \frac{2}{3}, 0)$	increasing
22	$\min(\max(X_1, X_2), \max(X_1, X_3, X_4), \max(X_2, X_3, X_4))$	$(0, \frac{1}{6}, \frac{5}{6}, 0)$	increasing
23	$X_{3:4}$	$(0, 0, 1, 0)$	increasing
24	$\max(X_1, \min(X_2, X_3, X_4))$	$(0, \frac{1}{2}, \frac{1}{4}, \frac{1}{4})$	$\times$
25	$\max(X_1, \min(X_2, X_3), \min(X_3, X_4))$	$(0, \frac{1}{6}, \frac{1}{12}, \frac{1}{4})$	$\times$
26	$\max(X_{2:3}, X_4)$	$(0, 0, \frac{3}{4}, \frac{1}{4})$	$\times$
27	$\min(\max(X_1, X_2, X_3), \max(X_2, X_3, X_4))$	$(0, 0, \frac{1}{2}, \frac{1}{2})$	increasing
28	$X_{4:4} = \max(X_1, X_2, X_3, X_4)$	$(0, 0, 0, 1)$	increasing

**Theorem 2.2.** Let  $\mathbf{s} = (s_1, s_2, \dots, s_n)$  be the signature of a coherent or mixed system based on  $n$  i.i.d. components with common lifetime distribution  $F$ . If  $F$  is DRFR (decreasing reversed failure rate) and  $is_i / \sum_{j=1}^i s_j$  (whenever defined) is non-increasing in  $i$ , then the system's lifetime is also DRFR.

The signatures of order 4 for all coherent systems with 1–4 components can be found in Navarro and Rubio (2011). In Table 1, we have determined that whether each system satisfies the assumption of Theorem 2.1. It is seen that from 28 coherent systems listed in Table 1, only 6 systems do not fulfill the stated assumption (in the case of IFR).

It is known from Block et al. (1998) that there does not exist any nonnegative random variable having increasing reversed failure rate function on its interval of support. In the following theorem, we have proved a stronger result implying that the reversed failure rate function of any continuous and nonnegative random variable must be first decreasing

at the beginning of the interval of support. In other words, such a random variable has initially DRFR distribution.

**Theorem 2.3.** *Let  $X$  be an absolutely continuous random variable with distribution  $F(x)$  and interval of support  $(a, b)$ , where  $0 \leq a < b \leq \infty$ . Furthermore, let  $\delta \in (0, b - a]$  be an arbitrary value. Then the reversed failure rate function of  $X$  is not increasing on  $(a, a + \delta)$ .*

**Proof.** The case  $\delta = b - a$  have been proved in Block et al. (1998). Thus, assume that  $\delta < b - a$  which implies that  $\delta < \infty$ .

Let  $r(t)$ , the reversed failure rate function of  $X$ , be increasing on  $(a, a + \delta)$ . It follows that

$$\int_t^{a+\delta} r(x)dx \leq \int_t^{a+\delta} r(a + \delta)dx = (a + \delta - t)r(a + \delta),$$

for all  $t \in (a, a + \delta)$ . This, in turn, implies that

$$F(a + \delta) \leq F(t)e^{a+\delta-t}r(a + \delta).$$

By considering the limits as  $t \rightarrow a^+$  and using the continuity of  $F$ , we conclude that  $F(a + \delta) \leq 0$ . Therefore,  $F(a + \delta) = 0$  which is a contradiction to the fact that  $a = \inf\{x : F(x) > 0\}$ . ■

The bathtub-shaped failure rate plays an important role in reliability practices such as burn-in or environmental-stress-screening to manufactured products. In this case, the curve has the property that in so-called “early failure” period, the failure rate decreases over time. Klutke et al. (2003) noticed some limitations of the bathtub-shaped failure rate in mixture models, and showed that a sufficient condition for the mixture of distributions with concave reliability functions in a neighborhood of 0 to have an IFR at 0 is that at least one of the distributions has a strictly concave reliability function in a neighborhood of 0. This means that the mixture of such distributions is initially IFR and thus cannot follow the classical bathtub shape. The following theorem reveals an application of this result to coherent systems.

**Theorem 2.4.** *Consider a coherent system consisting of  $n$  independent and identical components with a sufficiently smooth (continuous second time-derivatives) cdf  $F$  and signature  $(s_1, s_2, \dots, s_n)$ . Let  $k = \min\{i : s_i > 0\}$ . If  $\bar{F}_{k:n}(t)$  is strictly concave in a neighborhood of any point  $t_0$ , then the failure rate function of the system is strictly increasing in this neighborhood.*

**Proof.**

Under the assumption of the theorem, the mixture representation of the system reliability function reduces to

$$\bar{F}_T(t) = \sum_{i=k}^n s_i \bar{F}_{i:n}(t).$$

Observe that if  $\bar{F}_{k:n}(t)$  is strictly concave in a neighborhood of  $t_0$ , then  $f'_{k:n}(t)$  is strictly increasing in that interval. It follows that

$$f'_{k+1:n}(t) = \frac{n-k}{k} \{ [1 + \phi(t)] h(t) f_{k:n}(t) + \phi(t) f'_{k:n}(t) \},$$

where  $\phi(t) = F(t)/\bar{F}(t)$ , is positive and hence  $\bar{F}_{i:n}(t)$ ,  $i = k, k+1, \dots, n$ , is strictly concave in the neighborhood of  $t_0$ . The proof is completed by using the result obtained in Klutke et al. (2003). ■

Let  $f(t)$  be the probability density function of the component lifetimes of a coherent system with i.i.d. components. If  $f(0) = 0$  and  $f$  is strictly increasing in a neighborhood of 0, then the system is initially IFR.

**Remark 2.2.** It follows from the proof of Theorem 2.4 that a sufficient condition for  $\bar{F}_{k:n}(t)$  to be strictly concave on an interval is that  $\bar{F}^n(t)$  be strictly concave on that interval. Many important lifetime distributions such as the Weibull distribution and the gamma distribution with shape parameters greater than 1 satisfy the assumptions of Corollary 2. We conclude that if the lifetimes of the components of a coherent system have such distributions, then the system lifetime is initially IFR.

**Example 2.1.** Let the component lifetimes of the system has the beta distribution with strictly concave reliability

$$\bar{F}(t) = 1 - t^2, \quad 0 \leq t \leq 1.$$

It is easy to verify that  $\bar{F}^n(t)$  is strictly concave in the interval  $[0, \frac{1}{\sqrt{2n-1}}]$  and therefore, using the last remark and Theorem 2.4, any  $n$ -component system with  $\bar{F}(t)$  as component reliability is initially IFR.

### 3 Characterization results

Let  $X$  be a lifetime random variable with an absolutely continuous distribution function  $F(x)$ , reliability function  $\bar{F}(x) = 1 - F(x)$  and probability density function  $f(x)$ . Let  $a = \inf\{x : F(x) > 0\}$  and  $b = \sup\{x : F(x) < 1\}$ . The generalized reversed failure rate of  $X$  is defined as

$$r(x, y) = \lim_{\delta \rightarrow 0^+} \frac{1}{\delta} P\{y - \delta < X \leq y \mid x < X \leq y\} = \frac{f(y)}{\bar{F}(y) - \bar{F}(x)}, \quad (2)$$

whose domain of definition is the set  $D = \{(x, y) \in [a, b] : F(x) < F(y)\}$ . It is easily seen that  $r(x, x+y)$  is in fact the reversed failure rate function of  $(X - x \mid X > x)$  at time  $y$  after it is put into operation, which is the lifetime of a used component or system of age  $x$ .

Most of the characterization results on exponential distribution are based on the “lack of memory” property. The next result reveals a characterization of the exponential distribution involving the the generalized reversed failure rate function. In the proof, a result stronger than the lack of memory property of exponential distribution is utilized.

