Proceeding of The 3rd Seminar on Reliability Theory and its Applications

May 17-16 2017

Department of Statistics ,Ferdowsi University of Mashhad

Proceeding of

the 3rd Seminar on

Reliability Theory and its Applications

Department of Statistics

and

Ordered and Spatial Data Center of Excellence
Ferdowsi University of Mashhad,
Mashhad, Iran

16-17 May, 2017

Disclaimer

This book contains the proceeding of the 3rd Seminar on **Reliability Theory and its Applications**. Authors are responsible for the contents and accuracy. Opinions expressed may not necessarily reflect the position of the scientific and organizing committees.

عنوان: مجموعه مقالات انگلیسی سمینار تخصصی نظریه قابلیت اعتماد و کاربردهای آن

تدوین کننده : جعفر احمدی

صفحه آرا : زهره پاکدامن

طرح جلد: احمد دادرس جوان

ناشر: کمیته برگزاری سمینار

شمارگان: ۱۲۰

تاریخ انتشار: خرداد ۱۳۹۶

Preface

Continuing the series of workshops on "Reliability Theory and its Applications" in Ferdowsi University of Mashhad and two Seminars in University of Isfahan (2015) and University of Tehran (2016), we are pleased to organize the 3rd Seminar on "**Reliability Theory and its Applications**" during 16-17 May, 2017 at the Department of Statistics, Ferdowsi University of Mashhad. 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 Mashhad 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 and organizations that have contributed to the success of this seminar, in which around 100 colleagues, researchers, and postgraduate students have participated.

Finally, we would like to extend our sincere gratitude to the administration of Ferdowsi University of Mashhad and Faculty of Mathematical Sciences, the Iranian Statistical Society, the Ordered and Spatial Data Center of Excellence, the Scientific Committee, the Organizing Committee and the students of the Department of Statistics at Ferdowsi University of Mashhad for their kind cooperation.

Jafar Ahmadi (Chair)
May, 2017

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
- Degradation models
- Information theory topics in reliability
- Maintainability and availability
- Measures of dependence in reliability
- Network reliability
- Reliability of coherent systems
- Statistical inference for reliability data
- Stochastic aging
- Stochastic ordering and reliability
- Stress-strength modeling
- Optimization methods in reliability

Scientific Committee

1. Ahmadi, J., Ferdowsi University of Mashhad (Iran)
2. Arghami, N.R., Ferdowsi University of Mashhad (Iran)
3. Asadi, M., University of Isfahan (Iran)
4. Asgharzadeh, A., Mazandarn University (Iran)
5. Balakrishnan, N., McMaster University (Canada)

6. Bayramoglu, I., Izmir University of Economics (Turkey)
7. Di Crescenzo. A., Universit degli Studi di Salerno (Italy)
8. Doostparast, M. Ferdowsi University of Mashhad (Iran)
9. Eryilmaz, S., Atilim University, (Turkey)
10. Haghighi, F., University of Tehran (Iran)
11. Hamamadani, A. Z., Isfahan University of Technology (Iran)
12. Khaledi, B.k, Razi University (Iran)
13. Khanjar, M., Birjand University (Iran)
14. Khodadadi, A., Shahid Beheshti University (Iran)
15. Mothasihemi Borzadaran, G. R., Ferdowsi University of Mashhad (Iran)
16. Navarro, J., Universidad de Murcia (Spain)
17. Rezaei Roknabadi, A. H. Ferdowsi University of Mashhad (Iran)
18. Singpurwalla, N. D., City University of Hong Kong (Hong Kong)

Organizing Committee

1. Ahmadi, J., Ferdowsi University of Mashhad
2. Asadi, M., University of Isfahan
3. Amini, M. Ferdowsi University of Mashhad
4. Emadi, M., Ferdowsi University of Mashhad
5. Fashandi, M. Ferdowsi University of Mashhad
6. Jabari, Nooghabi, H. Ferdowsi University of Mashhad
7. Jabari Nooghabi, M. Ferdowsi University of Mashhad
8. Habibi Rad A., Ferdowsi University of Mashhad
9. Mothasihemi Borzadaran, G., Ferdowsi University of Mashhad
10. Razmkhah.M. Ferdowsi University of Mashhad
11. Rezaei Roknabadi, A. H. Ferdowsi University of Mashhad
12. Sadeghpour Gildeh, B. Ferdowsi University of Mashhad

Contents

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

A note on estimation based on joint progressively first-failure-censored data

Ahmadi, M. V. [1]

Department of Statistics, University of Bojnord

Abstract

In this paper, a new life test plan called a joint progressive first-failure-censoring scheme is introduced. The maximum likelihood estimates and the Bayes estimates for the parameters of two exponential distributions are discussed for the new censoring scheme. Moreover, the Bayes estimates are investigated under symmetric and asymmetric loss functions. Finally, a simulation study is performed and an illustrative example is also given.

Keywords: Bayes estimation, Joint progressive first-failure-censoring scheme, Exponential distribution, LINEX loss function, General entropy loss function.

1 Introduction

Censoring is usual in lifetime data due to time and cost restrictions. There are various types of censoring in survival analysis and progressive censoring is one of the most common for consideration. This censoring allows the experimenter to remove the units from a life test at various stages during the experiment. For a comprehensive review of theory, methods and applications of the progressive censoring, we refer the reader to Balakrishnan and Aggarwala [1] and the references contained therein. Progressive first-failure censoring, introduced by Wu and Kuş [9], is a type of progressive censoring in which n disjoint groups with k identical units within each group are placed on a life-testing experiment at time zero. The experimenter terminates the life testing when r failures

[1] mv.ahmadi@ub.ac.ir

occur. When the i-th failure occurs ($i = 1, \ldots, r-1$), randomly selected R_i groups and the group including the i-th failure are withdrawn from the experiment. When the r-th failure occurs, all of the remaining groups are withdrawn from the experiment. For $k = 1$, the progressive first-failure censoring is reduced to the case of progressive Type-II censoring. If $R_1 = \cdots = R_r = 0$, we have the first-failure censoring. If $k = 1$, $R_1 = \cdots = R_{r-1} = 0$ and hence $R_r = n - r$, this censoring is reduced to the Type-II censoring and finally, if $k = 1$, $R_1 = \cdots = R_r = 0$, we have the complete sample.

Almost all of the types of censoring are concerned with the one-sample problems. But, there are situations in which the experimenter plans to compare different populations. In such problems, the joint censoring scheme has been suggested in the literature. Suppose the products are being manufactured by two different lines within the same facility, and that two independent samples of sizes m and n are selected from these lines, denoted by L_1 and L_2, and put simultaneously on a life-testing experiment. Then, to save time and money, the experimenter follows a joint progressive Type II censoring scheme and terminates the life testing when r failures occur. Immediately following the first failure, R_1 of the surviving units are randomly selected and withdrawn from the experiment (we can partition R_1 into S_1 and T_1 such that S_1 and T_1 are the the number of units withdrawn at the time of the i-th failure that belong to lines L_1 and L_2, respectively, and hence $R_1 = S_1 + T_1$). Then, immediately following the second failure, R_2 of the surviving units are randomly selected and withdrawn from the experiment ($R_2 = S_2 + T_2$) and so on. When the r-th failure occurs,all of the remaining $R_r = m + n - r - R_1 - \cdots - R_{r-1}$ surviving units are withdrawn from the experiment. Note that if $R_1 = \cdots = R_{r-1} = 0$, then $R_r = m + n - r$, which corresponds to the joint Type-II censoring. Also, if $R_1 = \cdots = R_r = 0$, then $r = m + n$, which corresponds to the complete sample.

Statistical inferences on the basis of joint censoring have been considered in the literature. Under the joint progressive Type-II censoring scheme, Rasouli and Balakrishnan [7] studied the exact likelihood inference for two exponential populations, while Doostparast *et al.* [4] investigated the Bayesian inference for two Weibull populations. Recently, Balakrishnan *et al.* [2] generalize the work of Rasouli and Balakrishnan [7] from the two-sample to the r-sample situation and discuss the exact likelihood inference under the r exponential populations.

In this paper, we will combine the concepts of joint progressive Type-II censoring and progressive first-failure censoring to develop a new life test plan called a joint progressive first-failure-censoring scheme. The rest of this paper is organized as follows. In Section 1.2, we describe the formulation of a joint progressive first-failure-censoring scheme. In Sections 1.3 and 1.4, we discuss the maximum likelihood estimates and the Bayes estimates for the unknown parameters based on a joint progressively first-failure-censored sample coming from two exponential distributions. In Section 3, a simulation study is conducted on the basis of the Monte Carlo method for comparing the estimated risks of the estimates obtained in Sections 1.3 and 1.4. Finally, an illustrative example is given in Section 4.

2 A joint progressive first-failure-censoring scheme

In this section, progressive first-failure censoring is combined with joint progressive Type-II censoring. Suppose that products are being manufactured by two different lines L_1 and L_2 under the same conditions. Also, suppose that m independent groups with k_1 units within each group and n independent groups with k_2 units within each group have been selected randomly from the two lines L_1 and L_2, respectively. Moreover, consider the distribution functions of the two populations as $F(\cdot)$ and $G(\cdot)$, respectively. All $N = m + n$ groups are placed on a life-testing experiment at time zero and the experiment is terminated as soon as r failures are observed. In order to carry out the experiment according to a joint progressive first-failure-censoring scheme, immediately following the first failure, randomly selected R_1 groups and the group in which the first failure is observed, are withdrawn from the experiment. The R_1 withdrawn groups are divided by two types according to the populations L_1 and L_2 and so S_1 and T_1 denote the number of withdrawn groups belonging to the populations L_1 and L_2, respectively. It is clear that S_1 and T_1 are random quantities and $S_1 + T_1 = R_1$. Then, immediately following the second failure, randomly selected R_2 groups (S_2 groups from L_1 and T_2 groups from L_2) and the group in which the second failure is observed, are withdrawn from the experiment an so on. When the r-th failure occurs, all remaining $R_r = m + n - r - R_1 - \cdots - R_{r-1}$ groups (including S_r groups from L_1 and T_r groups from L_2) and the group in which the r-th failure is observed are withdrawn from the experiment. In this joint censoring scheme $\mathbf{R} = (R_1, \ldots, R_r)$ and r are pre-determined. Similar to the usual joint progressive Type-II censoring, the available data consist of $(\mathbf{W}, \mathbf{Z}, \mathbf{S})$ where $\mathbf{W} = (W_{1:r:N}, \ldots, W_{r:r:N})$ and $\mathbf{Z} = (Z_1, \ldots, Z_r)$ with $Z_i = 1$ if the i-th failure is from population L_1 and $Z_i = 0$, otherwise, and $\mathbf{S} = (S_1, \ldots, S_r)$. By assuming $\bar{F}(\cdot) = 1 - F(\cdot)$ and $\bar{G}(\cdot) = 1 - G(\cdot)$, the joint density function of $(\mathbf{W}, \mathbf{Z}, \mathbf{S})$ is

$$f^*(\mathbf{w}, \mathbf{z}, \mathbf{s}) = C \prod_{i=1}^{r} f(w_i)^{z_i} g(w_i)^{1-z_i} \bar{F}(w_i)^{k_1(s_i+z_i)-z_i} \bar{G}(w_i)^{k_2(t_i+1-z_i)-(1-z_i)} \tag{2.1}$$

where $C = D_1 D_2$,

$$D_1 = k_1^{m_1} k_2^{m_2} \prod_{j=1}^{r} \Big[\Big\{ m - \sum_{i=1}^{j-1} z_i - \sum_{i=1}^{j-1} s_i \Big\} z_j + \Big\{ n - \sum_{i=1}^{j-1} (1 - z_i) - \sum_{i=1}^{j-1} t_i \Big\} (1 - z_j) \Big],$$

$$D_2 = \prod_{j=1}^{r} \frac{\binom{m - \sum_{i=1}^{j} z_i - \sum_{i=1}^{j-1} s_i}{s_j} \binom{n - \sum_{i=1}^{j} (1-z_i) - \sum_{i=1}^{j-1} t_i}{t_j}}{\binom{N - j - \sum_{i=1}^{j-1} R_i}{R_j}},$$

with $m_1 = \sum_{i=1}^{r} z_i$, $m_2 = r - m_1$, $t_i = R_i - s_i$ $(i = 1, \ldots, r)$ and $w_1 < \cdots < w_r$. It need to be mentioned here:

1. If $k_1 = k_2 = 1$, the joint progressive first-failure censoring reduces to the joint progressive Type-II censoring.

2. If $R_1 = \cdots = R_r = 0$, the joint progressive first-failure censoring corresponds to the joint first-failure censoring.

3. If $k_1 = k_2 = 1$ and $R_1 = \cdots = R_{r-1} = 0$, then $R_r = N - r$ and the joint Type-II censoring is obtained.

4. If $k_1 = k_2 = 1$ and $R_1 = \cdots = R_r = 0$, the joint progressively first-failure-censored data are reduced to complete sample.

Obviously, it can be seen that the progressively first-failure-censored sample $(\mathbf{W}, \mathbf{Z}, \mathbf{S})$ can be viewed as a progressively Type-II censored sample coming from two populations with distribution functions $1-\bar{F}(\cdot)^{k_1}$ and $1-\bar{G}(\cdot)^{k_2}$. Therefore, all statistical inferences based on the joint progressive Type-II censoring can be extended to those based on the joint progressive first-failure censoring. However, the joint progressive first-failure censoring is worth to be used in terms of reducing test time and a saving of resources especially for product with long lifetime but low price.

3 Maximum likelihood estimation

Suppose that the lifetimes of units produced by L_1 follow a one-parameter exponential distribution, denoted by $\text{Exp}(\theta_1)$, with the density and distribution functions as

$$f(x) = \theta_1 \exp(-\theta_1 x) \text{ and } F(x) = 1 - \exp(-\theta_1 x), \ x > 0, \tag{3.1}$$

respectively. Similarly, Suppose that the lifetimes of units produced by L_2 follow another one-parameter exponential distribution, denoted by $\text{Exp}(\theta_2)$, with the density and distribution functions as

$$g(x) = \theta_2 \exp(-\theta_2 x) \text{ and } G(x) = 1 - \exp(-\theta_2 x), \ x > 0, \tag{3.2}$$

respectively. Now, by substituting Equations (3.1) and (3.2) into Equation (2.1), the likelihood function of (θ_1, θ_2) is reduced to

$$L(\theta_1, \theta_2; \mathbf{w}, \mathbf{z}, \mathbf{s}) = C\, \theta_1^{m_1} \theta_2^{m_2} \exp\Big\{ -(\theta_1 u_1 + \theta_2 u_2)\Big\}, \tag{3.3}$$

where

$$u_1 = k_1 \sum_{i=1}^{r} w_i(z_i + s_i) = k_1 \left(\sum_{i=1}^{m_1} x_{(i)} + \sum_{i=1}^{r} w_i s_i \right),$$

and

$$u_2 = k_2 \sum_{i=1}^{r} w_i(1 - z_i + t_i) = k_2 \left(\sum_{i=1}^{m_2} y_{(i)} + \sum_{i=1}^{r} w_i t_i \right),$$

while $x_{(1)} < \cdots < x_{(m_1)}$ and $y_{(1)} < \cdots < y_{(m_2)}$ are the order statistics of the failure times $w_1 < \cdots < w_r$ from L_1 and L_2, respectively.

The ML estimates of the parameters θ_1 and θ_2 are obtained by maximizing the likelihood function in Equation (3.3) with respect to θ_1 and θ_2, respectively. After some algebraic computations, the ML estimates of θ_1 and θ_2 are derived as

$$\hat{\theta}_{i,ML} = \frac{M_i}{U_i}, \quad i = 1, 2. \tag{3.4}$$

Remark 3.1. From the ML estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ in Equation (3.4), it can be seen immediately that when $M_1 = \sum_{i=1}^{r} Z_i = 0$ or r, $\hat{\theta}_1$ and $\hat{\theta}_2$ does not exist. Hence, $\hat{\theta}_1$ and $\hat{\theta}_2$ are only conditional ML estimates, conditioned on $1 \leq M_r \leq r-1$. We, therefore, need to discuss the properties of the ML estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ only conditional on $1 \leq M_r \leq r-1$.