**Theorem 3.1.** *Let  $F(x)$  be an absolutely continuous cdf concentrated on  $\mathbb{R}_+ = [0, \infty)$  with corresponding reversed failure rate and generalized reversed failure rate functions  $r(x)$  and  $r(x, y)$ . Then*

$$r(x_n, x_n + y) = r(y), \quad n = 1, 2; y > 0, \quad (3)$$

for two non-negative and incommensurable values  $x_1$  and  $x_2$  (that is,  $x_1/x_2$  is irrational), if and only if  $F$  is an exponential distribution.

**Proof.** Let  $X$  be a random variable with cdf  $F$ . As we have already mentioned at the beginning of this section,  $r(x_n, x_n + y)$  is the reversed failure rate function of  $(X - x_n \mid X > x_n)$ . Also it is known that the reversed failure rate function uniquely identifies the distribution. Therefore, it follows from (3) that  $(X - x_n \mid X > x_n) \stackrel{d}{=} X$ ,  $n = 1, 2$ , which is equivalent to

$$\bar{F}(x_n + y) = \bar{F}(x_n)\bar{F}(y), \quad n = 1, 2; y > 0.$$

The result then follows from Theorem 1 in Marsaglia and Tubilla (1975). ■

Suppose  $X$  is a non-negative real random variable. We say that  $X$  has a generalized Pareto distribution with parameter vector  $(c, .)$  if  $1 + cX > 0$  almost surely and  $\phi_c^{-1}(X)$  is exponential, where  $c \in \mathbb{R}$  and

$$\phi_c(x) = \begin{cases} \frac{e^{cx}-1}{c}, & \text{if } c \neq 0, \\ x, & \text{if } c = 0. \end{cases}$$

This family of distributions includes the exponential distribution (when  $c = 0$ ), the Pareto distribution (when  $c > 0$ ), and a rescaled beta distribution (when  $c < 0$ ). The GPD has many interesting characteristic properties. For example it has a linear mean residual life and its hazard rate is reciprocal of a linear function. These facts make the GPD to be a flexible model for different kinds of data. Many results have been appeared in the literature characterizing the GPD based on various properties. Oakes and Dasu (1990) consider a characterization of the GPD of the type of lack of memory property of the exponential distribution. Motivated by Oakes and Dasu's result, Asadi et al. (2001) extended and unified many characterization results on the exponential and geometric distributions, specially those that are based on order statistics and record values to arrive at the GPD. In the following theorem, a characterization of the GPD based upon the generalized reversed failure rate is provided.

**Theorem 3.2.** *Let  $F(x)$  be a cdf satisfying assumptions of Theorem 2. Furthermore, assume that  $F$  has finite mean. Let  $\theta : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , which is continuous from the right at 0 and  $\theta(0) = 1$ . Then there exist a point  $x_0 \in \mathbb{R}_+$  with  $F(x_0) < 1$  and a sequence  $\{x_n : n = 1, 2, \dots\}$  of points lying in  $(x_0, \infty)$  such that it converges to  $x_0$  and*

$$\theta(x_n)r(x_n, x_n + y\theta(x_n)) = r(y), \quad n = 0, 1, 2, \dots; y > 0, \quad (4)$$

*if and only if  $F$  is a GPD.*

**Proof.** The proof follows by integrating both sides of (4) with respect to  $y$  from  $t > 0$  to  $\infty$  and then using Theorem 3.4 of Asadi et al. (2001). The details are omitted. ■

## References

- [1] Asadi, M., Rao, C.R. and Shanbhag, D.N. (2001), Some unified characterization results on the generalized Pareto distribution. *Journal of Statistical Planning and Inference*, **93**, 29–50.
- [2] Barlow, R.E. and Proschan, F. (1975), *Statistical Theory of Reliability and Life Testing*. Holt, Rinehart: Winston, New York.
- [3] Block, H.W., Savits, T.H. and Singh, H. (1998), The reversed hazard rate function. , **12**(1), 69–90.
- [4] Boland, P.J. and Samaniego, F. (2004). The signature of a coherent system and its applications in reliability. In: Soyer R., Mazzuchi T., Singpurwalla, N. (eds.) *Mathematical reliability: an expository perspective*, international series in operational research and management science. Kluwer, Boston, pp. 1–29.
- [5] Karlin, S. (1968), *Total positivity*. Vol. I, Stanford University Press, Standford.
- [6] Klutke, G., Kiessler, P.C., and Wortman, M.A. (2003), A critical look at the bathtub curve. *IEEE Transactions on Reliability*, **52**(1), 125–129.
- [7] Marsaglia, G. and Tubilla, A. (1975), A note on the “lack of memory” property of the exponential distribution. *The Annals of Probability*, **3** (2), 353–354.
- [8] Navarro, J. and Rubio, R. (2011), A note on necessary and sufficient conditions for ordering properties of coherent systems with exchangeable components. *Naval Research Logistics*, **58** (5), 478–489.
- [9] Oakes, D. and Dasu, T. (1990), A note on residual life. *Biometrika*, **77**, 409–410.
- [10] Samaniego, F.J. (1985), On the closure of the IFR class under formation of coherent systems. *IEEE Transactions on Reliability*, **R-34**, 69–72.
- [11] Samaniego, F.J. (2007), *System Signatures and Their Applications in Engineering Reliability*. Springer, New York.
- [12] Shaked, M. and Shanthikumar, J.G. (2007), *Stochastic Orders*. Springer, New York.



## On the estimation of stress strength reliability parameter of new Burr distribution

Yari, G. H.<sup>1</sup>      Tondpour, Z.<sup>2</sup>

<sup>1,2</sup> Department of Statistics, Iran University of Science and Technology, Tehran

### Abstract

In this article, the estimation of the stress-strength reliability parameter  $R = P(X > Y)$ , where  $X$  and  $Y$  are independent random variables with New Burr distribution, has been considered. The maximum likelihood and bayes estimators of  $R$  are obtained. The bayes estimation of reliability parameter has been discussed under the assumption of exponential prior and under two different loss functions ( entropy and linex loss functions).

**Keywords:** Bayes estimator, Maximum Likelihood estimator, New Burr distribution, Stress-Strength reliability parameter.

## 1 Introduction

In statistical literature Burr family of distributions have attracted many statisticians for many years due to their applicability in different areas such as lifetime analysis. The system of Burr distributions was first introduced in the literature by Burr (1942). The Burr system of distributions includes twelve types of cumulative distribution functions which yield a variety of density shapes. Family of Burr type distributions is a widely applicable distributions family for modelling and analysis lifetimes (see for example, Okasha and Matter (2015) and Surles and Padgett (2001) ).

Analogous to the Pearson system of distributions, the Burr distributions are solutions to

<sup>1</sup>yari@iust.ac.ir

<sup>2</sup>ztondpoor@alumni.iust.ac.ir

a differential equation, which has the form

$$\frac{dy}{dx} = y(1-y)g(x,y), \quad (1)$$

where  $y$  equal to  $F(x)$  and  $g(x,y)$  must be positive for  $y$  in the unit interval and  $x$  in the support of  $F(x)$ .