4 Bayes estimation

Statistical inference based on the Bayesian approach is fundamentally different from the non-Bayesian one. The Bayesian approach allows one to incorporate prior subjective knowledge or technical information concerning the lifetime parameters into the inferential procedures. In lifetime data analysis, such prior knowledge is usually summarized into a prior density, denoted by $\pi(\cdot)$. In what follows, we assume that the the paramaters θ_1 and θ_2 are two independent random variables with gamma prior distributions $Gamma(a_1, b_1)$ and $Gamma(a_2, b_2)$, respectively. Then, the joint prior distribution of (θ_1, θ_2) is

$$\pi(\theta_1, \theta_2) = \pi_1(\theta_1)\pi_2(\theta_2), \quad \theta_1 > 0, \ \theta_2 > 0, \tag{4.1}$$

where

$$\pi_i(\theta_i) = \frac{b_i^{a_i}}{\Gamma(a_i)} \theta_i^{a_i - 1} \exp\left(-b_i \theta_i\right), \quad i = 1, 2,$$

and $\Gamma(\cdot)$ denotes the complete gamma function. On the basis of the observed censored sample $(\mathbf{w}, \mathbf{z}, \mathbf{s})$, the joint posterior density function of (θ_1, θ_2) is obtained from Equations (3.3) and (4.1) as

$$\pi(\theta_1, \theta_2 | \mathbf{w}, \mathbf{z}, \mathbf{s}) = \pi_1(\theta_1 | \mathbf{w}, \mathbf{z}, \mathbf{s}) \pi_2(\theta_2 | \mathbf{w}, \mathbf{z}, \mathbf{s}), \quad \theta_1 > 0, \ \theta_2 > 0, \tag{4.2}$$

where

$$\pi_i(\theta_i | \mathbf{w}, \mathbf{z}, \mathbf{s}) = \frac{(u_i + b_i)^{m_i + a_i}}{\Gamma(m_i + a_i)} \theta_i^{m_i + a_i - 1} \exp\left(-\theta_i (u_i + b_i)\right), \quad i = 1, 2.$$

From Equation (4.2), we see that the joint posterior density function of (θ_1, θ_2) is a product of two gamma density functions. So the posterior density functions of θ_1 and θ_2 are $Gamma(m_1 + a_1, u_1 + b_1)$ and $Gamma(m_2 + a_2, u_2 + b_2)$, respectively.

4.1 Loss function

In the Bayesian setup, the choice of a loss function is an integral part. A wide variety of loss functions have been discussed in the literature. One of the most popular loss functions is the Squared Error (SE) loss defined by $L(\delta, \theta) = (\delta - \theta)^2$, where δ is an estimate of θ. But, the SE loss function is justified only when losses are symmetric in nature. The symmetric nature of this loss function gives equal weight to overestimation as well as underestimation, while in practice, overestimation may be more serious than underestimation or vice versa. Such conditions are very common in engineering, medical and biomedical sciences. In this case, an asymmetric loss function might be more appropriate. A suitable alternative to the SE loss function is a convex but

asymmetric loss function, known as the LINear-EXponential (LINEX) loss function, proposed by Varian [8] and defined by

$$L(\delta,\theta) \propto \exp\left\{\tau(\delta-\theta)\right\} - \tau(\delta-\theta) - 1, \quad \tau \neq 0,$$

where τ is the shape parameter. Obviously, the nature of the LINEX loss function changes according to the choice of τ. The sign and magnitude of the shape parameter τ represents the direction and degree of symmetry, respectively. ($\tau > 0$ means overestimation is more serious than underestimation and $\tau < 0$ means the opposite). The LINEX loss converges to the SE loss as $\tau \to 0$. Since the Bayes estimate of θ is the value that minimizes the posterior mean of $L(\delta,\theta)$; therefore, it is easy to verify that the Bayes estimate of θ under the LINEX loss function is obtained as

$$\delta_{BL,\tau} = -\frac{1}{\tau}\ln E\big(\exp(-\tau\theta)|\mathbf{X}\big), \tag{4.3}$$

provided that the expectation $E\big(\exp(-\tau\theta)|\mathbf{X}\big)$ exists.

Another useful asymmetric loss function is the General Entropy (GE) loss, proposed by Calabria and Pucini [3], defined by

$$L(\hat{\delta},\delta) \propto \left(\hat{\delta}/\delta\right)^{\eta} - \eta\log\left(\hat{\delta}/\delta\right) - 1,$$

where η is the shape parameter. This loss is a generalization of the entropy loss used by several authors where the shape parameter $\eta = 1$. The parameter η reflects the departure from symmetry. When $\eta > 0$, overestimation is considered to be more serious than underestimation and vice versa.

It is easy to verify that the value of δ that minimizes $E(L(\delta,\theta)|\mathbf{X})$ under the GE loss is

$$\delta_{BE,\eta} = \left(E(\theta^{-\eta}|\mathbf{X})\right)^{-\frac{1}{\eta}}, \tag{4.4}$$

provided that $E(\theta^{-\eta}|\mathbf{X})$ exists. Note that if we put $\eta = -1$ in $\delta_{BE,\eta}$, it provides the Bayes estimate under the SE loss function.

4.2 Bayes estimation under the LINEX and GE loss functions

From Equations (4.3) and (4.4), the Bayes estimates of θ_1 and θ_2 under the LINEX and GE loss functions with respect to the joint prior density in Equation (4.1) can be shown that

$$\hat{\theta}_{i,BL,\tau} = \tau^{-1}(M_i + a_i)\ln\left(1 + \frac{\tau}{U_i + b_i}\right), \quad i = 1,2, \tag{4.5}$$

and

$$\hat{\theta}_{i,BE,\eta} = \left[\frac{\Gamma(M_i + a_i - \eta)}{\Gamma(M_i + a_i)}\right]^{-\frac{1}{\eta}} (U_i + b_i)^{-1}, \quad i = 1,2, \tag{4.6}$$

respectively. As a special case of (4.6), the Bayes estimates of θ_1 and θ_2 under the SE loss function are

$$\hat{\theta}_{i,BS} = \frac{M_i + a_i}{U_i + b_i}, \quad i = 1,2. \tag{4.7}$$

Note that $\hat{\theta}_{i,BL,\tau}$, $\hat{\theta}_{i,BE,\eta}$ and $\hat{\theta}_{i,BS}$ in Equations (4.5), (4.6) and (4.7), respectively, are the unique Bayes estimates of θ_1 and θ_2 under the LINEX, GE and SE loss functions, respectively, with respect to the proper density (4.1) and hence are admissible; see, for example, Lehmann and Casella [5, p. 323].

Table 1: Air-conditioning systems data.

Plane	Ordered failure times													
7914	3	5	5	13	14	15	22	22	23	30	36	39	44	46
	50	72	79	88	97	102	139	188	197	210				
7913	1	4	11	16	18	18	18	24	31	39	46	51	54	63
	68	77	80	82	97	106	111	141	142	163	191	206	216	

5 Simulation study

To study the performance of of the ML and the Bayes estimates discussed in the preceding sections, we simulated $10,000$ joint progressively first-failure-censored samples from two exponential distributions with the values of parameters $(\theta_1, \theta_2) = (2, 3)$ and different combinations of (k_1, k_2), N, r, (n, m) and $(R_1, \ldots, R_r)$. We also obtained the results for some other choices of (θ_1, θ_2), but as the findings were quite similar, we present here only the results corresponding to $(\theta_1, \theta_2) = (2, 3)$ for the sake of brevity. For these cases, we computed the ML and the Bayes estimates of θ_1 and θ_2 under the SE, LINEX and GE loss functions under the hyperparameters $(a_1, b_1) = (2, 1.1)$ and $(a_2, b_2) = (5, 1.2)$. Then, the estimated risks (ERs) for the estimates of θ_1 and θ_2 are calculated by using the root mean square error. The results are not presented for brevity.

From the simulation results, we concluded that the ERs of the Bayes estimates on the basis of the LINEX and GE loss functions are smaller than those of the ML estimates and the Bayes estimates on the basis of the SE loss function. We also observe that the ERs of all the estimates decrease with increasing r when all other components are kept fixed. Moreover, the ERs of the Bayes estimates based on the LINEX and GE loss functions are sensitive with respect to the shape parameters τ and η, respectively.

6 Illustrative example

Proschan[6] presented data on intervals between failures (in hours) of the air-conditioning system of a fleet of 13 Boeing 720 jet airplanes. After analyzing the data, Proschan[6] observed that the failure distribution of the air-conditioning system for each of the planes was well approximated by exponential distributions. For illustration purposes, we chose here the planes "7913" and "7914" and the corresponding failure time data are presented in Table 1.

By employing a joint progressive first-failure-censoring scheme with $m = 6$, $n = 9$, $k_1 = 4$, $k_2 = 3$, $r = 7$ and $(R_1, \ldots, R_7) = (1, 1, 1, 1, 1, 1, 2)$, the data presented in Table 2 were obtained. Using the data in Table 2 and Equations (3.4), (4.5), (4.6) and (4.7) with $(a_1, b_1) = (1, 1)$ and $(a_2, b_2) = (2, 2)$, we can get the ML and Bayes estimates of the parameters θ_1 and θ_2. These estimates present in Table 3 .

Table 2: The joint progressive first-failure-censoring scheme with $m = 6$, $n = 9$, $k_1 = 4$, $k_2 = 3$, $r = 7$ and $(R_1, \ldots, R_7) = (1, \ldots, 1, 2)$, employed on air-conditioning systems data.

w	1	4	14	15	16	24	30
z	0	0	1	1	0	0	1
s	1	1	1	0	0	0	0
t	0	0	0	1	1	1	2

Table 3: The estimates of $\hat{\theta}_1$ and $\hat{\theta}_2$.

	ML	BS	BL		BE	
			$\tau = -2$	$\tau = 2$	$\eta = -2$	$\eta = 2$
$\hat{\theta}_1$	0.00962	0.01278	0.01282	0.01247	0.01429	0.01345
$\hat{\theta}_2$	0.00833	0.01245	0.01274	0.01242	0.00783	0.00928

References

[1] Balakrishnan, N. and Aggarwala, R. (2000) *Progressive Censoring: Theory, Methods and Applications*, Birkhäuser, Boston.

[2] Balakrishnan, N., Su, F. and Liu, K.Y. (2015) Exact likelihood inference for k exponential populations under joint progressive Type-II censoring, *Communications in Statistics-Simulation and omputation*, **44**, 902-923.

[3] Calabria, R. and Pulcini, G. (1994) An engineering approach to Bayes estimation for the Weibull distribution, *Microelectronics Reliability*, **34**, 789-802.

[4] Doostparast, M., Ahmadi, M.V. and Ahmadi, J. (2013) Bayes estimation based on joint progressive Type II censored data under LINEX loss function, *Communications in Statistics-Simulation and Computation*, **42**, 1865-1886.

[5] Lehmann, E.L. and Casella, G. (1998) *Theory of Point Estimation*, Second edition, Springer, New York.

[6] Proschan, F. (1963) Theoretical explanation of observed decreasing failure rate, *Technometrics*, **15**, 375-383.

[7] Rasouli, A. and Balakrishnan, N. (2010) Exact likelihood inference for two exponential populations under joint progressive Type II censoring, *Communications in Statistics-Theory and Methods*, **39**, 2172-2191.

[8] Varian, H.R. (1975) A Bayesian Approach to Real Estate Assessment, *Studies in Bayesian econometrics and statistics in honor of Leonard J. Savage*, 195-208.

[9] Wu, S.J. and Kuş, C. (2009) On estimation based on progressive first-failure-censored sampling, *Computational Statistics and Data Analysis*, **53**, 3659-3670.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Univariate stochastic ordering for near-records based on different classes of life time distributions

Akbari, M. [1]

Department of Statistics, Mazandaran University

Abstract

The observations between the nth and $(n+1)$th upper record value that fall in the interval $(X^U(n)-a, X^U(n))$, where $X^U(n)$ is upper record value and $a>0$ is a constant, are called the observations near the nth upper record value. In this paper, is obtained some stochastic orderings for number of observations near the nth upper and lower $k-$records based on different classes of life time distributions. Also, It is shown that only in exponential distribution is possible that the number of observations near the nth and mth upper $k-$record or lower $k-$record be equal in distribution.

Keywords: Near-records, Stochastic ordering, Exponential distribution.

1 Introduction

Let $X_1, X_2, ...$ be a sequence of independent and identically distributed (iid) random variables with continuous cumulative distribution function (cdf) F and the probability density function (pdf) f. The statistics $X_{1:n} < X_{2:n} < ... < X_{n:n}$ denote the order statistics from above random sample. So far, a lot of results are provided about of properties and applications of order statistics from kind of different distributions. The reader is referred to Arnold et al. (2008) and David and Nagaraja (2003). Also, many studies of record statistics that are extreme order statistics in partial sequence of a sample have been obtained. X_i is called an upper record value if $X_i > \max\{X_1, ..., X_{i-1}\}$. Let

[1] m.akbari@umz.ac.ir

us recall the sequences of upper $k-$record times, $T^U(n,k)$, and upper $k-$record values, $X^U(n,k)$, which are defined as follows: $T^U(1,k)=k$, $X^U(1,k)=X_{1:k}$ and for $n\geq 2$

$$T^U(n,k)=\min\{j: j>T^U(n-1,k),\ X_j>X_{T^U(n-1,k)-k+1:T^U(n-1,k)}\},$$

and the nth upper $k-$record is defined by

$$X^U(n,k)=X_{T^U(n,k)-k+1:T^U(n,k)},\quad \text{for } n\geq 1.$$

For $k=1$, usual records are recovered. The pdf of $X^U(n,k)$ is given by

$$f_{X^U(n,k)}(x)=\frac{k^n[-\log(\bar{F}(x))]^{n-1}}{(n-1)!}\{\bar{F}(x)\}^{k-1}f(x),\qquad x\in S_X, \tag{1.1}$$

where $\bar{F}(x)=1-F(x)$. Lower $k-$record statistics are defined similarly. We denote the nth lower $k-$record time and lower record value by $T^L(n,k)$ and $X^L(n,k)$, respectively. See Arnold et al. (1998) and Nevzorov (2001) for more details on the theory and applications of record values.
It is well known that plain records are scarce in some states but they have some serious implications. For example, in destructive industrial stress - testing when specimens to be tested are expensive, in climatological or financial phenomena, and in insurance business. So in these cases, the record sequence are quite short and it is not good for statistical method. Gouet et al. (2011) show that near-record are simple to obtain and have longer series than records mostly. Near-record is introduced by Balakrishnan et al. (2005) for first time. They defined the number of observations near the nth upper record value as follows

$$\xi_n(a)=\#\{j:\ \ T_n^U<j<T_{n+1}^U, X_j\in(X^U(n)-a,X^U(n)]\}, \tag{1.2}$$

where $a>0$ is a constant. In fact, near-record observations are not real records, but by considering small values of $a>0$, they have larg values as record values and can be considered as records. So far, several researchers pay attention to properties and applications of $\xi_n(a)$ for example Pakes (2007), Gouet et al. (2007), (2011), (2012a) and (2012b).
In this paper, some stochastic orderings for near-records when parent population is one of life time distribution with increasing, decreasing or constant hazard rate are gotten. The related results are stated in two sections auxiliary results and main results.

1.1 Auxiliary results

Similar (1.2), it can be defined $\xi_n(a,k)$ as the number of observations near the nth upper $k-$record as follows

$$\xi_n(a,k)=\#\{j:\ T^U(n,k)<j<T^U(n+1,k), X_j\in(X^U(n,k)-a,X^U(n,k)]\}, \tag{1.3}$$

where $a>0$. It's pmf is given by (see Balakrishnan et al. (2005))

$$P(\xi_n(a,k)=j)=\int \beta_U(x,a)(1-\beta_U(x,a))^j f_{X^U(n,k)}(x)dx, \tag{1.4}$$

where $\beta_U(x,a) = \frac{\bar{F}(x)}{\bar{F}(x-a)}$. Similarly, the number of observations near the nth lower $k-$record can be expressed as

$$\eta_n(a,k) = \#\{j : T^L(n,k) < j < T^L(n+1,k), X_j \in [X^L(n,k), X^L(n,k)+a)\}, \tag{1.5}$$

and it's pmf is given by

$$P(\eta_n(a,k) = j) = \int \beta_L(x,a)(1-\beta_L(x,a))^j f_{X^L(n,k)}(x)dx, \tag{1.6}$$

where $\beta_L(x,a) = \frac{F(x)}{F(x+a)}$. Using (1.4) and (4), a closed form of survival functions of $\xi_n(a,k)$ and $\eta_n(a,k)$ can be provided easily which are mentioned below, respectively

$$P(\xi_n(a,k) \geq j) = E(1 - \frac{\bar{F}(X^U(n,k))}{\bar{F}(X^U(n,k)-a)})^j, \tag{1.7}$$

and

$$P(\eta_n(a,k) \geq j) = E(1 - \frac{F(X^L(n,k))}{F(X^L(n,k)+a)})^j. \tag{1.8}$$

Equations (1) and (1.8) are used in the next section to get main results.
The stochastic orderings and inequalities are useable in more fields such as probability, reliability theory, economic, insurance and so on. Now, some of agin notions and stochastic orderings that will use to get new results are remined shortly. See Barlow and Proshan (1981) and Shaked and Shantikumar (2007) for more details on aging properties and stochastic ordering, respectively.

Definition 1. *Let X and Y be continuous random variables with cdf's F and G, respectively. Then*

(a) The random variable X is said to be stochastically less than or equal to Y, denoted by $X \leq_{st} Y$, if $\bar{F}(t) \leq \bar{G}(t)$, for all t. Also, its equivalent condition is as follows

$$E(\phi(X)) \leq E(\phi(Y)), \quad \text{for every increasing function } \phi(\cdot). \tag{1.9}$$

(b) A random variable X is said to have increasing [decreasing] failure rate (IFR)[(DFR)] if its failure rate function $r(t) = \frac{f(t)}{\bar{F}(t)}$ is increasing [decreasing] for $t > 0$.