Yari and Tondpour [4] introduced a new distribution of Burr type distributions which is more flexible by replacing  $g(x,y)$  with  $g(x) = \frac{3px^2e^{-x^3}(1+e^{-x^3})^{p-1}}{(1+e^{-x^3})^p - 1}$ , ( $p > 0$ ). In this case, using the solution of the differential equation (1.1) cdf and pdf of new distribution are, respectively, given by:

$$F(x;p) = (1 + e^{-x^3})^{-p}, \quad -\infty < x < \infty, \quad (p > 0),$$

and

$$f(x;p) = 3px^2e^{-x^3}(1 + e^{-x^3})^{-p-1}, \quad -\infty < x < \infty.$$

This new distribution was named the New Burr distribution (NBurr distribution). If the location parameter  $\mu$  and the scale parameter  $\sigma$  are introduced, we have

$$F(x;\mu,\sigma,p) = (1 + e^{-(\frac{x-\mu}{\sigma})^3})^{-p}, \quad -\infty < x < \infty, \quad (p, \sigma > 0, \mu \in \mathbb{R}) \quad (2)$$

and

$$f(x;\mu,\sigma,p) = \frac{3p}{\sigma} \left(\frac{x-\mu}{\sigma}\right)^2 e^{-(\frac{x-\mu}{\sigma})^3} (1 + e^{-(\frac{x-\mu}{\sigma})^3})^{-p-1}.$$

Hence equation (1.2) is three parameter New Burr distribution.

The shapes of density and hazard functions of the New Burr distribution for different values of shape parameter  $p$  are illustrated in Figure 1.

As shown in Figure 1, New Burr distribution has unimodal and bimodal pdfs and modified unimodal (unimodal followed by increasing) hazard function that's it, the attractiveness of New Burr distribution. None of the 12 types of Burr distributions has this feature. Data that exhibit bimodal behavior arises in many different disciplines. In medicine, urine mercury excretion has two peaks, see for example, Ely et al., 1999. In material characterization, a study conducted by Dierickx et al. (2000), grain size distribution data reveals a bimodal structure. In meteorology, Zhang et al. (2003) indicated that , water vapor in tropics, commonly have bimodal distributions. To see more applications of bimodal

distributions, see Garboś and Świecicka, 2015, Knapp, 2007 and Sewell and Young, 1997.

Yari and Tondpour (2017) showed that this particular skewed distribution can be used quite effectively in analyzing lifetime data. Also they obtained various properties of the New Burr distribution such that moments, quantile functions, hazard function, Shannon's entropy and the exact form of the probability density function and moments of  $i^{th}$ -order statistics in a sample of size  $n$  from New Burr distribution. Estimation of parameters and change-point of hazard function by the maximum likelihood method were discussed. Also the flexibility of the new model was illustrated with an application to a real data set and a goodness of fit test statistic based on the Rényi Kullback-Leibler information was used.

Here, let  $X$  and  $Y$  are the independent random variables. Then the stress-strength reliability  $R$  is calculated as

$$\begin{aligned} R = P(X > Y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^x f(x, y) dy dx = \int_{-\infty}^{+\infty} \left[ \int_{-\infty}^x f_Y(y) dy \right] f_X(x) dx \\ &= \int_{-\infty}^{+\infty} F_Y(x) f_X(x) dx. \end{aligned}$$

In this article, we let  $X$  and  $Y$  have New burr distribution with parameters  $p_1$  and  $p_2$ , respectively, and  $\mu_1 = \mu_2 = 0$ ,  $\sigma_1 = \sigma_2 = 1$ . then, the parameter  $R$  is calculated as:

$$R = P(X > Y) = \int_{-\infty}^{+\infty} 3p_1 x^2 e^{-x^3} (1 + e^{-x^3})^{-p_1 - p_2 - 1} dx,$$

by using  $u = 1 + e^{-x^3}$ , we have

$$R = \frac{p_1}{p_1 + p_2}. \quad (3)$$

In this article, we estimate  $R$  by maximum likelihood and Bayesian methods.

The remainder of the paper is organized as follows:

In section 1, New Burr distribution is defined. in section 2 and 3, estimation of reliability parameter by the maximum likelihood and Baysian methods is discussed, respectively and in section 4 a simulation study is provided to compare various estimators.

## 2 Maximum likelihood estimation of reliability parameter

Let that  $x_1, \dots, x_n$  and  $y_1, \dots, y_m$  are two sets of independent observations from NBurr( $p_1$ ) and NBurr( $p_2$ ), respectively. Then log-likelihood function is given by:

$$l(p_1, p_2) = n \log(3p_1) + 2 \sum_{i=1}^n \log(x_i) - \sum_{i=1}^n x_i^3 - (p_1 + 1) \sum_{i=1}^n \log(1 + e^{-x_i^3}) +$$

$$m \log(3p_2) + 2 \sum_{i=1}^m \log(y_i) - \sum_{i=1}^m y_i^3 - (p_2 + 1) \sum_{i=1}^m \log(1 + e^{-y_i^3}).$$

By differentiating of log-likelihood function with respect to  $p_1$  and  $p_2$  get the following equations

$$\frac{\partial l}{p_1} = \frac{n}{p_1} - \sum_{i=1}^n \log(1 + e^{-x_i^3}),$$

$$\frac{\partial l}{p_2} = \frac{m}{p_2} - \sum_{i=1}^m \log(1 + e^{-y_i^3}).$$

Then the maximum likelihood estimations of  $p_1$  and  $p_2$  are, respectively, as

$$\hat{p}_{1mle} = \frac{n}{\sum_{i=1}^n \log(1 + e^{-x_i^3})},$$

$$\hat{p}_{2mle} = \frac{m}{\sum_{i=1}^m \log(1 + e^{-y_i^3})}.$$

Hence, using the invariance properties of maximum likelihood estimations, the maximum likelihood estimation of parameter  $R$  is given by

$$\hat{R}_{mle} = \frac{\hat{p}_{1mle}}{\hat{p}_{1mle} + \hat{p}_{2mle}}. \quad (4)$$

### 3 Bayes estimation of reliability parameter

In this section, we use the bayesian estimation procedure for the estimation of  $R$ . Let that  $x_1, \dots, x_n$  and  $y_1, \dots, y_m$  are two sets of independent observations from NBurr( $p_1$ ) and NBurr( $p_2$ ), respectively. Since  $p_1$  and  $p_2$  are independent, then the joint prior distribution of  $p_1$  and  $p_2$  is given by

$$\pi(p_1, p_2) = \pi(p_1) \times \pi(p_2)$$

Now, if priors of  $p_1$  and  $p_2$  are exponential distributions with parameters  $\lambda_1$  and  $\lambda_2$ , respectively, therefore

$$\pi(p_1, p_2) = \lambda_1 \lambda_2 e^{-\lambda_1 p_1 - \lambda_2 p_2}; \quad p_1, p_2, \lambda_1, \lambda_2 > 0$$

Hence, the joint posterior distribution of  $p_1$  and  $p_2$  is given by

$$\begin{aligned} \pi(p_1, p_2 | x_1, \dots, x_n, y_1, \dots, y_m) &\propto L(x_1, \dots, x_n, y_1, \dots, y_m) \times \pi(p_1, p_2) \\ &\propto \prod_{i=1}^n 3p_1 x_i^2 e^{-x_i^3} (1 + e^{-x_i^3})^{-p_1-1} \times \end{aligned}$$

$$\begin{aligned}
& \prod_{j=1}^m 3p_2 y_j^2 e^{-y_j^3} (1 + e^{-y_j^3})^{-p_2-1} \lambda_1 \lambda_2 e^{-\lambda_1 p_1 - \lambda_2 p_2} \\
& \propto p_1^n e^{-(T_1 + \lambda_1)p_1} p_2^m e^{-(T_2 + \lambda_2)p_2}, \\
& \propto \text{gamma}(n+1, T_1 + \lambda_1) \times \text{gamma}(m+1, T_2 + \lambda_2),
\end{aligned}$$

where  $T_1 = \sum_{i=1}^n \ln(1 + e^{-x_i^3})$  and  $T_2 = \sum_{j=1}^m \ln(1 + e^{-y_j^3})$ . Then, the posterior distributions of  $p_1$  and  $p_2$  are as

$$p_1 | x_1, \dots, x_n \propto \text{gamma}(n+1, T_1 + \lambda_1), \quad p_2 | y_1, \dots, y_m \propto \text{gamma}(m+1, T_2 + \lambda_2) \quad (5)$$

Let  $X$  and  $Y$  are independent random variables where  $X \sim \text{gamma}(\alpha_1, \beta_1)$  and  $Y \sim \text{gamma}(\alpha_2, \beta_2)$ . Then the pdf of  $W = \frac{X}{X+Y}$  is given by

$$f_W(w) = \frac{\beta_1^{\alpha_1} \beta_2^{\alpha_2}}{B(\alpha_1, \alpha_2)} \frac{w^{\alpha_1-1} (1-w)^{\alpha_2-1}}{(\beta_1 w + \beta_2 (1-w))^{\alpha_1+\alpha_2}}, \quad 0 < w < 1 \quad (6)$$

where  $B(.,.)$  is beta function.

Therefore, according to (1.3), (3.1) and (3.2), the posterior density of  $R$  is given by

$$\begin{aligned}
& \pi(r | x_1, \dots, x_n, y_1, \dots, y_m) = \\
& \frac{(T_1 + \lambda_1)^{n+1} (T_2 + \lambda_2)^{m+1}}{B(n+1, m+1)} \frac{r^n (1-r)^m}{((T_1 + \lambda_1)r + (T_2 + \lambda_2)(1-r))^{n+m+2}}, \quad 0 < r < 1
\end{aligned} \quad (7)$$

- Types of loss functions used in this article

If  $\hat{r}$  represent of estimator for parameter  $r$ .

1) Entropy loss function (ELF)

$$L(r, \hat{r}) = b(\frac{\hat{r}}{r} - \ln(\frac{\hat{r}}{r}) - 1) \quad ; \quad \hat{r}_{ELF} = (E(\frac{1}{r}))^{-1}$$

2) Linex (linear-exponential) loss function (LLF)

$$L(r, \hat{r}) = e^{c(\hat{r}-r)} - c(\hat{r}-r) - 1 \quad ; \quad \hat{r}_{LLF} = -\frac{1}{c} \ln E(e^{-cr})$$

Hence, under Entropy loss function, the Bayes estimation of  $R$  is given by

$$\begin{aligned}
& (\hat{r}_{ELF})^{-1} = \\
& \int_0^1 \frac{(T_1 + \lambda_1)^{n+1} (T_2 + \lambda_2)^{m+1}}{B(n+1, m+1)} \frac{r^{n-1} (1-r)^m}{((T_1 + \lambda_1)r + (T_2 + \lambda_2)(1-r))^{n+m+2}} dr,
\end{aligned} \quad (8)$$

and under Linex loss function, the Bayes estimation of  $R$  is given by

$$\hat{r}_{LLF} = -\frac{1}{c} \ln \left( \int_0^1 \frac{(T_1 + \lambda_1)^{n+1} (T_2 + \lambda_2)^{m+1}}{B(n+1, m+1)} \frac{r^n (1-r)^m e^{-cr}}{((T_1 + \lambda_1)r + (T_2 + \lambda_2)(1-r))^{n+m+2}} dr \right). \quad (9)$$

Table 1: reliability parameter estimation for various values of  $p_1$  and  $p_2$  and  $\lambda_1 = \lambda_2 = \frac{1}{2}$

$n$	$m$	$p_1$	$p_2$	Simulated	Actual	Approximated Value by maximum likelihood method (Absolute deviations)	Approximated Value by Bayesian method (ELF) (Absolute deviations)	Approximated Value by Bayesian method (LLF) (Absolute deviations)
10	10	0.1	0.1	0.4990	0.5019(0.0029)	0.6230(0.1240)	0.5990(0.1)	
20	20	0.1	0.1	0.4920	0.4970(0.005)	0.5233(0.0313)	0.5568(0.0648)	
30	30	0.1	0.1	0.5003	0.5008(0.0005)	0.4890(0.0113)	0.5239(0.236)	
40	40	0.1	10	0.4962	0.4973(0.0011)	0.4567(0.0395)	0.3399(0.1563)	
50	50	0.1	10	0.5072	0.5060(0.0012)	0.6198(0.1126)	0.7034(0.1962)	
10	10	1	1	0.5110	0.5005(0.0105)	0.2436(0.2674)	0.3648(0.1462)	
20	20	1	1	0.5015	0.4969(0.0046)	0.1234(0.3781)	0.2569(0.0254)	
30	30	1	1	0.5003	0.4905(0.0053)	0.2098(0.2905)	0.3519(0.1484)	
40	40	1	1	0.5013	0.5041(0.0028)	0.3378(0.1635)	0.3445(0.1568)	
50	50	1	1	0.4988	0.4998(0.0010)	0.3389(0.1599)	0.4500(0.0498)	
10	10	10	10	0.5100	0.4964(0.0136)	0.6778(0.1678)	0.7233(0.2133)	
20	20	10	10	0.4890	0.5054(0.0164)	0.6990(0.21)	0.3458(0.1432)	
30	30	10	10	0.5207	0.4868(0.0339)	0.4567(0.064)	0.4100(0.1107)	
40	40	10	10	0.5033	0.4930(0.0103)	0.6659(0.1626)	0.7700(0.2667)	
50	50	10	10	0.4846	0.5080(0.0234)	0.7033(0.2187)	0.6745(0.1904)	

The analytical solution of the above integrals is not possible. Therefore, we propose the use of approximation technique to solve such integrals. Here, we use Gauss quadrature method.

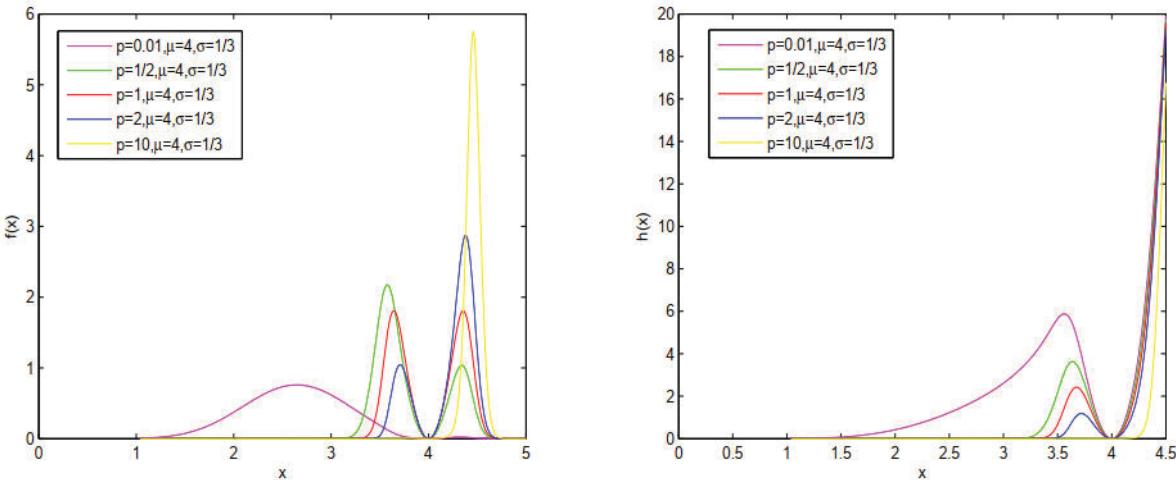


Figure 1. graphs of density and hazard functions of the New Burr distribution for different values of shape parameter  $p$

## 4 Numerical study

In this section, we investigate the performances of the estimators through Monte Carlo simulation. The simulation study is carried out for different variation of  $n$ ,  $m$  and model parameters  $p_1$ ,  $p_2$ , respectively. For exponential prior, we took  $\lambda_1 = \lambda_2 = \frac{1}{2}$ . Note that for other values of  $\lambda_1$  and  $\lambda_2$ , the results are similarly obtained.