(c) A random variable X is said to have increasing [decreasing] reverse hazard rate (IRHR)[DRHR] if its reverse hazard rate function $\tilde{r}(t) = \frac{f(t)}{F(t)}$ is increasing [decreasing] for $t > 0$.

(d) Let X and Y be two random variables with hazard rate functions r and q, respectively, such that

$$r(t) \geq q(t), \qquad \forall t \in R.$$

Then X is said to be smaller than Y in the hazard rate order, denoted as $X \leq_{hr} Y$. Also, its equivalent condition is as follows

$$\frac{\bar{G}(t)}{\bar{F}(t)} \quad \textit{increasing in } t.$$

(e) Let X and Y be two random variables with reversed hazard rate functions $\tilde{r}$ and $\tilde{q}$, respectively, such that

$$\tilde{r}(t) \leq \tilde{q}(t), \qquad \forall t \in R.$$

Then X is said to be smaller than Y in the reversed hazard rate order, denoted as $X \leq_{rh} Y$.

(f) Let X and Y be continuous [discrete] random variables with densities [discrete densities] f and g, respectively, such that

$$\frac{g(t)}{f(t)} \quad \textit{indreases in } t$$

Then X is said to be smaller than Y in the likelihood ratio order, denoted as $X \leq_{lr} Y$.

The above stochastic orderings are connected to each other as follows:

$$\begin{array}{ccc} X \leq_{lr} Y & \Longrightarrow & X \leq_{hr} Y \\ \Downarrow & & \Downarrow \\ X \leq_{rh} Y & \Longrightarrow & X \leq_{st} Y \end{array}$$

2 Main results

It is well known that there are several parameter models that have successfully served as population models for failure times arising from a wide range of products and failure mechanisms. For example, exponential, weibull, extreme value and so on. We can divide the family of life time distributions into three classes based on increasing, decreasing and constant failure rate functions. In this section, it is obtained some stochastically monotone properties for $\xi_n(a,k)$ and $\eta_n(a,k)$, based on different groups of life time distributions which are stated in the following theorems.

Proposition 2.1. *Let $X^U(n,k)$ and $X^L(n,k)$ be the nth upper and lower $k-$record value from a sequence of iid random variables with common distribution function F. Then, for any $n \leq m$, following statements hold*

(a) $X^U(n,k) \leq_{lr} X^U(m,k)$

(b) $X^L(m,k) \leq_{lr} X^L(n,k)$

Theorem 2.2. *Let $\{X_n,\ n \geq 1\}$ be a sequence of iid random variables with common distribution function F. If F is IFR, then*

$$\xi_n(a) \leq_{st} \xi_m(a), \quad \text{for} \quad n < m.$$

Proof. From Proposition 2.1 (a), we have

$$X^U(n,k) \leq_{st} X^U(m,k) \tag{2.1}$$

Also, it can be shown easily that the function $\frac{\bar{F}(x)}{\bar{F}(x-a)}$ is decreasing in x, when x is IFR. By this fact and equations (1.9) and (2.1), it is concluded that

$$E\left(1 - \frac{\bar{F}(X^U(n,k))}{\bar{F}(X^U(n,k)-a)}\right)^j \leq E\left(1 - \frac{\bar{F}(X^U(m,k))}{\bar{F}(X^U(m,k)-a)}\right)^j, \tag{2.2}$$

for $j \geq 1$. Then from (1), $\xi_n(a,k) \leq_{st} \xi_m(a,k)$, for any $n < m$. It completes the proof. □

Remark 2.3. Assume the assumptions of Theorem 1 hold, and F is DFR. Then

$$\xi_m(a,k) \leq_{st} \xi_n(a,k), \quad \text{for} \quad n < m.$$

Proof. The proof is the same as the proof of Theorem 1. □

Theorem 2.4. *Let $\{X_n,\ n \geq 1\}$ be a sequence of iid random variables with common cdf F that its support is $(0,\infty)$. Then F has constant failure rate c if and only if*

$$P(\xi_n(a,k) = 0) = e^{-ca},$$

for any $a > 0$ and $n \geq 1$.

Proof. It is well known that failure rate function is constant among continuous distribution functions if and only if exponential distribution be the distribution of the population. So, for necessity, it is assumed that X_i's have exponential distribution with parameter c. Then

$$\begin{aligned} P(\xi_n(a,k) = 0) &= \int \frac{e^{-cx}}{e^{-c(x-a)}} dF_{X^U(n,k)}(x). \\ &= e^{-ca} \end{aligned}$$

To proof of the sufficiency, suppose that

$$P(\xi_n(a,k) = 0) = e^{-ca}.$$

From (1.4), it is concluded

$$\begin{aligned} \int_a^\infty \frac{\bar{F}(x)}{\bar{F}(x-a)} dF_{X^U(n,k)}(x) &= e^{-ca} \\ &= e^{-ca} \int_0^\infty dF_{X^U(n,k)}(x) \end{aligned}$$

By simplification of above equality, is concluded

$$\lim_{a\longrightarrow 0}\int_0^{\infty}\{\frac{e^{-u}}{\bar{F}(F^{-1}(1-e^{-u})-a)}-e^{-ca}\}u^{n-1}e^{-ku}du=0.$$

Following result is based on property of completeness of the sequence $\{u^{n-1}\}$:

$$\lim_{a\longrightarrow 0}\{\frac{e^{-u}}{\bar{F}(F^{-1}(1-e^{-u})-a)}-e^{-ca}\}=0 \tag{2.3}$$

Equation (2.3) results in

$$\frac{\bar{F}(t)}{\bar{F}(t-a)}=e^{-ca}. \tag{2.4}$$

The most general solution of (2.4) is the function $\bar{F}(x)=e^{-cx}$. This completes the proof. □

The reader is referred to Higgins (2004) for details about complete sequence function and Aczél (1966) for functional equations.

Theorem 2.5. *Let $\{X_n,\ n\geq 1\}$ be a sequence of iid random variables with common cdf F and constant failure rate c. Then*

$$\xi_n(a,k)\stackrel{d}{=}\xi_m(a,k),\qquad m\neq n.$$

Proof. Suppose that X has exponential distribution with parameter c because the failure rate is assumed constant. Then can be obtained

$$\begin{aligned}
&P\Big(\xi_n(a,k)\geq j\Big)-P\Big(\xi_{n+1}(a,k)\geq j\Big)\\
&=\int_0^{\infty}\left(1-\frac{\bar{F}(x)}{\bar{F}(x-a)}\right)^j\left\{\frac{k^n\{-\log\bar{F}(x)\}^{n-1}}{(n-1)!}\right\}\bar{F}^{k-1}(x)f(x)dx\\
&\quad-\int_0^{\infty}\left(1-\frac{\bar{F}(x)}{\bar{F}(x-a)}\right)^j\left\{\frac{k^m\{-\log\bar{F}(x)\}^{m-1}}{(m-1)!}\right\}\bar{F}^{k-1}(x)f(x)dx\\
&=(1-e^{-ac})^j\{\int_0^{\infty}\frac{k^n}{(n-1)!}y^{n-1}e^{-ky}dy-\int_0^{\infty}\frac{k^m}{(m-1)!}y^{m-1}e^{-ky}dy\\
&=(1-e^{-ac})^j(1-1)=0.
\end{aligned}\tag{2.5}$$

From (2.5), it is concluded that $\xi_n(a,k)\stackrel{d}{=}\xi_m(a,k)$. Then the proof is completed. □

In the next theorem, it is shown the same results as $\xi_n(a,k)$ for $\eta_n(a,k)$.

Theorem 2.6. *Let $\{X_n,\ n\geq 1\}$ be a sequence of iid random variables with continuous cdf F. Then*

(a) $\eta_n(a,k) \leq_{st} \eta_m(a,k)$, *for* $n < m$, *when* F *is DRHR.*

(b) $\eta_m(a,k) \leq_{st} \eta_m(a,k)$, *for* $n < m$, *when* F *is IRHR.*

Proof. The proof is similar to the proof of Theorem 1. □

As mentioned in Balakrishnan et al. (2005), an important use of near-records is in the insurance companies. They defined following quantity that it measures the sume of insurance claims close in value to an unusually large claim $X^U(n,k)$

$$S_n(a,k) = \sum_{i=T_n^U+1}^{T_{n+1}^U-1} X_i I(X_i \in (X^U(n,k)-a, X^U(n,k))). \tag{2.6}$$

Corollary 2.7. *Let* $\{X_n,\ n \geq 1\}$ *be a positive sequence of iid random variables with common cdf* F *that has infinite upper end point. Then*

(a) $\xi_n(a,k) \overset{a.s.}{\sim} \frac{S_n(a,k)}{X^U(n,k)}$

(b) $S_n(a,k) \leq_{st} S_m(a,k)$, *for* $n < m$ *and* F *is IFR*

(c) $S_m(a,k) \leq_{st} S_n(a,k)$, *for* $n < m$ *and* F *is DFR*

Proof. In view of positivity of random variables, is obtained following inequality

$$\xi_n(a,k)(X^U(n,k)-a) \leq S_n(a,k) \leq \xi_n(a,k)X^U(n,k) \tag{2.7}$$

It is well known that $X^U(n,k) \xrightarrow{a.s.} \infty$. This fact and inequality (2.7) complete the proof of part (a). To prove part (b), observe that

$$\begin{aligned} P(S_n(a,k) \leq t) &= P(\frac{S_n(a,k)}{X^U(n,k)} \leq \frac{t}{X^U(n,k)}) \\ &\geq P(\frac{S_m(a,k)}{X^U(m,k)} \leq \frac{t}{X^U(n,k)}) \\ &= P(S_m(a,k) \leq t\frac{X^U(m,k)}{X^U(n,k)}) \\ &\geq P(S_m(a,k) \leq t), \end{aligned} \tag{2.8}$$

where first inequality is concluded from Theorem 1. It completes the proof (b). Similarly, part (c) can be proved. □

Theorem 2.8. *Suppose that* $\{X_n,\ n \geq 1\}$ *and* $\{Y_n,\ n \geq 1\}$ *be two sequences of continuous random variables with cdfs* F *and* G*, respectively. If* $X \leq_{hr} Y$ *and* F *is DFR, then*

$$\xi_n^X(a,k) \leq_{st} \xi_n^Y(a,k),$$

where $\xi_n^X(a,k)$ *denotes the number of observations near the* n*th upper* $k-$ *records when* F *is parent distribution.*

Proof. From Ahmadi and Arghami (2001) is concluded that

$$X^U(n,k) \leq_{st} Y^U(n,k) \tag{2.9}$$

Equation (2.9) and DFR property of F will result in

$$E(1-\frac{\bar{F}(X^U(n,k))}{\bar{F}(X^U(n,k)-a)})^j \geq E(1-\frac{\bar{F}(Y^U(n,k))}{\bar{F}(Y^U(n,k)-a)})^j \tag{2.10}$$

From equivalent condition of hazard rate ordering, the following inequality holds

$$E(1-\frac{\bar{F}(Y^U(n,k))}{\bar{F}(Y^U(n,k)-a)})^j \geq E(1-\frac{\bar{G}(Y^U(n,k))}{\bar{G}(Y^U(n,k)-a)})^j \tag{2.11}$$

Equations (2.10) and (2.11) result in

$$E(1-\frac{\bar{F}(X^U(n,k))}{\bar{F}(X^U(n,k)-a)})^j \geq E(1-\frac{\bar{G}(Y^U(n,k))}{\bar{G}(Y^U(n,k)-a)})^j \tag{2.12}$$

From (2.12), the proof is completed. □

Example 2.9. Assume that X and Y be continuous random variables with pareto(2,1) and pareto(1,1), respectively. Then, one can easily show that $X \leq_{hr} Y$. Also, X has DFR property. Thus based on Theorem 2.8, $\xi_n^X(a,k) \leq_{st} \xi_n^Y(a,k)$.

Theorem 2.10. *Suppose that* $\{X_n,\ n \geq 1\}$ *and* $\{Y_n,\ n \geq 1\}$ *be two sequences of continuous random variables with cdfs* F *and* G*, respectively. If* $X \leq_{rh} Y$ *and* F *is DRHR, then*

$$\eta_n^X(a,k) \leq_{st} \eta_n^Y(a,k),$$

Example 2.11. Assume that X and Y be continuous random variables with following density functions f and G, respectively

$$f(x)=1, \quad 0<x<1 \quad \text{and} \quad g(x)=2x, \quad 0<x<1.$$

Then, one can easily show that $X \leq_{rh} Y$. Also, X has DRHR property. Thus based on Theorem 2.10, $\eta_n^X(a,k) \leq_{st} \eta_n^Y(a,k)$.

Conclusion

Because of complex form of pmfs of $\xi_n(a,k)$ and $\eta_n(a,k)$, it is difficult to obtain some stochastic orderings. But, it is shown usual stochastic ordering in special class of life time distribution. Also, as mentioned in Balakrishnan et al. (2005), an important use of near-records is in the insurance companies. So, the results of this paper can be used in insurance analysis.

References

[1] Aczél, J. (1966). *Lectures on Functional Equations and Their Applications*. Academic Press, New York and London.

[2] Ahmadi, J. and Arghami, N. R. (2001). On the fisher information in record variables. *Metrika*, **53**, 195-206.

[3] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1998). *Records*. John Wiley, New York.

[4] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (2008). *A First Course in Order Statistics*. Unabridged republication of the 1992 original. Classics in Applied Mathematics, No. 54. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pennsylvania.

[5] Balakrishnan, N., Pakes, A. and Stepanov, A. (2005). On the number and sum of near record observations. *Advances in Applied Probability*, **37**, 765–780.

[6] Barlow, R. E. and Proschan, F. (1981), *Statistical Theory of Reliability and Life Testing Probability Models*. To Begin With: Silver Springs, Maryland.

[7] David, H. A. and Nagaraja, H. N. (2003). *Order Statistics*. John Wiley & Sons, New York.

[8] Gouet, R., López, F. J., Sanz, G. (2007), Asymptotic normality for the counting process of weak records and δ-records in discrete models. *Bernoulli*, **13**, 754–781.

[9] Gouet, R., López, F. J., Sanz, G. (2011). Asymptotic normality for the number of records from general distributions. Adv Appl Probab, 43, 2, 422436.

[10] Gouet, R., López, F. J., Sanz, G. (2012a), On δ -record observations: asymptotic rates for the counting process and elements of maximum likelihood estimation. *Test*, **21**, 188–214.

[11] Gouet, R., López, F. J. and Sanz, G. (2012b), Central limit theorem for the number of near-records. *Communications in Statistics: Theory and Methods*, **41**, 309–324.

[12] Higgins, J. R. (2004). *Completeness and Basis Properties of Sets of Special Functions*. Cambridge University press, New York.

[13] Nevzorov, V. B. (2001). *Records: Mathematical Theory (Transl. Math. Monogr.***194***)*. American Mathematical Society, Providence, RI.

[14] Pakes, A. G. (2007). Limit theorems for numbers of near records. *Extremes*, **10**, 207–224.

[15] Shaked, M., Shanthikumar, J. G. (2007), *Stochastic Orders*, Springer.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Selection of the least risky minimal repair system

Amini, M. [1]

Department of Statistics, University of Tehran

Abstract

Bayesian selection rule is proposed for selection of the subset of the least risky systems based on minimal repair times of repairable systems. The Weibull distribution is considered as the lifetime distribution of the systems. The Laplace approximation is used for computation of the Bayesian selection rule. It is observed through a simulation study that selection of the least risky system results to a high level of accuracy and efficiency.

Keywords: Computational Bayes, Qualifying index set, Ranking, Reliability, Time interval between failures.

1 Introduction

Around 50 productive years of statistical ranking, selection and multiple comparison theories from mid-50's has witnessed a wide range of methods for selection of the best population and multiple comparisons of populations based on the random and censored samples. For an overview of methods and the key references, one can refer to Dudewicz, 1980, Balakrishnan et al., 2007 and Gupta and Berger, 2012.

There are yet many complications for utilizing the ranking, selection and multiple comparison methods in the reliability theory of the systems. There are several criteria for comparing the systems, including reliability, the hazard rate, and the mean time to failure. Based on the type of the life experiment or the usage of the system, there might be different types of lifetime data available,

[1] mort.amini@gmail.com

including failure times, repair times, censored samples, degradation measurements, step-stress, and experimental data. The best fitted distribution for the lifetime of the underlying system is much more complicated than the usual exponential or gamma distributions, studied in the literature. These complications urge the need for certain customized computational methods.

In this paper, we consider the problem of Bayesian selection of the least risky system based on the available minimal repair times of the systems, when the weibull distribution is considered as the lifetime distribution of the systems.

2 The model

In the sequel, the Weibull distribution is assumed for the underlying lifetime distribution of the systems. The two-parameter Weibull distribution plays a central role in reliability analysis, mainly due to its flexibility and the ability to capture a wide variety of shapes. If X is a Weibull distributed random variable with scape and shape parameter α and β, respectively, it will be denoted by $X \sim \text{Weib}(\alpha, \beta)$. The associated probability density function (pdf) is given by

$$g(x|\alpha,\beta) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] \ , x>0,\ \alpha>0,\ \beta>0\ . \tag{2.1}$$

The corresponding reliability and hazard rate functions are $R(x|\alpha,\beta) = \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]$ and $h(x|\alpha,\beta) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1}$, respectively, while the mean time to failure is $\mu(\alpha,\beta) = \alpha\Gamma(1+\beta^{-1})$.