All the simulation results are based on 1000 replications and presented in Table 1. As Table 1 shows maximum likelihood method is better than bayesian method for estimation of reliability parameter as distribution of stress and strength are New Burr distribution.

## References

- [1] Burr, I. W. (1942), Cumulative frequency distribution, *Annals of Mathematical Statistics*, **13**, 215-232.
- [2] Dierickx, B., Vleugels, J. and Van der Biest, O. (2000). Statistical extreme value modeling of particle size distributions: experimental grain size distribution type estimation and parameterization of sintered zirconia D. *Materials characterization* 45:61-70.
- [3] Ely, J.T.A., Fudenberg, H.H., Muirhead, R. J., LaMarche, M. G., Krone, C. A., Buscher, D., and Stern, E. A. (1999). Urine mercury in micromercurialism: bimodal distribution and diagnostic implications. *Bulletin of environmental contamination and toxicology* 63(5):553-559.
- [4] Garboś, S. and Świecicka, D. (2015), Application of bimodal distribution to the detection of changes in uranium concentration in drinking water collected by random daytime sampling method from a large water supply zone, *Chemosphere*, 377-382.
- [5] Knapp, T. R. (2007), Bimodality Revisited, *Journal of Modern Applied Statistical Methods*, **6**, 8-20.
- [6] Okasha, M. K. and Matter, M. Y. (2015), On The Three-Parameter Burr Type XII Distribution and its Application to Heavy Tailed Lifetime Data, *Journal Of Advances In Mathematics*, **10**, 3429-3442.
- [7] Surles, J. G. and Padgett, W. J. (2001), Inference for reliability and stress-strength for a scaled Burr Type X distribution, *Lifetime Data Analysis*, **7**, 187-200.
- [8] Sewell, M. A. and Young, C. M. (1997), Are Echinoderm Egg size Distributions Bimodal?, *Biological Bulletin*, 297-305.
- [9] Yari, G. H. and Tondpour, Z. (2017), The new Burr distribution and its application, *Mathematical Sciences*, under publication.
- [10] Zhang, C., Mapes, B. E., and Soden, B. J. (2003). Bimodality in tropical water vapor. *Quarterly Journal of the Royal Meteorological Society* 129:2847-2886.



*Forth Seminar on*  
**Reliability Theory and its Applications**  
*April 25-26, 2018*



## An RSS-based estimator of mean residual life

Zamanzade, E.<sup>1</sup>      Parvardeh, A.<sup>2</sup>      Asadi, M.<sup>3</sup>

<sup>1,2,3</sup> Department of Statistics, University of Isfahan

### Abstract

Mean residual life (MRL) is one of the most applicable measures in survival analysis and reliability theory. In this article, an empirical MRL based on a ranked set sampling (RSS) is described and some of its theoretical properties are explored. Our simulation and numerical results show that the RSS-based estimator of MRL beats its counterpart in simple random sampling (SRS).

**Keywords:** Estimation, Ranked set sample, Mean residual life.

## 1 Introduction

The mean residual life (MRL) function of a non-negative random variable  $X$  at age  $t$  with the survival function  $\bar{F}(t) = P(X > t)$  and a finite mean  $\mu$  is defined as

$$M(t) = E(X - t | X > t) = \frac{\int_t^\infty \bar{F}(x)dx}{\bar{F}(t)}, \quad (1)$$

for  $t$  such that  $\bar{F}(t) > 0$ . It is worth mentioning that  $M(0) = \mu$  and thus it follows from finiteness of  $\mu$  that  $M(t) < \infty$ , for all  $t \in (0, \infty)$ .

The MRL function has wide applications for modeling of lifetime data, reliability and survival analysis.

---

<sup>1</sup>eli.zaman@sci.ui.ac.ir

<sup>2</sup>a.parvardeh@sci.ui.ac.ir

<sup>3</sup>m.asadi@sci.ui.ac.ir

In this article, we describe an estimator for MRL function based on a viable alternative to simple random sampling (SRS), which is called ranked set sampling (RSS). RSS was proposed by [7] for improving performance of mean estimator of a pasture yield and then finds many applications in other branches including medicine ([1]), forestry ([3]), biometrics ([9]), and reliability ([6]).

To draw a ranked set sample of size  $n$ , one first determines the design parameters set size ( $k$ ) and cycle size ( $m$ ) such that  $n = km$ . He then draws  $mk$  simple random samples of size  $k$  from the population. Each sample of size  $k$  is ranked in increasing magnitude without referring to their precise values. Then from the first  $m$  simple random samples of size  $k$ , the units with rank 1 are selected for actual measurement, for the second  $m$  simple random samples of size  $k$ , the units with rank 2 are selected for actual measurement, and the process is continued until the total sample of size  $n = mk$  is collected. The ranking is called perfect if there is no error in the ranking of sample units, otherwise it is called imperfect.

The aim of this paper is to investigate the RSS-based estimation of MRL. The rest of paper is structured as follows. In Section 2, we study the estimation of MRL based on RSS design and its asymptotic properties. A comparison between RSS and SRS based estimators via Monte Carlo simulation are presented in Section 3. Section 4 is devoted to a real data example of described estimator.

## 2 RSS-based estimation of MRL

Let  $X_1, \dots, X_n$  be a simple random sample of size  $n$  from a population with continuous distribution function  $F$  and survival function  $\bar{F} = 1 - F$ . An SRS-based estimator of MRL  $M(t)$  is given by

$$\widehat{M}_{SRS}(t) = \frac{\int_t^{+\infty} \bar{F}_{SRS}(x) dx}{\bar{F}_{SRS}(t)} I(\bar{F}_{SRS}(t) > 0) \quad (2)$$

where  $\bar{F}_{SRS}(x) = \frac{1}{n} \sum_{i=1}^n I(X_i > x)$  is the survival function and  $I(A)$  is the usual indicator function on set  $A$ . The asymptotic properties of this estimator is analysed by [8], [4].

Consider a ranked set sample  $\{X_{[i:k]j}, i = 1, \dots, k; j = 1, \dots, m\}$  of size  $n = km$  from distribution function  $F$ , where  $X_{[i:k]j}$  is the sample unit with judgement rank  $i$  in  $i$ th sample of size  $k$  and in the  $j$ th cycle. Note that  $X_{[i:k]j}$ 's are independent and follow the distribution of the  $i$ th judgement order statistic from a sample of size  $k$ . In this article, we use the subscript  $[.]$  for imperfect ranking and the subscript  $(.)$  for perfect ranking. We also assume that the ranking is consistent which means that the equality  $F(t) = \frac{1}{k} \sum_{i=1}^k F_{[i:k]}(t)$  holds, where  $F_{[i:k]}$  is the distribution function of  $i$ th judgement order statistic of sample of size  $k$ .

Under the above assumptions, [10] showed

$$M(t) = \frac{1}{k} \sum_{i=1}^k \frac{\bar{F}_{[i:k]}(t)}{\bar{F}(t)} M_{[i:k]}(t), \quad (3)$$

where

$$M_{[i:k]}(t) = \frac{\int_t^{+\infty} \bar{F}_{[i:k]}(x) dx}{\bar{F}_{[i:k]}(t)},$$

is the MRL of random variable  $X_{[i:k]j}$ . Note that the empirical estimator of the distribution function  $F$  is given by

$$\hat{F}_{RSS}(t) = \frac{1}{k} \sum_{i=1}^k \hat{F}_{[i:k]}(t),$$

and  $\hat{F}_{[i:k]}(t) = \frac{1}{m} \sum_{j=1}^m I(X_{[i:k]j} \leq t)$  is the empirical distribution function of  $F_{[i:k]}$ .

Therefore, an RSS estimator of  $M(t)$  can be constructed by replacing  $\bar{F}$ ,  $\bar{F}_{[i:k]}$  and  $M_{[i:k]}$  with their empirical counterparts in (3) as follows

$$\widehat{M}_{RSS}(t) = \frac{1}{k} \sum_{i=1}^k \frac{\bar{\hat{F}}_{[i:k]}(t)}{\bar{\hat{F}}_{RSS}(t)} \widehat{M}_{[i:k]}(t) I(\bar{\hat{F}}_{RSS}(t) > 0), \quad (4)$$

where  $\bar{\hat{F}}_{[i:k]}(t) = 1 - \hat{F}_{[i:k]}(t)$ ,  $\bar{\hat{F}}_{RSS}(t) = 1 - \hat{F}_{RSS}(t)$  and

$$\widehat{M}_{[i:k]}(t) = \frac{\int_t^{+\infty} \bar{\hat{F}}_{RSS}(x) dx}{\bar{\hat{F}}_{RSS}(t)} I(\bar{\hat{F}}_{[i:k]}(t) > 0),$$

is an estimate of  $M_{[i:k]}(t) = E(X_{[i:k]} - t | X_{[i:k]} > t)$ . It is worth noting that the estimator in (4) was firstly proposed by [10].

In what follows, we describe the asymptotic properties of MRL estimator in RSS and compare its asymptotic performance with its counterpart in SRS. It is shown that the estimator in (4) is asymptotically at least as efficient as the estimator in (2). Note that the following theorems have been already proved by [10], so we only provide a sketch proof for them and refer the interested reader to [10] for more details.

In the first theorem, the asymptotic normality of (4) is established.

**Theorem 2.1.** *Let  $\{X_{[i:k]j}, i = 1, \dots, k; j = 1, \dots, m\}$ , be a ranked set sample of size  $n = km$  from a population with distribution function  $F$ . Then for all  $T > 0$ ,*

$$\left\{ Z_n^{RSS}(t) \equiv \sqrt{n} (\widehat{M}_{RSS}(t) - M(t)), t \in [0, T] \right\},$$

*converges in distribution to a mean zero Gaussian process  $\{Z_{RSS}(t), t \in [0, T]\}$  with variance function*

$$\sigma_{ARSS}^2(t) = \frac{1}{k \bar{F}^2(t)} \left\{ \sum_{i=1}^k \sigma_{[i:k]}^2(t) \bar{F}_{[i:k]}(t) + \bar{F}_{[i:k]}(t) F_{[i:k]}(t) [M_{[i:k]}(t) - M(t)]^2 \right\}, \quad (5)$$

where

$$\sigma_{[i:k]}^2(t) = Var(X_{[i:k]} - t | X_{[i:k]} > t) = \int_t^\infty \frac{(x-t)^2}{\bar{F}_{[i:k]}(t)} dF_{[i:k]}(x) - M_{[i:k]}^2(t).$$

*Proof.* By [8], [4], it can be simply shown that the corresponding limiting process  $Z_{RSS}$  is obtained as

$$\begin{aligned} Z_{RSS}(t) &= \frac{-1}{\sqrt{k}\bar{F}(t)} \sum_{i=1}^k \int_t^{+\infty} U_{[i:k]}(x)dx + \frac{M(t)}{\sqrt{k}\bar{F}(t)} \sum_{i=1}^k U_{[i:k]}(t) \\ &= \frac{-1}{\sqrt{k}\bar{F}(t)} \left[ \sum_{i=1}^k \left( \int_t^{+\infty} U_{[i:k]}(x)dx - M(t)U_{[i:k]}(t) \right) \right], \quad t \in [0, T], \end{aligned}$$

where  $U_{[i:k]} = \{U_{[i:k]}(t), t \geq 0\}$  a Gaussian process with mean zero and covariance function

$$E(U_{[i:k]}(t_1)U_{[i:k]}(t_2)) = F_{[i:k]}(t_1 \wedge t_2) - F_{[i:k]}(t_1)F_{[i:k]}(t_2), \quad t_1, t_2 \geq 0.$$

Then by calculating the covariance function  $Cov(Z_{RSS}(t_1), Z_{RSS}(t_2))$  the result will be obtained.  $\square$

The next theorem shows that the RSS-based estimator of MRL is at least as efficient asymptotically as the SRS-based one. Recall that in the SRS,  $\sqrt{n}(\widehat{M}_{SRS} - M)$  converges in distribution to a Gaussian process  $Z_{SRS}$  with mean zero and variance function

$$\sigma_{ASRS}^2 = \frac{\sigma^2(t)}{\bar{F}(t)}, \quad 0 \leq t < \infty, \quad (6)$$

as  $n$  goes to infinity, where

$$\sigma^2(t) = Var(X - t|X > t), \quad (7)$$

(see, [4]).

**Theorem 2.2.** For all  $t \geq 0$ , we have  $Var(Z_{RSS}(t)) \leq Var(Z_{SRS}(t))$ .

*Proof.* It follows from the identity  $k\bar{F}(t) = \sum_{i=1}^k \bar{F}_{[i:k]}(t)$  and (7) that

$$k\sigma^2(t) = \sum_{i=1}^k \frac{\bar{F}_{[i:k]}(t)}{\bar{F}(t)} \sigma_{[i:k]}^2(t) + \sum_{i=1}^k \frac{\bar{F}_{[i:k]}(t)}{\bar{F}(t)} [M_{[i:k]}(t) - M(t)]^2.$$

Therefore, we have

$$\begin{aligned} kVar(\bar{F}(t)Z_{SRS}(t)) &= k\sigma^2(t)\bar{F}(t) \\ &= \sum_{i=1}^k \bar{F}_{[i:k]}(t)\sigma_{[i:k]}^2(t) + \sum_{i=1}^k \bar{F}_{[i:k]}(t)[M_{[i:k]}(t) - M(t)]^2 \\ &\geq \sum_{i=1}^k \bar{F}_{[i:k]}(t)\sigma_{[i:k]}^2(t) + \sum_{i=1}^k \bar{F}_{[i:k]}(t)F_{[i:k]}(t)[M_{[i:k]}(t) - M(t)]^2 \\ &= kVar(\bar{F}(t)Z_{RSS}(t)). \end{aligned}$$

$\square$

### 3 Comparison

A comparison between the RSS and SRS-based estimators of MRL for gamma, and exponential distributions was made by [10]. In this section, we compare the estimators when the parent distribution is Weibull.

We first compare the RSS and SRS-based estimator of MRL for finite sample sizes. To do so, we define the relative efficiency (RE) as follows

$$RE(t) = \frac{MSE(\widehat{M}_{SRS}(t))}{MSE(\widehat{M}_{RSS}(t))}. \quad (8)$$

A value of  $RE(t)$  greater than 1 indicates that the RSS-based estimator of MRL is better than the SRS-based one. A Monte Carlo simulation study with 1,000,000 repetitions is used to estimate  $RE(t)$ . To do so, we first set  $k \in \{3, 5\}$  and  $n \in \{15, 30, 90\}$  and then for each  $(k, n)$ , 1,000,000 random samples of size  $n = mk$  are generated from the Weibull distribution with shape parameter 2 and scale parameter 2 (denoted by Weibull(2,2)) which has decreasing MRL. The ranking mechanism considered in the simulation study to generate an RSS sample is fraction of random ranking model due to [2]. In this model, the distribution function of  $i$ th judgement order statistics is considered as a mixture of distribution function of  $i$ th order statistics and parent distribution function as following

$$F_{[i:k]} = pF_{(i:k)} + (1 - p)F \quad p \in [0, 1]$$

where  $p \in [0, 1]$  controls the quality of ranking process. The selected values of  $p$  are as  $p \in \{1, 0.8, 0.5, 0.2\}$ . We estimate  $RE(t)$  values for  $t \in \{Q_{0.05}, \dots, Q_{0.95}\}$ , where  $Q_s$  is the  $s$ th quantile of the Weibull distribution for  $s = 0.05, \dots, 0.95$ .