The minimal repair, is a repair process such that, after a failure, the system is put back into operation in the same state as immediately before the failure. Consider k independent minimal repair systems S_i, $i = 1,\ldots,k$, and for each of them, let $T_{i(1)}, T_{i(2)}, \ldots,$ denote minimal repair times from some absolutely continuous distribution function $G(\cdot|\boldsymbol{\theta}_i)$ with probability density function $g(\cdot|\boldsymbol{\theta}_i)$ and the reliability function $R(\cdot|\boldsymbol{\theta}_i) = 1 - G(\cdot|\boldsymbol{\theta}_i)$, in which $\boldsymbol{\theta}_i$ is the associated parameter vector of the ith system, $i = 1,\ldots,k$. It is well known that the minimal repair times possess the same joint distribution as upper record values from the underlying G. For details one may refer to Arnold et al. (1998, p. 10) and Kamps (1995, p.31). Thus, the joint density function of the first m_i minimal repair times of the i-th system, $\mathbf{T}_i = (T_{i(1)},\ldots,T_{i(m_i)})$, for some $m_i \in \mathbb{N}$, is given by

$$\begin{aligned} f_{\mathbf{T}_i}(\mathbf{t}_i|\boldsymbol{\theta}_i) &= g(t_{im_i}|\boldsymbol{\theta}_i)\prod_{j=1}^{m_i-1}\frac{g(t_{ij}|\boldsymbol{\theta}_i)}{R(t_{ij}|\boldsymbol{\theta}_i)}, \quad 0 \le t_{i1} \le \ldots \le t_{im_i} \\ &= \frac{\beta_i^{m_i}\left(\prod_{j=1}^{m_i} t_{ij}\right)^{\beta_i-1}}{\alpha_i^{m_i\beta_i}} \exp\{-\left(\frac{t_{im_i}}{\alpha_i}\right)^{\beta_i}\},\ i=1,\ldots,k, \end{aligned} \tag{2.2}$$

where $\mathbf{t}_i = (t_{i1},\ldots,t_{im_1})$ with $0 \le t_{i1} \le \ldots \le t_{im_i}$.

3 Bayesian selection of the least risky system

Let $S_1,\ldots,S_k$ be k reliable systems, such that the lifetime of S_i has pdf $f_i(\cdot;\boldsymbol{\theta}_i)$, mean time to failure $\mu_i(\boldsymbol{\theta}_i)$ and the hazard rate $h(t_0|\boldsymbol{\theta}_i)$ at a given time point t_0, $i = 1,\ldots,k$, where $\boldsymbol{\theta}_i$ is the

parameter vector associated to the lifetime of ith system. Considering the prespecified control value h_0 and define the qualifying set

$$Q_h = \{i \in \{1, \ldots, k\};\ h(t_0|\theta_i) \leq h_0,\ \mu(\theta_i) \geq \mu_0\}.$$

The system S_i is considered as the least risky system at time t_0, if it simultaneously satisfies

(i) $i \in Q_h$, and

(ii) $h(t_0|\theta_i) \leq h(t_0|\theta_j),\ \forall j \in Q_h$.

One may mimic the loss function introduced by Lai et al. (2012) and consider

$$L(\boldsymbol{\delta}', \tilde{\boldsymbol{\theta}}|t_0) = \sum_{j=0}^{k} \delta'_j h'(t_0|\boldsymbol{\theta}_j) - \min_{0\leq i\leq k} h'(t_0|\boldsymbol{\theta}_i),$$

where

$$h'(t_0|\boldsymbol{\theta}_i) = h(t_0|\boldsymbol{\theta}_i) I_{\{\mu(\boldsymbol{\theta}_i)\geq\mu_0\}} + (h_0 + \epsilon) I_{\{\mu(\boldsymbol{\theta}_i)<\mu_0\}},$$

for a small enough positive number ϵ and $h'(t_0|\boldsymbol{\theta}_0) = h_0$. Then, the Bayes risk of the decision $\boldsymbol{\delta}$, for a given prior π, is obtained in a similar way as

$$r(\boldsymbol{\delta}, \pi|t_0) = \int_T \sum_{i=0}^{k} \delta'_i(\tilde{\mathbf{t}}) \phi'_i(\tilde{\mathbf{t}}) m(\tilde{\mathbf{t}})\ \mathrm{d}\tilde{\mathbf{t}} - C',$$

where C' is a constant, $\phi'_0(\tilde{\mathbf{t}}) = h_0 + \epsilon$ and for $i = 1, \ldots, k$,

$$\phi'_i(\tilde{\mathbf{t}}) = (h_0 + \epsilon) Pr(\mu(\boldsymbol{\theta}_i) \leq \mu_0|\tilde{\mathbf{t}}) + \int_{\boldsymbol{\Theta}_i} h(t_0|\boldsymbol{\theta}_i) \pi(\boldsymbol{\theta}_i|\mathbf{t}_i) I(\mu(\boldsymbol{\theta}_i) > \mu_0)\ \mathrm{d}\boldsymbol{\theta}_i. \tag{3.1}$$

Hence, a selection rule for selection of the least risky system with minimum Bayes risk would be

$$\delta'^{B_1}_i(\tilde{\mathbf{t}}) = \begin{cases} 1, & i = \min\{j;\ \phi'_j(\tilde{\mathbf{t}}) \leq \phi'_h(\tilde{\mathbf{t}}),\ \forall\, 0 \leq h \leq k\} \\ 0, & \text{otherwise,} \end{cases} \qquad i = 0, 1, \ldots, k. \tag{3.2}$$

However, if there are systems with equal values of $\phi'_i(\tilde{\mathbf{t}})$, then selection of the first system with maximum value of $\phi'_i(\tilde{\mathbf{t}})$, as in (3.2), is not appropriate. In such a situation, one might be interested in selection of all systems, which their value of $\phi_i(\tilde{\mathbf{t}})$ is equal to $\max_{1\leq j\leq k} \phi_j(\tilde{\mathbf{t}})$, that is

$$\delta^{B_2}_i(\tilde{\mathbf{t}}) = \begin{cases} 1, & \phi'_i(\tilde{\mathbf{t}}) = \max_{0\leq j\leq k} \phi'_j(\tilde{\mathbf{t}}) \\ 0, & \text{otherwise,} \end{cases} \qquad i = 0, 1, \ldots, k. \tag{3.3}$$

For the selection rule (3.3) to be well-defined, the restrictions $\sum_{i=0}^{k} d_i = 1$ in the decision space as well as $\sum_{i=0}^{k} \delta_i(\tilde{\mathbf{t}}) = 1$ for the selection rules must be relaxed, and instead of a selection rule $\boldsymbol{\delta}$ with minimum Bayes risk $r(\boldsymbol{\delta}, \pi)$, we are interested in the selection rule $\boldsymbol{\delta}$, such that

$$r(\boldsymbol{\delta}, \pi) < \frac{\#\{i \in \{1, \ldots, k\};\ \boldsymbol{\delta}_i = 1\}}{\#\{i \in \{1, \ldots, k\};\ \boldsymbol{\delta}'_i = 1\}} r(\boldsymbol{\delta}', \pi). \tag{3.4}$$

The inequality (3.4) guarantees that if there are more than one system with maximum Bayes risk, then the selection rule $\boldsymbol{\delta}$ selects all of them. The selection rule (3.3) satisfies the inequality (3.4) and thus is is the optimal Bayes selection rule based on the criterion (3.4).

Table 1: Simulation results for different values of parameters.

c	ϵ	(h_0,μ_0,t_0)	p_1	p_2	p_3	p_4	p_5	p_6	p_7
0.1	0.001	(20,110,150)	0.92	0.61	0.61	0.32	0.08	**0.32**	**0.32**
		(10,100,200)	0.98	0.91	0.91	0.07	0.02	0.00	0.00
	0.01	(20,110,150)	0.97	0.59	0.59	0.37	0.03	0.27	0.27
		(10,100,200)	0.96	0.85	0.85	0.12	0.04	0.00	0.00
	0.1	(20,110,150)	0.94	0.56	0.56	0.38	0.06	0.26	0.26
		(10,100,200)	0.93	0.80	0.80	0.13	0.07	0.00	0.00
0.5	0.001	(20,110,150)	0.95	0.57	0.57	0.38	0.05	0.25	0.25
		(10,100,200)	0.98	0.88	0.88	0.10	0.02	0.00	0.00
	0.01	(20,110,150)	0.96	0.57	0.57	0.38	0.04	0.26	0.26
		(10,100,200)	0.98	0.89	0.89	0.09	0.02	0.00	0.00
	0.1	(20,110,150)	0.97	0.57	0.57	0.40	0.03	0.27	0.27
		(10,100,200)	0.97	0.89	0.89	0.08	0.03	0.00	0.00
1	0.001	(20,110,150)	0.97	0.57	0.57	0.40	0.03	0.27	0.27
		(10,100,200)	**0.99**	**0.92**	**0.92**	0.07	**0.01**	0.00	0.00
	0.01	(20,110,150)	0.97	0.58	0.58	0.38	0.40	0.23	0.23
		(10,100,200)	0.98	0.90	0.90	0.08	0.02	0.00	0.00
	0.1	(20,110,150)	0.97	0.57	0.57	0.40	0.04	0.25	0.25
		(10,100,200)	**0.99**	**0.92**	**0.92**	0.07	**0.01**	0.00	0.00
2	0.001	(20,110,150)	0.96	0.57	0.57	0.39	0.04	0.27	0.27
		(10,100,200)	0.99	0.91	0.91	0.08	**0.01**	0.00	0.00
	0.01	(20,110,150)	0.96	0.58	0.58	0.38	0.04	0.25	0.25
		(10,100,200)	**0.99**	0.90	0.90	0.09	**0.01**	0.00	0.00
	0.1	(20,110,150)	0.97	0.53	0.53	**0.44**	0.03	0.25	0.25
		(10,100,200)	**0.99**	0.90	0.90	0.08	**0.01**	0.00	0.00

4 Simulation study

In order to examine the performance of the proposed method, we perform a simulation study by considering $k = 10$ repairable systems, with $\mathbf{m} = (m_1, \ldots, m_{10}) = (4, 6, 3, 8, 5, 3, 6, 7, 4, 5)$ observed minimal repair times. We have generated the minimal repair times under Weibull populations, with paramaters $\theta_1 = (110, 15)$, $\theta_2 = (120, 10)$, $\theta_3 = (100, 20)$, $\theta_4 = (105, 12)$, $\theta_5 = (130, 15)$, $\theta_6 = (140, 10)$, $\theta_7 = (120, 20)$, $\theta_8 = (110, 30)$, $\theta_9 = (125, 25)$ and $\theta_{10} = (100, 25)$. Thus, the mean failure times vector of the systems, $(\mu(\boldsymbol{\theta}_1), \ldots, \mu(\boldsymbol{\theta}_{10}))$ is as follows

$$(106.22, 114.16, 97.35, 100.62, 125.54, 133.19, 116.82, 108.00, 122.30, 97.84).$$

The algorithm is repeated for $N = 10^4$ times and the following probabilities are estimated

$$p_1 = P(S \ni i^*),\ p_2 = P(S \subset Q_R),\ p_3 = P(i^* \in S \subset Q_R),\ p_4 = P(i^* \in S \not\subset Q_R),$$

$$p_5 = P(i^* \notin S \not\subset Q_R),\ p_6 = P(i^* \in S = Q_R),\ p_7 = P(S = Q_R),$$

where S is the selected subset.

Here, we have chosen the hyperparameter c to take different values $c = 0.1, 0.5, 1, 2$ and $d = c/(\hat{\beta}_i/\hat{\alpha}_i(\hat{\tau}_i/\hat{\alpha}_i)^{\hat{\beta}_i - 1}$, where $\hat{\boldsymbol{\theta}}_i = (\hat{\alpha}_i, \hat{\beta}_i)$ is the MLE of $\boldsymbol{\theta}_i = (\alpha_i, \beta_i)$ and $\hat{\tau}_i = \mu(\hat{\boldsymbol{\theta}}_i)$. These values are chosen so that the mean of the prior distribution be fix and its variance vary by c.

Table 1 presents the values of the estimated $p_1 - p_7$ for $c = 0.1, 0.5, 1, 2$, $\epsilon = 0.1, 0.01, 0.001$ and $(h_0, \mu_0, t_0) = (20, 110, 150),\ (10, 100, 200)$.

It is observed from Table 1, that $c = 1$ results in the best values of $p_1 - p_3$, while $c = 2$ results in the best values of p_4, p_5, and $c = 0.1$ and $\epsilon = 0.001$ results in best values of p_6 and p_7. Furthermore, $(h_0, \mu_0, t_0) = (10, 110, 200)$ results in better selection rules.

References

[1] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1998). *Records*, John Wiley & Sons, New York.

[2] N. Balakrishnan, N. Kannan and H. N. Nagaraja (2007). *Advances in Ranking and Selection, Multiple Comparisons, and Reliability: Methodology and Applications.* Springer Science & Business Media, USA.

[3] E. J. Dudewicz (1980). Ranking (ordering) and selection: an overview of how to select the best. *Technometrics*, 22 (1), 113-119.

[4] S. S. Gupta and J. O. Berger (2012). *Statistical Decision Theory and Related Topics V*. Springer Science & Business Media, New York, USA.

[5] Kamps, U. (1995). *A Concept of Generalized Order Statistics*. Teubner, Stuttgart.

[6] Y. T. Lai, H. H. Huang and Y. P. Lin (2012). Selecting the Best Exponential Populations in Terms of Reliability: Empirical Bayes Approach. *Tamsui Oxford Journal of Information and Mathematical Sciences*, 28 (3), 327–340.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Stochastic comparisons of series systems with independent heterogeneous components

Asgari, T. [1] and Kelkinnama, M. [2]

[1,2] Department of Mathematical Sciences, Isfahan University of Technology

Abstract

In this paper, we investigate stochastic properties of the smallest order statistics from independent heterogeneous Chen's random variables with different parameters.

Keywords: Series system, Heterogeneous components, Stochastic comparisons.

1 Introduction

Order statistics play an important role in statistical inference, life testings, reliability theory and many other areas. Let $X_1, ..., X_n$ be n random variables and let $X_{i:n}$ denotes their ith order statistic, $i = 1, ..., n$. In reliability theory, the kth order statistic $X_{k:n}$ corresponds to the lifetime of a $(n-k+1)$-out-of-n system. Parallel and series systems are the special ceses of such coherent systems, wherein the lifetime of a series system corresponds to the smallest order statistic $X_{1:n}$ and the lifetime of a parallel system corresponds to the largest order statistic $X_{n:n}$.

Many authors have studied stochastic comparisons of lifetimes of series and parallel systems with s-independent heterogeneous distribution; for example, [5] considered paralle system with heterogeneous exponentiated Weibull distribution, [7] and [4] considered series system with heterogeneous Weibull distribution, [8] considered parallel system with heterogeneous exponenetial

[1]tahere.asgari1371@gmail.com
[2]m.kelkinnama@cc.iut.ac.ir

distribution, [9] and [10] considered parallel system with heterogeneous gamma distribution. The references [3] and [2] investigated stochastic comparisions for series and paralle systems with heterogeneous lower truncated Weibull and exponenetial-weibull distributions, respectively.

Chen [1] proposed a two-parameter distribution with bathtub shape or increasing failure rate function. The probability density function (PDF) and cumulative distribution function (CDF) are respectively of the form,

$$f_X(x;\alpha,\beta) = \alpha\beta x^{\beta-1} exp\{\alpha(1-e^{x^\beta}) + x^\beta\}, x > 0, \alpha > 0, \beta > 0,$$
$$F_X(x;\alpha,\beta) = 1 - exp\{\alpha(1-e^{x^\beta})\}, x > 0.$$

In fact $f_X(x;\alpha,\beta)$ has a bathtub shaped hazard function when $\beta < 1$ and when $\beta \geq 1$, it has an increasing hazard function. The case $\alpha = 1$ corresponds to the exponential power distribution.

In this work, we investigate the stochastic properties of the smallest order statistics from Chen's distributions.

Before proceeding the main results, let us first recall some stochastic orders that will be used in the sequel.

Definition 1. *Let X and Y be two nonnegative random variables having support $(0,+)$. Then:*

(i) X is said to be smaller than Y in the hazard rate order if $h_X(x) \geq h_Y(x)$, or equivalently, if $\bar{G}(x)/\bar{F}(x)$ is increasing in x, and denoted by $X \leq_{hr} Y$;

(ii) X is said to be smaller than Y in the likelihood ratio order if $g(x)/f(x)$ is increasing in x, and denoted by $X \leq_{lr} Y$;

(iii) X is said to be smaller than Y in the usual stochastic order if $\bar{F}(x) \leq \bar{G}(x)$, and denoted by $X \leq_{st} Y$.

Definition 2. *Let $\mathbf{x} = (x_1, ..., x_n)$ and $\mathbf{y} = (y_1, ..., y_n)$ be two real vectors, denote $x_{(1)} \leq ... \leq x_{(n)}$ the increasing arrangement of $x_1, ..., x_n$. $\mathbf{x}$ is said to be majorized by $\mathbf{y}$ (denoted as $\mathbf{x} \preceq^m \mathbf{y}$) if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ and $\sum_{i=1}^j x_{(i)} \geq \sum_{i=1}^j y_{(i)}$ for all $j = 1, ..., n-1$.*

2 Main results

Theorem 2.1. *Let $X_1, X_2, ..., X_n$ be independent random variables with $X_i \sim Chen(\alpha_i, \beta)$ where $\alpha_i > 0$, $i = 1, ..., n$. Let $Y_1, Y_2, ..., Y_n$ be another set of independent random variables with $Y_i \sim Chen(\theta_i, \beta)$ where $\theta_i > 0, i = 1, ..., n$. Then $\sum_{i=1}^n \alpha_i \leq (\geq) \sum_{i=1}^n \theta_i$ implies that $X_{1:n} \geq_{hr} (\leq_{hr}) Y_{1:n}$ for all $\beta > 0$.*

Proof. It is well-known that for the series system, $h_{X_{1:n}}(x) = \sum_{i=1}^n h_i(x)$, where, $h_i(x)$ is the hazard rate function of random variable X_i. Hence, for the Chen's distribution we have

$$h_{X_{1:n}}(x) = \sum_{i=1}^n \alpha_i \beta x^{\beta-1} e^{x^\beta} = \beta x^{\beta-1} e^{x^\beta} \sum_{i=1}^n \alpha_i \tag{2.1}$$

Now

$$\begin{aligned} h_{X_{1:n}}(x) &= \beta x^{\beta-1} e^{x^\beta} \sum_{i=1}^{n} \alpha_i \\ &\leq (\geq) \quad \beta x^{\beta-1} e^{x^\beta} \sum_{i=1}^{n} \theta_i \\ &= h_{Y_{1:n}}(x) \end{aligned}$$

□

Theorem 2.2. *Let $X_1, X_2, ..., X_n$ be independent random variables with $X_i \sim Chen(\alpha_i, \beta)$ where $\alpha_i > 0$, $i = 1, ..., n$. Let $Y_1, Y_2, ..., Y_n$ be another set of independent random variables with $Y_i \sim Chen(\theta_i, \beta)$ where $\theta_i > 0, i = 1, ..., n$. Then $X_{1:n} \leq_{hr} Y_{1:n} \Leftrightarrow X_{1:n} \leq_{lr} Y_{1:n}$.*

Proof. For series system we have $f_{X_{1:n}}(x) = \prod_{i=1}^{n} \bar{F}_i(x) \sum_{i=}^{n} h_i(x)$, therefore,

$$\frac{f_{Y_{1:n}}(x)}{f_{X_{1:n}}(x)} = \frac{\prod_{i=1}^{n} \bar{F}_{Y_i}(x)}{\prod_{i=1}^{n} \bar{F}_{X_i}(x)} \frac{\sum_{i=1}^{n} h_{Y_i}(x)}{\sum_{i=1}^{n} h_{X_i}(x)} = \frac{\bar{F}_{Y_{1:n}}(x)}{\bar{F}_{X_{1:n}}(x)} \frac{\sum_{i=1}^{n} \theta_i}{\sum_{i=1}^{n} \alpha_i}$$

Hence, $\frac{f_{Y_{1:n}}(x)}{f_{X_{1:n}}(x)}$ is increasing in x iff $\frac{\bar{F}_{Y_{1:n}}(x)}{\bar{F}_{X_{1:n}}(x)}$ is increasingin x.

□

It will be interesting to see whether the above result can be changed to the case that under some conditions on the shape parameter, the series system lifetimes can be ordered. The following example show that the popular condition $(\beta_1, ..., \beta_n) \preceq^m (\beta_1^*, ..., \beta_n^*)$ does not conclude $X_{1:n} \leq_{st} Y_{1:n}$.

Example 2.3. Let (X_1, X_2) be a vector of heterogeneous Chen random variables, $Chen(\alpha, \beta_i)$ with $\alpha = 0.2$ and shape parameter vector $(1.5, 2)$. Let (X_1^*, X_2^*) be a vector of heterogeneous Chen random variables, $Chen(\alpha, \beta_i^*)$ with $\alpha = 0.2$ and shape parameter vector $(1, 2.5)$.It can be easily shown that $(1.5, 2) \preceq^m (1, 2.5)$ however $X_{1:2}$ and $X_{1:2}^*$ are not ordered in the stochastic order as can be seen in Figure 1.

At the following theorem, we compare the series systems with heterogeneous and homogeneous components, in likelihood ratio order.

Theorem 2.4. *Consider a series system with heterogeneous components with lifetimes $X_i \sim Chen(\alpha_i, \beta)$ for $i = 1, ..., n$ and a series system with i.i.d. components following the common distribution $Z_i \sim Chen(\alpha, \beta)$ for $i = 1, ..., n$. Then (i) $1/n \sum_{i=1}^{n} \alpha_i \leq \alpha$ implies that $X_{1:n} \geq_{lr} Z_{1:n}$ and (ii) $1/n \sum_{i=1}^{n} \alpha_i \geq \alpha$ implies that $X_{1:n} \leq_{lr} Z_{1:n}$.*

Proof. As denoted above, the first order statistics, $X_{1:n}$, has the density function $f_{X_{1:n}}(x) = \sum_{i=1}^{n} \alpha_i \beta x^{\beta-1} e^{\{(1-e^{x^\beta}) \sum_{i=1}^{n} \alpha_i + x^\beta\}}$, and similarly $Z_{1:n}$ has the density function $g_{Z_{1:n}}(x) =$

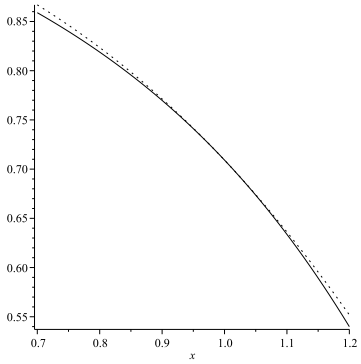


Figure 1: plot of reliability functions of $X_{1:2}, X^*_{1:2}$

$n\alpha\beta x^{\beta-1}e^{n\alpha(1-e^{x^\beta})+x^\beta}$. Then $l_n(x) = \frac{f_{X_{1:n}}(x)}{g_{Z_{1:n}}(x)} = \frac{\sum_{i=1}^n \alpha_i}{n\alpha} e^{(\sum_{i=1}^n \alpha_i - n\alpha)}(1-e^{x^\beta})$ satisfies

$$\frac{\partial l_n(x)}{\partial(x)} = \frac{\sum_{i=1}^n \alpha_i}{n\alpha}(\sum_{i=1}^n \alpha_i - n\alpha)(-e^{x^\beta})\beta x^{\beta-1}e^{(\sum_{i=1}^n \alpha_i - n\alpha)}(1-e^{x^\beta})$$

(i) If $\frac{1}{n}\sum_{i=1}^n \alpha_i \leq \alpha$ then $\frac{\partial l_n(x)}{\partial(x)} \geq 0$ for all $x > 0$ and hence $\frac{f_{X_{1:n}}(x)}{g_{Z_{1:n}}(x)}$ increases in all $x > 0$, that is $X_{1:n} \geq_{lr} Z_{1:n}$.

(ii) If $\frac{1}{n}\sum_{i=1}^n \alpha_i \geq \alpha$ then $\frac{\partial l_n(x)}{\partial(x)} \leq 0$ for all $x > 0$ and hence $\frac{g_{Z_{1:n}}(x)}{f_{X_{1:n}}(x)}$ increases in all $x > 0$, and hence $X_{1:n} \leq_{lr} Z_{1:n}$.

□

At the continue, we compare series systems with multiple-outlier model for Chen's distribution with different parameters.

Theorem 2.5. *Let $(X_1, ..., X_n)$ be a vector of independent Chen random variables such that $X_i \sim Chen(\alpha_1, \beta_1)$ for $i = 1, ..., p$ and $X_j \sim Chen(\alpha_2, \beta_2)$ for $j = p+1, ..., n$. Let $(Y_1, ..., Y_n)$ be another vector of independent Chen random variables such that $Y_i \sim Chen(\theta_1, \beta_1)$ for $i = 1, ..., p$ and $Y_j \sim Chen(\theta_2, \beta_2)$ for $j = p+1, ..., n$. If*

$$min(\alpha_1, \alpha_2) \leq min(\theta_1, \theta_2) \leq max(\alpha_1, \alpha_2) \leq max(\theta_1, \theta_2)$$

then $X_{1:n} \geq_{hr} Y_{1:n}$ for any $\beta_1, \beta_2 > 0$.

Proof. Without loss of generality, let us assume that $\alpha_1 \leq \alpha_2$ and $\theta_1 \leq \theta_2$. We then get $\alpha_1 \leq \theta_1 \leq \alpha_2 \leq \theta_2$. We have to prove that $h_{X_{1:n}}(x) \leq h_{Y_{1:n}}(x)$ for $x \geq 0$, i.e.,

$$p\alpha_1\beta_1 x^{\beta_1-1}e^{x^{\beta_1}} + q\alpha_2\beta_2 x^{\beta_2-1}e^{x^{\beta_2}} \leq p\theta_1\beta_1 x^{\beta_1-1}e^{x^{\beta_1}} + q\theta_2\beta_2 x^{\beta_2-1}e^{x^{\beta_2}},$$

where, $q = n - p$. It can be seen that the desired result holds. □

Theorem 2.6. *Let $(X_1, ..., X_n)$ be a vector of independent Chen random variables such that $X_i \sim Chen(\alpha_1, \beta_1)$ for $i = 1, ..., p$ and $X_j \sim Chen(\alpha_2, \beta_2)$ for $j = p+1, ..., n$. Let $(Y_1, ..., Y_n)$ be another vector of independent Chen random variables such that $Y_i \sim Chen(\theta_1, \beta_1)$ for $i = 1, ..., p$ and $Y_j \sim Chen(\theta_2, \beta_2)$ for $j = p+1, ..., n$. If*

$$min(\alpha_1, \alpha_2) \leq min(\theta_1, \theta_2) \leq max(\alpha_1, \alpha_2) \leq max(\theta_1, \theta_2)$$

and $\frac{min(\alpha_1,\alpha_2)}{max(\alpha_1,\alpha_2)} \geq \frac{min(\theta_1,\theta_2)}{max(\theta_1,\theta_2)}$ then $X_{1:n} \geq_{lr} Y_{1:n}$ for any $\beta_1, \beta_2 > 0$.

Proof. Without loss of generality, let us assume that $\alpha_1 \leq \alpha_2$, $\theta_1 \leq \theta_2$ and $\beta_1 \leq \beta_2$. We then get $\alpha_1 \leq \theta_1 \leq \alpha_2 \leq \theta_2$ and $\frac{\alpha_1}{\alpha_2} \geq \frac{\theta_1}{\theta_2}$. From Theorem 2.5, we know that $X_{1:n} \geq_{hr} Y_{1:n}$ for any $\alpha_1, \alpha_2 > 0$. From Theorem 1.c.4(a) of Shaked and Shanthikumar [8], it is enough to prove that the function

$$\phi(x) = \frac{h_{X_{1:n}}(x)}{h_{Y_{1:n}}(x)} = \frac{p\alpha_1\beta_1 x^{\beta_1-1}e^{x^{\beta_1}} + q\alpha_2\beta_2 x^{\beta 2-1}e^{x^{\beta_2}}}{p\theta_1\beta_1 x^{\beta_1-1}e^{x^{\beta_1}} + q\theta_2\beta_2 x^{\beta 2-1}e^{x^{\beta_2}}}$$

is increasing in $x \geq 0$. After some mathematical computations, it can be shown that under the assumptions, $\phi(x)$ is increasing in $x > 0$ and then the theorem is proven. □

Theorem 2.7. *Let $(X_1, ..., X_n)$ be a vector of independent Chen random variables such that $X_i \sim Chen(\alpha, \beta_1)$ for $i = 1, ..., p$ and $X_j \sim Chen(\gamma, \beta_2)$ for $j = p+1, ..., n$. Let $(Y_1, ..., Y_n)$ be another vector of independent Chen random variables such that $Y_i \sim Chen(\omega, \beta_1)$ for $i = 1, ..., p$ and $Y_j \sim Chen(\gamma, \beta_2)$ for $j = p+1, ..., n$. If $\alpha \leq \omega$ and $\beta_1 \leq \beta_2$ then $X_{1:n} \geq_{lr} Y_{1:n}$ for any $\beta_1, \beta_2 > 0$.*

Proof. Note that $X_j =^{st} Y_j$ for $j = p+1, ..., n$. It can be shown that for the Chen's distribution if $\alpha \leq \omega$ then $X_i \geq_{hr} Y_i$ for $i = 1, ..., p$. Then, we have that $X_{1:n} \geq_{hr} Y_{1:n}$ since the hazard rate order is closed under the operation of ordering. From Theorem 1.c.4(a) of Shaked and Shanthikumar [8], it is enough to prove that the ratio of their hazard rate functions is increasing, i.e.

$$\phi(x) = \frac{h_{X_{1:n}}(x)}{h_{Y_{1:n}}(x)} = \frac{p\alpha\beta_1 x^{\beta_1-1}e^{x^{\beta_1}} + q\gamma\beta_2 x^{\beta_2-1}e^{x^{\beta_2}}}{p\omega\beta_1 x^{\beta_1-1}e^{x^{\beta_1}} + q\gamma\beta_2 x^{\beta_2-1}e^{x^{\beta_2}}}$$

After some mathemartical computations, it can be shown that $\frac{\partial\phi(x)}{\partial x} = (\alpha - \omega)((\beta_1 + \beta_1 x^{\beta_1}) - (\beta_2 + \beta_2 x^{\beta_2}))$, which in nonnegative under the assumptions of the theorem, hence, $\phi(x)$ is increasing in $x > 0$ and then theorem is proven.

□

References

[1] Chen, Z., (2000). A new two parameter lifetime distribution with bathtub shaped or increasing failure rate function. *Statististics and Probability Letters*, **49**, 155161.

[2] Fang L. and Balakrishnan N.(2016), Ordering results for the smallest and largest order statistics from independent heterogeneous exponentialWeibull random variables. *Statistics*, **50**, 1195-1205.

[3] Fang L. and Balakrishnan N.(2016), Stochastic comparisons of series and parallel systems with independent heterogeneous lower truncated Weibull components. *Communications in Statistics - Theory and Method.*, **45**, 540-551.

[4] Fang L. and Zhang X. (2013), Stochastic comparison of series systems with heterogeneous Weibull components. *Statistics and Probability Letter.*, **97**, 25-31.

[5] Fang L. and Zhang X. (2013) Stochastic comparisons of parallel systems with exponentiated Weibull components. *Statist. Probab. Lett.*, **83**, 16491653.

[6] Shaked M., Shanthikumar J.G. (2007), Stochastic Orders, *Springer: New York.*

[7] Torrado N. (2015), Comparisons of smallest order statistics from Weibull distributions with different scale and shape parameters. *Journal of the Korean Statistical Society*, **44**, 68-76.

[8] Zhao P. and Balakrishnan N. (2011), Some characterization results for parallel systems with two heterogeneous exponential components. *Statistics.*, **65**, 593604.

[9] Zhao P. and Balakrishnan N.(2013), New results on comparisons of parallel systems with heterogeneous gamma components. *Statist. Probab.Lett.*, **81**, 3644.

[10] Zhao P. and Balakrishnan N. (2013), Hazard rate comparison of parallel systems with heterogeneous gamma components. *J.Mult. Anal.*, **113**, 153160.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

A shock-based preventive maintenance model for three-state systems

Ashrafi, S. [1]

Department of Statistics, University of Isfahan

Abstract

In this paper, we consider a three-state system with states up, partial performance and down. We assume that this system is subject to shocks. As a result of each shock more than one component may fail and the state of the system may change. We assume that the preventive maintenance is performed at the occurrence time of the shocks, and investigate an optimal shock number in which the mean cost per unit of time is minimized.

Keywords: Two dimensional tie-signature, Optimal shock, Emergency repair.

1 Introduction

The preventive maintenance (PM) is one of the noteworthy areas in the reliability that is first studied by Barlow and Hunter [1] in 1960. After this work, a large number of papers and books have been written in this area. One of the important problems that have been considered in PM is the optimization problem. In many models that are presented until now, it is usually assumed that the PMs have been done periodically at the specified or random time instants $T_1, T_2, \ldots$. In each period of time $(T_{i-1}, T_i]$, if the system fails the emergency repair (ER) is performed and otherwise a PM is performed at T_i. Therefore, in each interval $(T_{i-1}, T_i]$, some costs of repairing the system and rewards for its operating are considered. In such a situation one of the interesting problem is to find the best $T_1, T_2, \ldots$ that minimize the mean costs per unit of time.