The simulation results are presented in Figures 1. Figure 1 shows the simulation results when parent distribution is Weibull with decreasing MRL. The three top panels in Figure 1 presents the results for  $n = 15, 30$  and  $90$  for the perfect ranking case. It is clear from these panels that the efficiency gained of using the RSS-based estimator relative to SRS-based on can be as large as 225% for  $k = 5$  and small values of  $t$ . Moreover, even for large values of  $t$ , the RSS-based estimator is almost as efficient as the SRS-based one. It is of interest to note that the  $RE(t)$  does not have a clear dependence on sample size ( $n$ ) when  $k$  is fixed. This finding is consistent with what is observed in the most of the RSS literature. The results of  $p = 0.8$  and  $p = 0.5$  show that similar pattern to those  $p = 1$ , with a obvious difference that the  $RE(t)$  values are lower. We also observe that almost no efficiency gain is obtain for  $p = 0.2$ .

Next, we compare the asymptotic relative efficiency of the estimators for Weibull distribution with shape parameter 2. To do so, the asymptotic efficiency (ARE) of  $\widehat{M}_{RSS}(t)$  relative to  $\widehat{M}_{SRS}(t)$  is defined as

$$ARE(t) = \frac{\sigma_{ARSS}^2(t)}{\sigma_{ASRS}^2(t)}. \quad (9)$$

Note that one can conclude from Theorem 2.2,  $ARE(t) \geq 1$  for all  $t \geq 0$ . Figure 2 presents exact values of  $ARE(t)$  evaluated at the point  $t$  for  $t \in \{Q_{0.05}, \dots, Q_{0.95}\}$  where  $Q_s$  is the

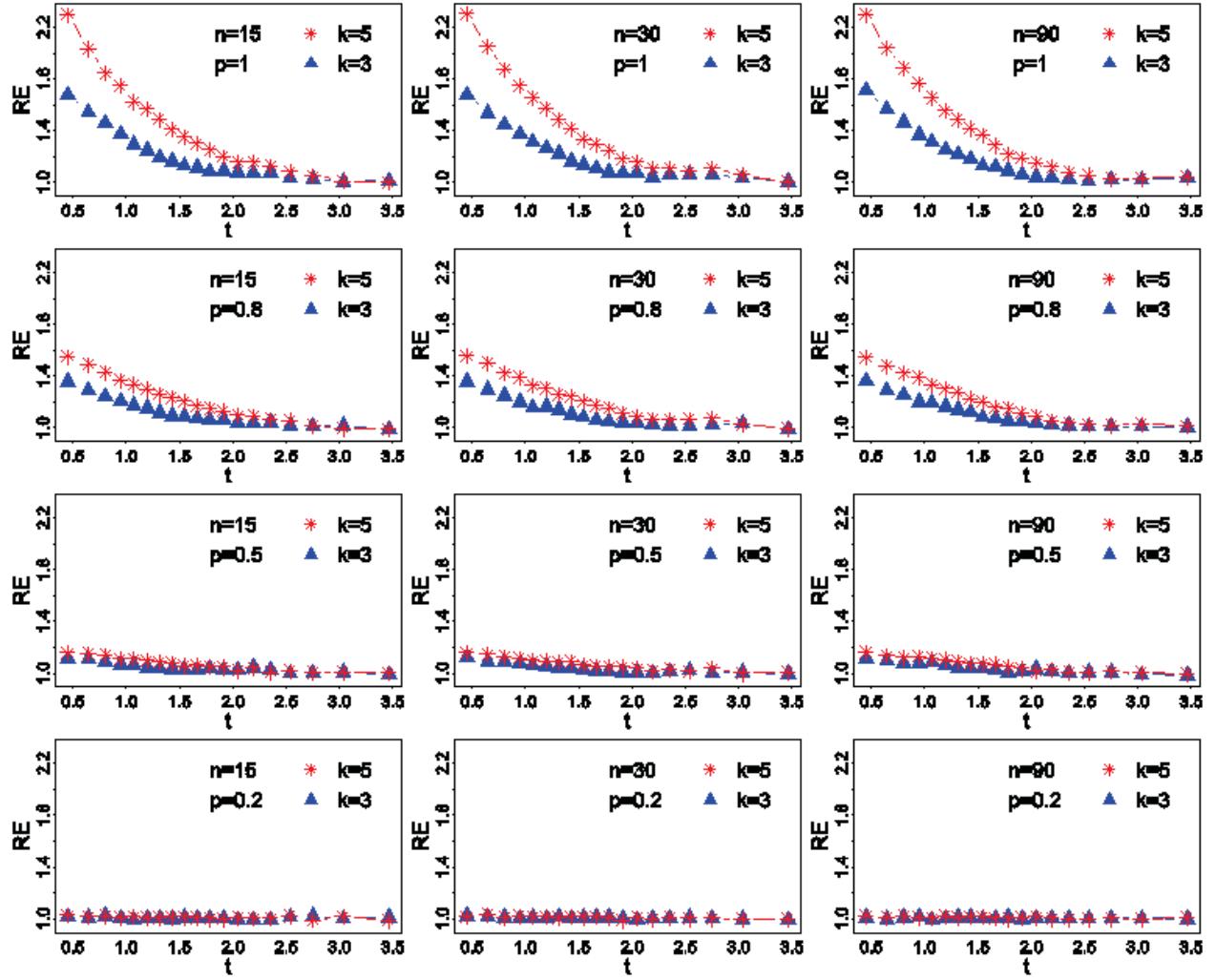


Figure 1: Simulated  $RE(t)$  as a function of  $t$  for  $k = 3$  and  $5$  (represented by  $\blacktriangle$  and  $*$  respectively),  $n \in \{15, 30, 90\}$  and  $p \in \{1, 0.8, 0.5, 0.2\}$  when the parent distribution is Weibull(2,2). This figure appears in color in the electronic version of this paper.

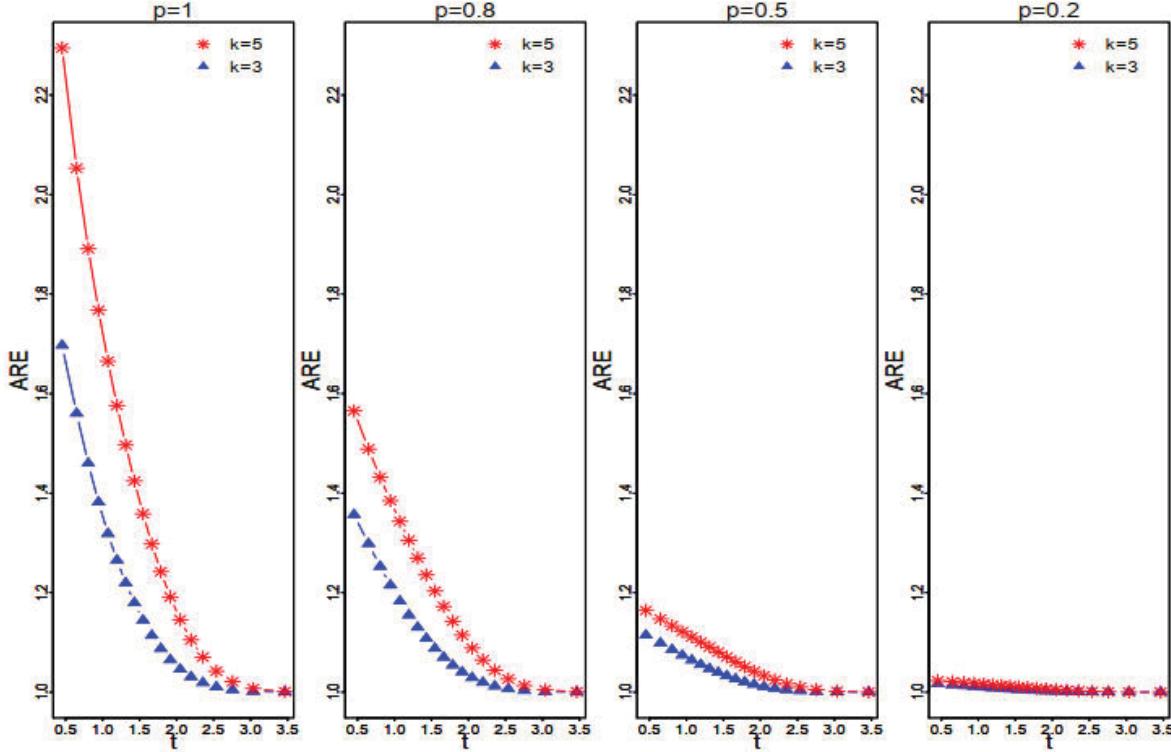


Figure 2: Exact values of  $ARE(t)$  as a function of  $t$  for  $k = 3, 5$  (represented by  $\blacktriangle$  and  $*$  respectively) and  $p \in \{1, 0.8, 0.5, 0.2\}$  when the parent distribution is Weibull(2,2). This figure appears in color in the electronic version of this paper.

sth quantile of Weibull(2,2) distribution and different values of  $p \in \{1, 0.8, 0.5, 0.2\}$ . By comparing Figure 2 with Figure 1, respectively, we observe that the  $ARE(t)$  curves are very close to  $RE(t)$ . Therefore, the relative efficiency for a finite sample size can be approximated by the asymptotic one even in the imperfect ranking case.

## 4 Empirical evidence

In this section, we use a real data set to compare MRL estimators in the SRS and RSS settings. The data is on Monoclonal gammopathy of undetermined significance (MGUS) which is one of the plasma cell disorders, and older people are more prone to suffer from it. This disorder occurs when some abnormal but non-cancerous plasma cells start to produce a large amount of monoclonal antibodies. MGUS has no symptoms and it is usually identified by chance when the patient gives some laboratory tests for other reasons. Although MGUS is benign and does not make any significant problems, but for some unknown reasons it may progress to some kinds of cancers such as blood cancer. Therefore, patients who suffer from MGUS should have regular checkups.

The observations used in this section is natural history of 241 patients with an apparently MGUS at Mayo Clinic on 12 variables before Jan. 1, 1971, underwent prospective follow-up for 20 to 35 years [5]. This data set is known as MGUS data set and is now available in `mgus` package of R statistical software. We consider MGUS data set as our given population and we are interested in point and interval estimation of residual lifetime of patients from diagnosis of a plasma cell malignancy until to death at a given time  $t$ . Therefore, we only consider 225 patients who died during the experiment and we

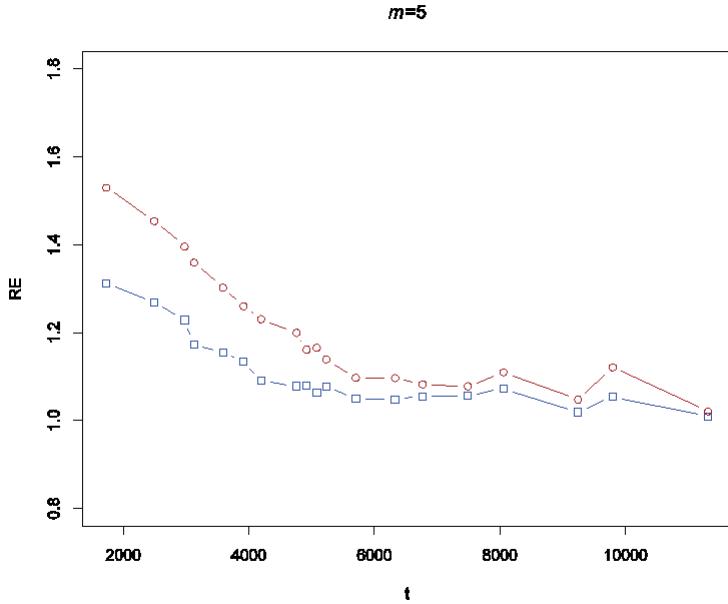


Figure 3: Estimated  $RE(t)$  for  $k = 3, 5$  (represented by  $\square$  and  $\circ$ , respectively),  $m = 5$  using MGUS data set. This figure appears in color in the electronic version of this paper.

take `futime`, days from diagnosis of a plasma cell malignancy in the patient to his/her last follow-up for those patients as our interest variable. To draw a ranked set sample from the population, we assume that the ranking process in each set of size  $k$  is made according to the concomitant variable `pctime`, days from detection of MGUS in the patient until diagnosis of a plasma cell malignancy in his/her blood. It is worth noting that the correlation coefficient between variable of the interest and the concomitant variable in this population is 0.83.

In order to compare the point estimators of MRL function in the SRS and RSS settings using this date set, we draw a set of 100,000 simple random samples and a set of 100,000 ranked set samples with cycle size  $m = 5$  and set sizes  $k = 3$  and 5. Sampling with replacement is considered here, so the assumption of independence is guaranteed. The relative efficiency of the RSS-based MRL estimator to its SRS-based counterpart is defined as in Section 3, and is evaluated for  $t \in \{Q_{0.05}, \dots, Q_{0.95}\}$ , where  $Q_s$  is the  $s$ th quantile of the population. The results are presented in Figure 3.

Figure 3 shows the preference of the RSS-based estimator to its SRS-based competitor. We observe that the  $RE(t)$  increases as set size increases while the other parameters are fixed, and the efficiency gain using the RSS estimator can be as large as 150% for  $k = 5$  and small values of  $t$ .

## 5 Conclusion

In this paper, the RSS-based estimator of MRL due to [10] is described and some of its theoretical properties is explored. A Monte Carlo simulation is used to show the superiority of the described procedure to its competitor in the literature. Finally, as an

application of the described estimator, a real data set, called MGUS which is one of the plasma cell disorders in some patients, was employed to estimate the MRL of days from diagnosis of a plasma cell malignancy until to death of the patients.

## References

- [1] Chen, H., Stasny, E. A., & Wolfe, D. A. (2007). Improved procedures for estimation of disease prevalence using ranked set sampling. *Biometrical Journal* **49**, 530-538.
- [2] Frey, J., Ozturk, O. and Deshpande, J. (2007), Nonparametric tests for perfect judgment rankings, *Journal of Statistical Planning and Inference*, **138**, 708–717.
- [3] Halls, L. K., & Dell, T. R. (1966). Trail of ranker set sampling for forage yields. *Forest Science* **12**, 22-26.
- [4] Hall, W. J., & Wellner, J. A. (1979). Estimation of mean residual life. Unpublished manuscript.
- [5] Kyle, R. (1993). Benign monoclonal gammopathy-after 20 to 35 years of follow-up. *Mayo Clinic Proceedings* **68**, 26-36.
- [6] Mahdizadeh, M. and Zamanzade, E. (2017). A new reliability measure in ranked set sampling, To appear in *Statistical Papers* <https://doi.org/10.1007/s00362-016-0794-3>
- [7] McIntyre, G.A. (1952). A method for unbiased selective sampling, using ranked sets. *Crop and Pasture Science* **3**, 385-390.
- [8] Yang, G. L. (1978). Estimation of a Biometric Function. *The Annals of Statistics* **6**, 112-116.
- [9] Zamanzade, E., and Mahdizadeh, M. (2017). A more efficient proportion estimator in ranked set sampling. *Statistics and Probability Letters*, 129, 28-33.
- [10] Zamanzade, E., Parvardeh, A., and Asadi, M. (2018). Estimation of mean residual life based on ranked set sampling, Submitted for publication.