[1] m.asadi@sci.ui.ac.ir

Many papers in PM area have been considered a component or a simple system consists of some components. Recently, Finkelsteina and Gertsbakh [3] studied a PM model for any binary complex system. They assumed that the system is subject to shocks in which as a result of each shock exactly one component fails. Then, they investigated an optimal shock number that after it, the PM should be performed. Finkelsteina and Gertsbakh [4] generalized this model to the case where at each shock one component fails with probability p. Zarezaded and Ashrafi [6], investigated the problem of finding the optimal shock number for the case that at each shock more than one component may fail. Finkelsteina and Gertsbakh [5], studied some PM models for the multi-state systems having binary components under the assumption that at each shock exactly one component fails. In this paper, we present a PM model for the three-state systems in the case where at each shock more than one component may fail.

Consider a three-state system with binary components having three states, $K = 0, 1, 2$ in which $K = 2$ denotes the up state, $K = 1$ denotes the partial performance state and $K = 0$ denotes the down state. Let the random variable T_1 denote the time that system enters into state $K = 1$ and T the time that system enters into state $K = 0$. Assume that the system is subject to shocks that appear according to a counting process at random time instants $\varrho_1, \varrho_2, \ldots$. At the occurrence time of each shock, more than one component may fail, and consequently, the system states change as one of the following cases:

(a) The state of system is changed from $K = 2$ to $K = 1$ at one shock and from $K = 1$ to $K = 0$ at another shock, that is $T_1 < T$.

(b) The system state is changed from $K = 2$ to $K = 0$ at one shock. In such a situation, transitions from $K = 2$ to $K = 1$ and from $K = 1$ to $K = 0$ happen simultaneously, that is $T_1 = T$.

Suppose that transitions from $K = 2$ to $K = 1$ and from $K = 1$ to $K = 0$ occur at two different shocks i.e. $T_1 < T$. Let the r.v. M_1 (M_2) denote the least number of the components whose failures change the system state from $K = 2$ to $K = 1$ ($K = 1$ to $K = 0$). Now, assume that the system state changes from $K = 2$ to $K = 0$ at one shock. Let M_3 denote the least number of components whose failures change the state of system from $K = 2$ to $K = 0$. Ashrafi and Zarezadeh [1], introduced a new variant of the notion of two dimensional signature, called two dimensional tie-signature (t-signature), to introduce the model in which it is possible to fail more than one component at each time instant as

$$s^{\tau}_{i,j} = \frac{n_{i,j}}{n^*}, \quad 1 \leq i < j \leq n,$$

and

$$s^{\tau}_{i,i} = \frac{n_i}{n^*}, \quad i = 1, \ldots, n,$$

where $n_{i,j}$ is the number of cases which the components fail such that $M_1 = i$ and $M_2 = j$ and n_i is the number of cases which the components fail such that $M_3 = i$. Also, n^* denotes the number

of all ways that the components of system may fail. The value of n^* has been computed in Lemma 1 of [6] as

$$n^* = \sum_{l=1}^{n} \sum_{i=0}^{l} \binom{l}{i} (-1)^i (l-i)^n.$$

The notion of two dimensional t-signature is a combinatorial property of the system. It does not depend on the mechanism that the components fail and only depends to the system structure.

Consider a three-state system consists of n binary components that is subject to shocks. Let the shocks appear according to the counting process $\{\zeta(t), t > 0\}$ at random times $\varrho_1, \varrho_2, \ldots$. Assume that the components fail independently with probability p and the number of components which have failed at the ith shock, $V_i, i = 1, 2, \ldots$ depends on $V_1, \ldots V_{i-1}$ via $\sum_{j=1}^{i-1} V_j$. Consider $q = 1 - p$ and $\bar{S}^{\tau}_{i,j} = \sum_{k=i+1}^{n} \sum_{l=max\{i,j+1\}}^{n} s^{\tau}_{k,l}$. Ashrafi and Zarezadeh [1] obtained the joint reliability function of T_1 and T as

$$\begin{aligned} P(T_1 > t_1, T > t) &= \sum_{k=0}^{\infty} \sum_{l=k}^{\infty} \eta^*_{k,l} P(\zeta(t_1) = k, \zeta(t) = l) \\ &= \sum_{k=1}^{\infty} \sum_{l=k}^{\infty} c^*_{k,l} P(\varrho_k > t_1, \varrho_l > t), \end{aligned}$$

where for $k = 0, 1, \ldots$, $l = k, k+1, \ldots$, we have

$$\eta^*_{k,l} = \sum_{r_1=0}^{n-1} \sum_{r_2=r_1}^{n-1} \bar{S}^{\tau}_{r_1,r_2} \frac{n!}{r_1!(r_2-r_1)!(n-r_2)!} (1-q^k)^{r_1} (q^k - q^l)^{r_2-r_1} q^{(n-r_2)l}$$

and

$$\begin{aligned} c^*_{k,l} &= \eta^*_{k-1,l-1} - \eta^*_{k-1,l} - \eta^*_{k,l-1} + \eta^*_{k,l}, \quad k < l; \\ &= \eta^*_{k-1,k-1} - \eta^*_{k-1,k}, \qquad\qquad\qquad\quad k = l. \end{aligned} \tag{1.1}$$

In this paper, we investigate a PM model for the three-state system under conditions that are described. We perform the optimal preventive maintenance after the mth shock or the emergency repair after the failure of the system, whichever happens first. Due to the costs of the PM or ER of the system, we are going to find the optimal shock number in which the PM is performed.

2 Shock-based Preventive Maintenance

Consider a three-state system that is subject to shocks that occur according to a counting process. Assume that as a result of each shock each component fail with probability p independently from the other components and hence at each shock some components may fail. Let R_2 (R_1) denote the mean reward per unit of time when the system is operating at state $K = 2$ ($K = 1$). Also, assume that c_0 denotes the cost of replacement a failed component with a new one. In the shock-based

PM, it is assumed that either after the occurrence of one of the shocks PM is performed or at the failure time of the system an emergency repair is done, whichever happens first. In PM (ER), all failed components are replaced with new ones and c_{PM} (c_{ER}) denotes the incurred cost by PM (ER). It is clear that $c_{PM} > c_{ER}$. Hence, we define the renewal cycle as one the following cases

(1) The system is failed before or at the mth shock and at the failure time of the system the state of the system is changed from $K = 2$ into $K = 0$ that is $T_1 = T$. In such a situation, the expected cost is obtained as

$$C_1(m) = \sum_{i=1}^{m} P(T = \varrho_i, T_1 = T)(E(\sum_{k=1}^{i} V_k)c_0 - iR_2 + c_{ER}).$$

Also, the mean length of cycle is

$$L_1(m) = \sum_{i=1}^{m} iP(T = \varrho_i, T_1 = T).$$

(2) The system is failed before or at the mth shock and at the failure time of the system the state of the system is changed from $K = 1$ into $K = 0$, that is $T_1 < T$. In such a situation, the expected cost is obtained as

$$C_2(m) = \sum_{i=1}^{m} \sum_{j=i+1}^{m} P(T_1 = \varrho_i, T = \varrho_j, T_1 < T)(E(\sum_{k=1}^{j} V_k)c_0 - iR_2 - (j-i)R_1 + c_{ER}).$$

Also, the mean length of cycle is

$$L_2(m) = \sum_{i=1}^{m} \sum_{j=i+1}^{m} jP(T_1 = \varrho_i, T = \varrho_j, T_1 < T).$$

(3) At the mth shock the system is in state $K = 1$. In such a situation, the expected cost is obtained as

$$C_3(m) = \sum_{i=1}^{m} \sum_{j=m+1}^{\infty} P(T_1 = \varrho_i, T = \varrho_j, T_1 < T)(E(\sum_{k=1}^{m} V_k)c_0 - iR_2 - (m-i)R_1 + c_{PM}).$$

Also, the mean length of cycle is

$$L_3(m) = m\sum_{i=1}^{m} \sum_{j=i+1}^{\infty} P(T_1 = \varrho_i, T = \varrho_j, T_1 < T).$$

(4) At the mth shock, the system is in state $K = 2$. In such a situation, the expected cost is obtained as

$$C_4(m) = P(T_1 > \varrho_m)(E(\sum_{k=1}^{m} V_k)c_0 - mR_2 + c_{PM}).$$

Also, the mean length of cycle is

$$L_4(m) = mP(T_1 > \varrho_m).$$

It can be seen that $P(T = \varrho_i, T_1 = T) = c^*_{i,i}$ and $P(T_1 = \varrho_i, T = \varrho_j, T_1 < T) = c^*_{i,j}$, where $c^*_{i,j}$ is defined in (1.1); see [1]. Hence,

$$\begin{aligned} P(T_1 > \varrho_m) &= P(T_1 > \varrho_m, T_1 < T) + P(T_1 > \varrho_m, T_1 = T) \\ &= \sum_{i=m+1}^{\infty} \sum_{j=i+1}^{\infty} c^*_{i,j} + \sum_{i=m+1}^{\infty} c^*_{i,i} \\ &= \sum_{i=m+1}^{\infty} \sum_{j=i}^{\infty} c^*_{i,j}. \end{aligned}$$

Zarezadeh et al. showed that $\sum_{k=1}^{i} V_k$ has binomial distribution with parameters n and $1 - q^i$. Hence, $E(\sum_{k=1}^{i} V_k) = n(1-q^i)$. From the above argument, it can be seen that mean cost per cycle is

$$\begin{aligned} C_{sh}(m) &= \sum_{i=1}^{5} C_i(m) \\ &= \sum_{i=1}^{m} \sum_{j=i}^{m} c^*_{i,j}(n(1-q^j)c_0 - iR_2 - (j-i)R_1 + c_{ER}) \\ &+ \sum_{i=1}^{\infty} \sum_{j=\max\{i,m+1\}}^{\infty} c^*_{i,j}(n(1-q^m)c_0 + c_{PM}) \\ &- \sum_{i=1}^{m} \sum_{j=m+1}^{\infty} c^*_{i,j}(iR_2 + (m-i)R_1) - mR_2 \sum_{i=m+1}^{\infty} \sum_{j=i}^{\infty} c^*_{i,j}. \end{aligned}$$

The length of each cycle is

$$L(m) = \sum_{i=1}^{5} L_i = \sum_{i=1}^{m} \sum_{j=1}^{m} j c^*_{i,j} + m \sum_{i=1}^{\infty} \sum_{j=\max\{i,m+1\}}^{\infty} c^*_{i,j}.$$

Therefore, the average cost per unit of time is obtained as

$$C(m) = \frac{C_{sh}(m)}{L(m)}, \qquad m = 1, 2, \ldots$$

Thus the optimal shock is the shock that minimizes $C(m)$. Now, we present an example.

Example 2.1. Suppose a network has 4 nodes, 4 links and terminals set $\{a, b, d\}$ with the graph presented in Figure 1. We define three states for the network. It is in up state if three terminals

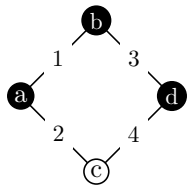


Figure 1: Network with 4 nodes and 4 links.

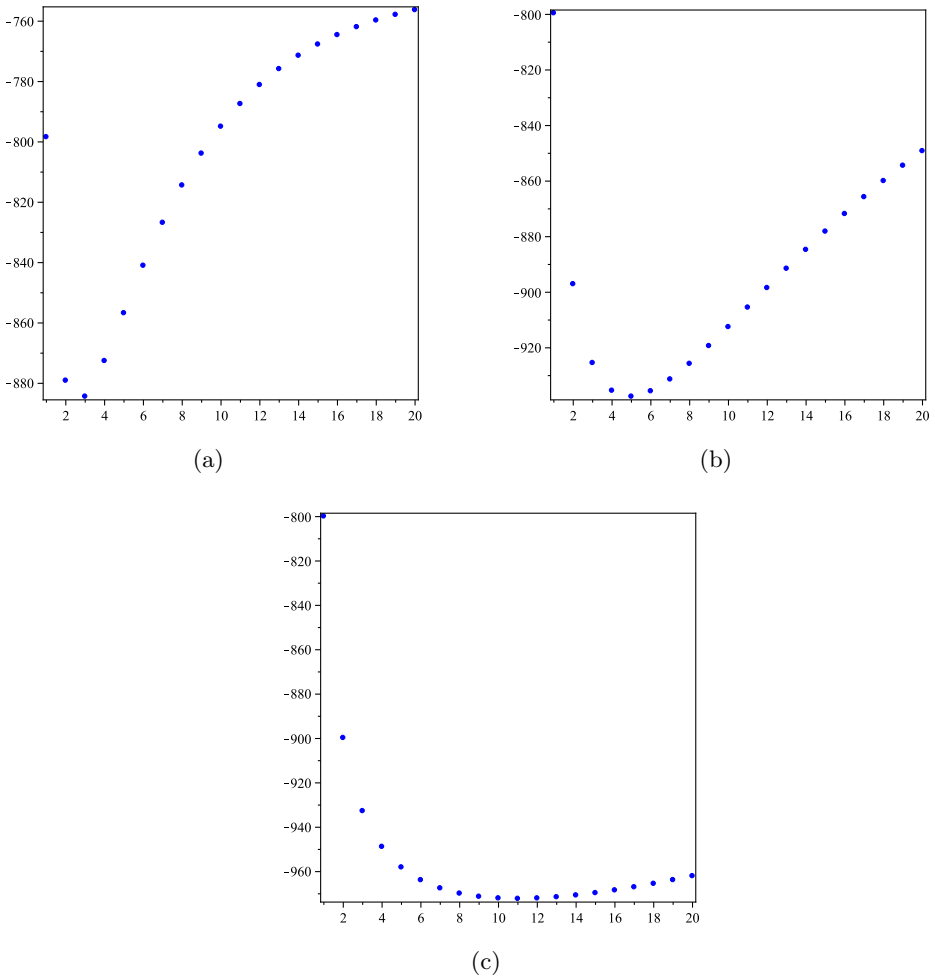


Figure 2: the plot of $C(m)$ when (a) $p = 0.15$, (b) $p = 0.05$, (c) $p = 0.01$

are joined, partial performance if just two terminals are joined, and down state if all terminals are disjoined.

Assume that the nodes are absolutely reliable and only the links may fail. The two dimensinal

t-signatures of this network are obtained as

$$s^{\tau}_{2,3} = \frac{33}{75},\ s^{\tau}_{2,4} = \frac{20}{75},\ s^{\tau}_{3,3} = \frac{13}{75},\ s^{\tau}_{3,4} = \frac{6}{75},\ s^{\tau}_{4,4} = \frac{3}{75}.$$

In this example, we investigate the optimal shock that minimizes the mean cost per unit of time,$C(m)$, when

$$R_1 = 500, \quad R_2 = 1000, \quad c_0 = 2, \quad c_{PM} = 200, \quad c_{ER} = 250.$$

We consider three values for p, $p = 0.15$, $p = 0.05$, $p = 0.01$. The Plots 1, 2 and 3 presents the plots of $C(m)$ for the cases $p = 0.15$, $p = 0.05$ and $p = 0.01$, respectively. It can be seen that the optimal numbers of shock are $m = 3$, $m = 5$ and $m = 11$ when $p = 0.15$, $p = 0.05$ and $p = 0.01$, respectively. Then, when p increases the optimal shock number decreases i.e. when the probability of failure of the components in each shocks increases the PM should be performed earlier.

References

[1] Ashrafi, S., and Zarezadeh, S. (2016). A shock-based model for the reliability of three-state networks, *submitted*.

[2] Barlow R.E., Hunter L.C. (1960). Optimum preventive maintenance policies, *Operations Research* , **8(1)**, 90-100.

[3] Finkelstein, M., and Gertsbakh, I. (2015). Time-free preventive maintenance of systems with structures described by signatures, *Applied Stochastic Models in Business and Industry*, **31(6)**, 836-845.

[4] Finkelstein, M., and Gertsbakh I. (2016). On preventive maintenance of systems subject to shocks, *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, **230(2)**, 220-227.

[5] Finkelstein, M., and Gertsbakh, I. (2017). Preventive maintenance of multistate systems subject to shocks. *Applied Stochastic Models in Business and Industry.*

[6] Zarezadeh, S., and Ashrafi, S. (2016). Preventive maintenance of a network subject to external shocks. *Reliability Theory and its Applications*, 178-181.

[7] Zarezadeh, S., Ashrafi, S. and Asadi, M. (2016). A shock model based approach to network reliability. *IEEE Transactions on Reliability*, **65(2)**, 992-1000.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Bayesian Laplace regression for C-inflated survival data

Aslani Akhore Oleiayi, N. [1] and Saberi, Z. [2]

[1,2] Department of Mathematical Sciences, Isfahan University of Technology

Abstract

In this paper, we propose the use of Bayesian quantile regression for the analysis of survival data with mixed discrete-continuous distribution with a point mass at C, where these observations of the response variable may be censored at C, where these observations may be right censored or true C. The censoring mechanism is assumed random and possibly dependent on the covariates. Quantile regression allow us to permit covariates to affect survival at different stages in the follow up period, so providing a comprehensive study of the survival distribution. We take a Baysian quantile regression approach for continuous part by considering the quantile process as a linear combination of covariates. Also the probability of being censored given that the observed value is equal to C, will be analyzed. We build up a Markov Chain Monte Carlo method from related models in the literature to obtain samples from the posterior distribution. We demonstrate the suitability of the model to analyze this censoring probability with a simulated study.

Keywords: Asymmetric Laplace distribution, Bayesian quantile regression, Right censoring, Survival analysis, Two-part model.

1 Introduction

Survival data analysis typically relies on a parametric assumption about the relationship between the covariates and the survival distribution, for example, the proportional hazards, proportional

[1] n.aslani@math.iut.ac.ir
[2] z_saberi@cc.iut.ac.ir

odds, or accelerated failure time models. In this paper, we consider a regression model where the error term is assumed to follow a type of asymmetric Laplace distribution. The linear quantile regression model assumes that each quantile of the survival (or log survival) distribution is a linear combination of the covariates. Koenker and Bassett (1978) developed quantile regression, where include the following advantages: it makes no distributional assumptions about the regression error term; its inference is invariant to monotone transformations of the outcome variable; in the presence of outlying values, it can be more efficient than ordinary least squares; and most importantly, it permits thorough inference on the entire shape of the conditional distribution and not just the mean. In this study, we consider inference on quantiles of a response variable conditional on a set of covariates when the observations on the life time as a response variable may be censored. The censoring mechanism is assumed random and possibly dependent on the covariates. Peng and Huang (2008) developed a new quantile regression method for survival data subject to conditionally independent censoring. Wang and Wang (2009) proposed to overcome the global-linearity assumption by directly estimating the conditional censoring distribution non-parametrically using the local Kaplan-Meier method. In this paper, for a Bayesian setting, the asymmetric Laplace distribution can be useful in obtaining posterior conditional quantile estimates. Yu and Moyeed (2001) proposed the use of this distribution in order to introduce a Bayesian quantile regression model. In this paper we propose a Bayesian parametric quantile regression model for right censored survival data. In Section 2, we introduce the Laplace method, in Section 3, we show the Bayesian model using quantile regression for censored survival data. In Section 4, we show the Bayesian two-part model using quantile regression for the continuous part, with its prior and posterior setting, we extend the two-part model to allow that C observations are either censored or are true C. In Section 5, we present the suitability of the method using a simulated study.

2 Laplace regression

The quantile regression model proposed by Koenker and Bassett (1978) was based in the following linear model:

$$y_i = \mathbf{x}_i'\boldsymbol{\beta}(\tau) + \varepsilon_i, \text{ i=1,...,n}$$

where ε_i's are the error terms assumed to have the τth quantile equal to zero. Then, the quantile regression estimator is obtained by minimizing

$$\sum_{i=1}^{n} \rho_\tau(y_i - \mathbf{x}_i'\boldsymbol{\beta}(\tau)), \tag{2.1}$$

where the check function $\rho_\tau(.)$ is defined as $\rho_\tau(u) = u(\tau - I(u < 0))$. It is important to note that, initially, there was no distribution assumption made for the response variable.

Santos and Balfarine(2015) proposed the use of the asymmetric laplace distribution for reponse variable in quantile regression. We used the equivariance property of the quantile regression to model a transformed response to meet the limited support of survival data to the real line support of the asymmetric laplace distribution. The probability density function of asymmetric laplace

distribution with location prameter $\mu \in \mathbb{R}$, scale prameter $\sigma > 0$, skewness parameter $\tau \in [0,1]$, can be written as,

$$f(y \mid \mu, \sigma, \tau) = \frac{\tau(1-\tau)}{\sigma} \exp\{-\rho_\tau(\frac{y-\mu}{\sigma})\}, y \in \mathbb{R}. \tag{2.2}$$

In quantile regression, the τth conditional quantile of y_i given $\mathbf{x}_i$ is then modeled by,

$$q_\tau(y_i \mid \mathbf{x}_i) = \mathbf{x}_i'\boldsymbol{\beta}(\tau), \tag{2.3}$$

now, for limited response variable, let $h(.)$ be a nondecreasing function on $\mathbb{R}$, then for any limited random variable Y such a survival data, we have,

$$q_\tau(h(y_i) \mid \mathbf{x}_i) = h(q_\tau(y_i \mid \mathbf{x}_i)), \tag{2.4}$$

where $q_\tau(y \mid \mathbf{x})$ represents the τth conditional quantile of Y given X. Then, we could assume that the τth quantile of transformed response variable follows the linear model, such as,

$$q_\tau(h(y_i) \mid \mathbf{x}_i) = \mathbf{x}_i'\boldsymbol{\beta}(\tau). \tag{2.5}$$

Now, let T_i, $i = 1, ..., n$, be independent response variables of interest and $\mathbf{x}_i$'s k-dimensional vectors of observed covariates. Suppose that T_i may be censored and instead of T_i, we observe $Y_i = \min\{T_i, C_i\}$ where the censoring variable C_i may depend on $\mathbf{x}_i$ but, conditionally on $\mathbf{x}_i$, is independent of T_i. We assume that C_i has no information about the parameters of interest. We define $\delta_i = I(T_i \leq C_i)$, where $I(A)$ is the indicator function of the set A.

Suppose that there exists a fixed k-dimensional parameter vector $\boldsymbol{\beta}(\tau)$ such that

$$T_i = \mathbf{x}_i'\boldsymbol{\beta}(\tau) + \varepsilon_i$$

where $\tau \in (0,1)$ is a fixed and given probability and ε_i an independent and identically distributed residual whose τth quantile equals zero, i.e. $P(\varepsilon_i \leq 0 \mid \mathbf{x}_i) = \tau$. Model (2.1) is equivalent to assuming that $\mathbf{x}_i'\boldsymbol{\beta}(\tau)$ is τth quantile of the conditional distribution of T_i given $\mathbf{x}_i$, i.e. $P(T_i \leq \mathbf{x}_i'\boldsymbol{\beta}(\tau) \mid \mathbf{x}_i) = \tau$.

A desirable property of the conditional quantile $\mathbf{x}_i'\boldsymbol{\beta}(\tau)$ is that it is equivariant to non-decreasing transforms h of the variable T_i simply because $P(T_i \leq \mathbf{x}_i'\boldsymbol{\beta} \mid \mathbf{x}_i) = P(h(T_i) \leq h(\mathbf{x}_i'\boldsymbol{\beta}) \mid \mathbf{x}_i)$. Accelerated failure time (AFT) models, for example, typically model the logarithm of time, i.e. $h(T_i) = log(T_i)$. Now, in the next section, we introduce Bayesian quantile regression for transformed survival data.

3 Bayesian Laplace regression

Yu and Moyeed (2001) proved that the posterior distribution obtained is proper and suggested a MetropolisHasting algorithm to sample from the posterior distribution of the regression parameters. Kozumi and Kobayashi (2011) considered a location-scale mixture representation of the asymmetric Laplace distribution to introduce a Gibbs sampling method for Bayesian quantile regression. We can say that if $h(Y) = log(Y)$ is distributed according to an asymmetric Laplace distribution with parameters $\{\mu, \sigma, \tau\}$, then we have,

$$h(Y)|\nu \sim N(\mu + \theta\nu, \psi^2\sigma\nu),$$

$$\nu \sim Exp(\sigma),$$

where $\theta = \frac{1-2\tau}{\tau(1-\tau)}$, $\psi^2 = \frac{2}{\tau(1-\tau)}$. Then, we assume that the τth quantile of transformed response variable follows the linear model:

$$q_\tau(h(y_i) \mid \mathbf{x}_i) = \mathbf{x}_i'\boldsymbol{\beta}(\tau),$$

so the likelihood function can be written as

$$L(\boldsymbol{\beta}(\tau), \sigma, \boldsymbol{\nu}) \propto \sigma^{n/2}(\prod_{i=1}^{n} \nu_i^{1/2}) \exp\{\sum_{i=1}^{n} \frac{(h(y_i) - \mathbf{x}_i'\boldsymbol{\beta} - \theta\nu_i)}{2\psi^2\sigma\nu}\}, \tag{3.1}$$

following, to complete the Bayesian specication, we assume normal priors for $\boldsymbol{\beta}(\tau)$, inverse Gamma for σ as in Kozumi and Kobayashi(2011). The hierarchical structure for Bayesian quantile regression for transformed survival data can be summarized as,

$$h(y_i) \mid \nu_i, \boldsymbol{\beta}(\tau), \sigma \sim N(\mathbf{x}_i'\boldsymbol{\beta}(\tau) + \theta\nu_i, \psi^2\sigma\nu_i),$$

where,

$$\nu_i \sim E(\sigma),$$

$$\boldsymbol{\beta}(\tau) \sim N(\mathbf{b}_0, \mathbf{B}_0),$$

$$\sigma \sim IG(n_0, s_0),$$

the posterior distribution and a Gibbs sampling algorithm to obtain samples from the posterior follow from Kozumi and Kobayashi (2011).

4 Two-part model

In this section, we assume that conditionally on $\mathbf{x}_i$, $log(Y_i)$ follows a form of asymmetric Laplace distribution with probability density function

$$f(\log(y_i) \mid \mathbf{x}_i) = \frac{\tau(1-\tau)}{\sigma(\tau)} \exp\{[I(\log(y_i) \leq \mathbf{x}_i'\boldsymbol{\beta}(\tau)) - \tau]\frac{\log(y_i) - \mathbf{x}_i'\boldsymbol{\beta}(\tau)}{\sigma(\tau)}\},$$

and cumulative distribution function

$$F(\log(y_i) \mid \mathbf{x}_i) = \{\tau - I(\log(y_i) > \mathbf{x}_i'\boldsymbol{\beta}(\tau))\}$$
$$\exp\{[I(\log(y_i) \leq \mathbf{x}_i'\boldsymbol{\beta}(\tau)) - \tau]\frac{\log(y_i) - \mathbf{x}_i'\boldsymbol{\beta}(\tau)}{\sigma(\tau)}\} + I(\log(y_i) > \mathbf{x}_i'\boldsymbol{\beta}(\tau)),$$

where $\boldsymbol{\beta}(\tau) \in \mathbb{R}^k$ and $\sigma(\tau) \in (0, \infty)$. In the presence of censored observations, Y_i is observed in place of T_i, and the likelihood function is proportional to

$$L(\boldsymbol{\beta}(\tau), \sigma(\tau) \mid \mathbf{y}, \mathbf{x}_i) = \prod_{i=1}^{n} (f(\log(y_i) \mid \mathbf{x}_i))^{\delta_i} (1 - F(\log(y_i) \mid \mathbf{x}_i))^{1-\delta_i}$$
$$= \prod_{i=1}^{n} (h(\log(y_i) \mid \mathbf{x}_i))^{\delta_i} s(\log(y_i) \mid \mathbf{x}_i),$$

that $h(.)$ is hazard function and $s(.)$ is survival function.

If we consider the possibility of modeling the response variable by a mixture of two distributions, a point mass distribution at c and a continuous distribution for the survival data, we can write the probability density function of Y as

$$g(y \mid \mathbf{x}, \mathbf{z}) = pI(Y = c) + (1 - p)f(y \mid \mathbf{x})I(y < c), \tag{4.1}$$

where $p = P(Y = c \mid \mathbf{z})$. Consider the interest in modelling $P(Y = c \mid \mathbf{z})$, as a function of covariates as well, then we can use a link function $\eta(.)$, $\eta : (0, 1) \to \mathbb{R}$, and write

$$\eta(p_i) = \mathbf{z}_i' \boldsymbol{\gamma}.$$

If we dene the sets $J = \{y_i : y_i = c\}$ and $K = \{y_i : y_i < c\}$, the likelihood function for the two-part model considering the location-scale mixture of the asymmetric Laplace distribution, can be defined as

$$\mathrm{L}(\boldsymbol{\beta}(\tau), \sigma, \boldsymbol{\gamma}) = \prod_{y_i \in J} \eta^{-1}(\mathbf{z}_i' \boldsymbol{\gamma}) \prod_{y_i \in K} (1 - \eta^{-1}(\mathbf{z}_i' \boldsymbol{\gamma})) f(y_i), \tag{4.2}$$

To complete the Bayesian specication, we assume priors distributions for the parameters,

$$h(Y_i)|\nu_i \sim p_i I(y_i \in J) + (1 - p_i) N(\mathbf{x}_i' \boldsymbol{\beta} + \theta\nu_i, \psi^2 \sigma \nu_i) I(y_i \in K),$$

$$\nu_i \sim E(\sigma),$$

$$\boldsymbol{\beta}(\tau) \sim N(\mathbf{b}_0, \mathbf{B}_0),$$

$$\sigma \sim IG(n_0, s_0),$$

$$\boldsymbol{\gamma} \sim N(\mathbf{g}_0, \mathbf{G}_0).$$

The full conditional distributions for all parameters, after we combine the likelihood with the prior information, are

$$\boldsymbol{\beta}(\tau) \mid \mathbf{h}(\mathbf{y}), \boldsymbol{\nu}, \sigma, \boldsymbol{\gamma} \sim N(\mathbf{b}_1, \mathbf{B}_1),$$

$$\nu_i \mid \mathbf{h}(\mathbf{y}), \boldsymbol{\beta}(\tau), \sigma, \boldsymbol{\gamma} \sim GIG(\tfrac{1}{2}, \hat{\delta}_i, \hat{\xi}_i),$$

$$\sigma \mid \mathbf{h}(\mathbf{y}), \boldsymbol{\nu}, \boldsymbol{\beta}(\tau), \boldsymbol{\gamma} \sim IG(\tfrac{\tilde{n}}{2}, \tfrac{\tilde{s}}{2}),$$

$$\pi(\boldsymbol{\gamma} \mid \mathbf{h}(\mathbf{y}), \boldsymbol{\nu}, \boldsymbol{\beta}(\tau), \sigma) \propto \prod_{i \in J} \eta^{-1}(\mathbf{z}_i' \boldsymbol{\gamma}) \prod_{i \in K} (1 - \eta^{-1}(\mathbf{z}_i' \boldsymbol{\gamma})) \exp\{-\tfrac{1}{2}(\boldsymbol{\gamma} - \mathbf{g}_0)' \mathbf{G}_0^{-1} (\boldsymbol{\gamma} - \mathbf{g}_0)\},$$

the posterior distribution for $\boldsymbol{\gamma}$ has not a recognizable distribution, so we suggest a random walk Metropolis-Hastings algorithm, where we can use as proposal a multivariate normal distribution centered at the current value of $\boldsymbol{\gamma}$. Then at the kth step of the MCMC, we draw $\boldsymbol{\gamma}_{(k)}$ from $N(\boldsymbol{\gamma}_{(k-1)}, \sigma^2_{\gamma}\Omega_{\boldsymbol{\gamma}})$, and $\boldsymbol{\gamma}_{(k)}$ is accepted with probability

$$\alpha(\boldsymbol{\gamma}_{(k)}, \boldsymbol{\gamma}_{(k-1)}) = \min\{1, \tfrac{\pi(\boldsymbol{\gamma}_{(k)}|\mathbf{h}(\mathbf{y}),\boldsymbol{\nu},\boldsymbol{\beta}(\tau),\sigma)}{\pi(\boldsymbol{\gamma}_{(k-1)}|\mathbf{h}(\mathbf{y}),\boldsymbol{\nu},\boldsymbol{\beta}(\tau),\sigma)}\},$$

where σ^2_{γ} is a tuning parameter that should be chosen carefully to give acceptance probabilities between 0.15 and 0.50 (see Gelman et al., 2003). We define $\boldsymbol{\Omega}_{\gamma}$ as the identity matrix, but other options could be considered.

In this paper, we considered censored observations, in this case, we should rewrite the density in (4.1) as,

$$g(y) = [p + (1-p)(1-F(c))]I(y = c) + (1-p)f(y)I(y < c), \tag{4.3}$$

where $F(.)$ is the cdf of the continuous part. For the complete cases, dene the sets $C = \{y_i : y_i = c, \delta_i = 1\}$, $D = \{y_i : y_i = c, \delta_i = 0\}$ and $K = \{y_i : y_i < c\}$. Then the likelihood function for $\xi = (\boldsymbol{\beta}(\tau), \sigma, \boldsymbol{\gamma})$, without writing the conditional parameters for $F(.)$ and $f(y_i|\nu_i)$ for notational simplicity, can be written as

$$\begin{aligned} L(\boldsymbol{\beta}(\tau), \sigma, \boldsymbol{\gamma}) = & \prod_{y_i \in D} \eta^{-1}(\mathbf{z}_i'\boldsymbol{\gamma}) \prod_{y_i \in C} (1 - \eta^{-1}(\mathbf{z}_i'\boldsymbol{\gamma}))(1 - F(c)) \\ & \prod_{y_i \in K} (1 - \eta^{-1}(\mathbf{z}_i'\boldsymbol{\gamma}))f(y_i|\nu_i)f(\nu_i) \end{aligned} \tag{4.4}$$

5 Simulation Study

We consider a model with just two covariates and the following structure as

$$\log(\tfrac{p_i}{1-p_i}) = \gamma_0 + \gamma_1 z_{i1} + \gamma_2 z_{i2},$$

$$\log(T_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i,$$

where $\varepsilon_i \sim N(0, 0.5^2)$, $\boldsymbol{\beta} = 3$ and $\gamma_0 = 1$, $\gamma_1 = -4$ and $\gamma_2 = 1$, we consider $x_{ij} = z_{ij}$, $j = 1, 2$, that these produce from a uniform distribution and we use the same covariates for both parts of the model. We believe that after taking into account the simulated studies throughout the literature of Bayesian quantile regression, this simple design could give a reasonable assessment of the performance of the MCMC proposed for the two-part model, as from the likelihood in Equation (4.4) the inference for each part can be made separately. For the prior hyperparameters, we assumed the $\mathbf{b}_0 = \mathbf{g}_0 = 0$ and $\mathbf{B}_0 = \mathbf{G}_0 = 100\mathbf{I}$, where $\mathbf{I}$ is the identity matrix, and for σ we considered $IG(3/2, 0.1/2)$, as in Kozumi and Kobayashi (2011). Our sample size in this study is 100 and we report our results for 10,000 replications of this model that for each parameter, we calculate the posterior mean from 10,000 draws, where we discard the first 3000 for burn-in. In Table 1, we show the summaries from the posterior means for each parameter obtained from 100 replications of this simulation. We observe that all mean and median estimates are reasonably close to their

true values, while the 95% credible interval composed by the posterior means from the replications always contain the true value.

Fig.2 shows the posterior histograms of $\boldsymbol{\beta}(0.5)$ and posterior trace of $\boldsymbol{\beta}(0.5)$ that shows of convergence parameter to true value.

Table 1: Summaries from posterior estimates regarding 100 replications of simulation.

Sample size	Parameter	True	Mean	Median	SD	2.5%	97.5%
n=100	γ_0	1	1.043	0.894	0.59	-0.229	2.04
	γ_1	-4	-4.996	-4.505	1.08	-6.84	-2.52
	γ_2	1	1.529	1.00	0.88	0.35	3.78
$\tau = 0.1$	β_0	3	2.71	2.747	0.23	2.14	3.09
	β_1	3	2.86	2.826	0.20	2.55	3.38
	β_2	3	3.07	3.084	0.23	2.58	3.51
$\tau = 0.5$	β_0	3	3.28	3.276	0.17	2.95	3.62
	β_1	3	2.75	2.761	0.17	2.41	3.07
	β_2	3	2.68	2.671	0.16	2.38	3
$\tau = 0.9$	β_0	3	3.48	3.474	0.15	3.22	3.78
	β_1	3	2.99	2.998	0.20	2.64	3.27
	β_2	3	2.76	2.769	0.19	2.43	3.07

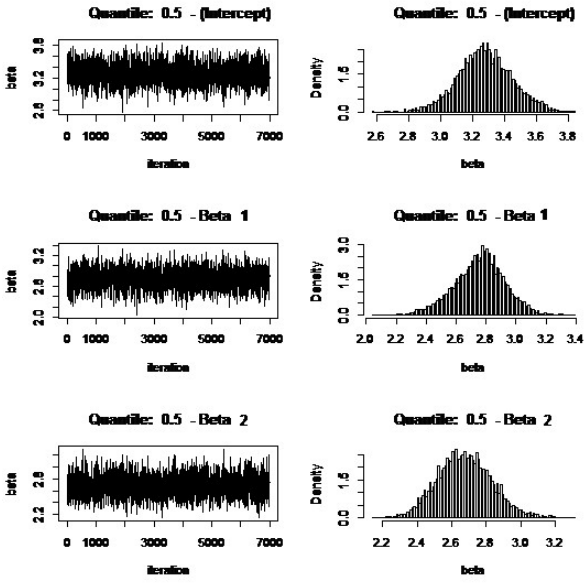


Figure 1: The posterior histograms of $\boldsymbol{\beta}(0.5)$ and posterior trace of $\boldsymbol{\beta}(0.5)$.

References

[1] Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (2003). *Bayesian data analysis*. CRC press.

[2] Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, **46**, 33-50.

[3] Kozumi, H. and Kobayashi, G. (2011). Gibbs sampling methods for Bayesian quantile regression. *Journal of Statistical Computation and Simulation*, **81**, 1565-1578.

[4] Peng, L. and Huang, Y. (2008). Survival analysis with quantile regression models. *Journal of American Statistical Association*, **103**, 637-649.

[5] Santos, B. and Bolfarine, H. (2015). Bayesian analysis for zero-or-one inated proportion data using quantile regression. *Journal of Statistical Computation and Simulation*, **85**, 3579-3593.

[6] Wang, H. J. and Wang, L. (2009). Locally weighted censored quantile regression. *Journal of the American Statistical Association*, **1046**, 1117-1128.

[7] Yu, K. and Moyeed, R. (2001). Bayesian quantile regression. *Statistics and Probability Letters*, **54**, 437-447.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Random preventive maintenance under mixed effects

Atlekhani, M. [1] and Doostparast, M. [2]

[1] Department of Statistics, Malayer University
[2] Department of Statistics, Ferdowsi University of Mashhad

Abstract

This paper extends a random preventive maintenance scheme, called repair alert model, when there exist environmental variables which effect on system lifetimes. It can be used for implementing age-dependent maintenance policies on engineering devices. The new model is flexible to including covariates with both fixed and random effects. The problem of estimating parameters is also investigated in details. To do this, it is assumed that the system lifetime distribution belongs to the log-location-scale family of distributions. A real data set is also analysed on the of the results obtained.

Keywords: Random signs censoring, Repair alert model, Log-location-scale family, Fixed and random effects.

1 Introduction

Suppose a device works for a job with random working times. The device is replaced at random time Z $(0 < Z \leq \infty)$ or at failure time X, whichever occurs first, where Z is a random variable with a general distribution function (DF) $F_Z(z) = P(Z \leq z)$ and is usually assumed to be independent of the failure time X. This policy is called *random age replacement maintenance*; See Chapter 2 in Nakagawa [10] for more information. Hence, one observes only the pair (Y, δ), where $Y = \min(X, Z)$

[1] m_akhani@yahoo.com
[2] doustparast@um.ac.ir

and $\delta = I(X < Z)$ is the indicator function of the event $\{X < Z\}$. When Z and X are independent, the marginal DFs of X and Z, denoted by F_X and F_Z respectively, are identifiable on the basis of the joint DF (Y, δ); See Theorem 1 in Cook [5] for more details. The maintenance engineer may have some information on the state of the device performance during the operation time and use it to (heuristically) do maintenance (replacement) for avoiding more costly random device failures. It is therefore reasonable to expect a (positive) dependence between the failure time X and the preventive maintenance (PM) time Z; See Lindqvist et al. [7] for more details. But from the *competing risk theory*, the marginal DFs are generally nonidentifiable from the observed (Y, δ) without making assumptions on the dependence between X and Z (Tsiatis [11]). Cook [6] proposed the concept of *random signs censoring* (called also *age-dependent censoring* by Cook [5]), in which the marginal DF of the failure time X is identifiable. As mentioned by Cook [5], in the random signs censoring (age-dependent censoring) the probability of a device being censored is independent of its age, but given that it is censored, the time at which it is censored may depend on its age. In other words, the random sign censoring means that the event that the failure of the device is preceded by PM is not influenced by the time X at which the device fails without PM; See Lindqvist et al. [7] for a greater detail. Lindqvist et al. [7] introduced the repair alert model (RAM) for analysing the mentioned PM scheme. The RAMs have been studied in literature; See for example, Christen et al. [4] and Atlekhani and Doostparast [2]. According to the authors knowledge the RAM has been investigated without considering the environmental factors such as temperature, voltage, pressure and etc. As mentioned by Meeker and Escobar [9], a model with explanatory variables sometimes explains or predicts why some units fail quickly and other units survive a long time. This paper considers the RAM with explanatory variables. The general idea of a given RAM with covariates is to express the failure-time distribution as a function of the explanatory variables. To do this, we allow one or more of the elements of the model parameter vector, say $\boldsymbol{\Theta} = (\theta_1, \cdots, \theta_k)$ where r is the number of parameters, to be a function of the explanatory variables. Therefore, the rest of this paper is organized as follows. The mixed RAMs which include both fixed and random effects are discussed in Section 2. A real data set on transmission belt lifetimes in a production line is analysed in Section 3 using the obtained results in the preceding sections. Section 4 concludes.

2 RAMs with mixed effects

In Table 1 notations and acronyms are used throughout the paper. Some basic definitions to introduce RAMs are now given.

Definition 1 (Cook [6]). *Let X and Z be lifetime variables with $Z = X - \xi$, where ξ is a random variable, $\xi \leq X$, $P(\xi = 0) = 0$, whose sign is independent of X. The random vector $(\min(\{X, Z\}, I(X < Z))$ is called "*random signs censoring of X by Z*".*

Definition 2. *The sub-DFs of X and Z are defined by $F_X^*(x) = P(X \leq x, X < Z)$, and $F_Z^*(z) = P(Z \leq z, Z < X)$, respectively. The conditional DFs of X and Z are defined by $\tilde{F}_X(x) = P(X \leq x | X < Z)$, and $\tilde{F}_Z(z) = P(Z \leq z | Z < X)$, respectively.*

Lindqvist et al. [7] provided the following formal definition of RAMs.

Table 1: Notations and acronyms

Notations and acronyms	Description
X	The system lifetime
Z	The repair time
$F_X(x)$	$P(X \leq x)$, distribution function of X
$F_Z(z)$	$P(Z \leq z)$, distribution function of Z
$F_X^*(x)$	$P(X \leq x, X < Z)$, sub distribution function of X
$F_Z^*(z)$	$P(Z \leq z, Z < X)$, sub distribution function of Z
$\tilde{F}_X(x)$	$P(X \leq x \mid X < Z)$, conditional distribution function of X
$\tilde{F}_Z(z)$	$P(Z \leq z \mid Z < X)$, conditional distribution function of Z
RAM	Repair alert model
RSCD	Random sign censoring data
PM	Preventive maintenance
LF	Likelihood function
LLF	Log-likelihood function
ML	Maximum likelihood
CRAF	Cumulative repair alert function
DF	Distribution function
RAF	Repair alert function

Definition 3. *The pair (X, Z) satisfies the requirements of the RAM if*

(i) Z is a random signs censoring of X; i.e. the event $\{Z < X\}$ is stochastically independent of X,

(ii) There exists an increasing function G defined on $[0, \infty)$ with $G(0) = 0$, such that for all $x > 0$

$$P(Z \leq z \mid Z < X, X = x) = \frac{G(z)}{G(x)}, \qquad 0 < z \leq x. \tag{2.1}$$

Theorem 2.1. *(Lindqvist et al. [7].) Suppose that the pair (X, Z) follows a RAM with parameters $F_X(x), G(t)$ and $q = P(Z < X)$. Then, for $t > 0$,*

$$\tilde{F}_X(t) = F_X(t), \tag{2.2}$$

$$F_X^*(t) = (1 - q)F_X(t), \tag{2.3}$$

$$\tilde{F}_Z(t) = F_X(t) + G(t) \int_t^\infty \frac{f_X(y)}{G(y)} dy, \tag{2.4}$$

$$\tilde{f}_Z(t) = g(t) \int_t^\infty \frac{f_X(y)}{G(y)} dy, \tag{2.5}$$

$$F_Z^*(t) = P(Z \leq t, Z < X) = q\tilde{F}_Z(t). \tag{2.6}$$

A natural approach to incorporate the covariates in the performance of a component lifetime is to allow some parameters of the RAM be a function of the covariates. To do this, we assume that the lifetime X belongs to the well-known log-location-scale family.

Definition 4. *A random variable X belongs to the log-location-scale family of distributions if*

$$F_X(t) = P(X \leq t) = \Phi\left(\frac{\log(t) - \mu}{\sigma}\right), \quad \forall\, t > 0, \tag{2.7}$$

where Φ is a baselines DF and does not depend on μ and σ.

Suppose that in addition to the fixed effects, there are some random effects. In the location-scale family, we can consider the location parameter μ in Equation (2.7) is a linear function of the fixed and random effects. Let there exist r fixed and b random effects, denoted by $w_1, \cdots, w_r$ and $U_1, \cdots, U_b$, respectively. Conditionally on the random effects $(U_1, \cdots, U_b) = (u_1, \cdots, u_b)$, we have $\mu = \sum_{i=1}^{r} w_i\beta_i + \sum_{j=1}^{b} u_j$. For a given RSCD, let $\boldsymbol{\mu} = \mathbf{w}\boldsymbol{\beta} + \boldsymbol{\alpha}_1\mathbf{u}_1 + \cdots + \boldsymbol{\alpha}_b\mathbf{u}_b$ where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_N)^T$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_r)^T$ and $[[w_{ij}]]_{1\leq i\leq N,\ 1\leq j\leq r}$ is the known model matrix for the fixed effects $w_1, \cdots, w_r$, $\mathbf{u}_i = (u_{i1}, \cdots, u_{ip_i})^T$, for $1 \leq i \leq b$, denotes the observed levels of the i-th random effects $\mathbf{U}_i$ and $\boldsymbol{\alpha}_i$ stands for the corresponding matrix model. In sequel, we need the following definition.

Definition 5. *The triple (X, Z, U) is called the conditional RAM (CRAM) on $\boldsymbol{u}$ if*

(i) The event $\{Z < X|\boldsymbol{u}\}$ is stochastically independent of $X|\boldsymbol{u}$;

(ii) There exists an increasing function $G(.|\boldsymbol{u})$ defined on $[0, \infty)$ with $G(0|\boldsymbol{u}) = 0$, such that for all $x > 0$,

$$P(Z \leq z|Z < X, X = x, \boldsymbol{U} = \boldsymbol{u}) = \frac{G(z|\boldsymbol{u})}{G(x|\boldsymbol{u})}, \qquad 0 < z \leq x. \tag{2.8}$$

Proposition 2.2. *(Atlekhani and Doostparast [3].) Suppose that the triple (X, Z, U) follows a CRAM then, for $t > 0$. Then*

$$\tilde{F}_{X|\boldsymbol{u}}(t|\boldsymbol{u}) = F_{X|\boldsymbol{u}}(t|\boldsymbol{u}), \tag{2.9}$$

$$F^*_{X|\boldsymbol{u}}(t|\boldsymbol{u}) = (1 - q_{\boldsymbol{u}})F_{X|\boldsymbol{u}}(t|\boldsymbol{u}), \tag{2.10}$$

$$\tilde{F}_{Z|\boldsymbol{u}}(t|\boldsymbol{u}) = F_{X|\boldsymbol{u}}(t|\boldsymbol{u}) + G(t|\boldsymbol{u}) \int_t^{\infty} \frac{f_{X|\boldsymbol{u}}(x|\boldsymbol{u})}{G(x|\boldsymbol{u})} dx, \tag{2.11}$$

$$\tilde{f}_{Z|\boldsymbol{u}}(t|\boldsymbol{u}) = g(t|\boldsymbol{u}) \int_t^{\infty} \frac{f_{X|\boldsymbol{u}}(x|\boldsymbol{u})}{G(x|\boldsymbol{u})} dx, \tag{2.12}$$

$$F^*_{Z|\boldsymbol{u}}(t|\boldsymbol{u}) = q_{\boldsymbol{u}}\tilde{F}_{Z|\boldsymbol{u}}(t|\boldsymbol{u}). \tag{2.13}$$

2.1 Weibull family

For sake of brevity, it is assumed that there exist only one fixed effect and one random effect, denoted by w and U, respectively. Then $r = b = 1$. Given $U = u$, the system lifetime X follows the Weibull distribution. For a given RSCD, the response variable can be represented as

$$y_{lj} = x_{lj}\delta_{lj} + z_{lj}(1 - \delta_{lj}) \quad j = 1, \ldots, N_l, \quad l = 1, \ldots, p, \tag{2.14}$$

where $\delta_{lj} = 1$ if $x_{lj} < z_{lj}$ and 0 if otherwise. Here, it is assumed that the covariates are not censored and are independent of the censoring mechanism. Thus, observations are $\{(y_{lj}, w_{lj}), \quad j = 1, \ldots, N_l, l = 1, \ldots, p\}$. Here $y_{lj} = \min(x_{lj}, z_{lj})$ so that the form of data is $\{(x_{li}, w_{li}), \quad i = 1, \ldots, m_l\}$ and $\{(z_{lj}, w_{lj}), \quad j = 1, \ldots, n_l\}$, where $m_l + n_l = N_l, l = 1, \ldots, p$, and N_l is the number of observation in the l-th level of the random effect. Note that $\sum_{l=1}^{p} N_l = N$. Hence, for the j-th observation (for $1 \leq j \leq N_l$) at the l-th level of the random effect U, the location parameter is

$$\mu_{lj} = \beta_0 + \beta_1 w_{lj} + u_l, \quad 1 \leq j \leq N_l, \ 1 \leq l \leq p. \tag{2.15}$$

Here the random effects u_l are independent and $E(u_l) = 0$ and $Var(u_l) = \sigma_u^2$ for $l = 1, \cdots, p$; See e.g. McCulloch et al. [8]. In order to calculate the LF from Equations (2.9)-(2.13) we have

$$\tilde{F}_{X|u}(t|u) = 1 - \exp\Big(- \exp\Big(\frac{\log(t) - \mu}{\sigma}\Big)\Big), \tag{2.16}$$

$$F^*_{X|u}(t|u) = (1 - q_u)\Big(1 - \exp\Big(- \exp(\frac{\log(t) - \mu}{\sigma})\Big)\Big), \tag{2.17}$$

$$\begin{aligned}\tilde{F}_{Z|u}(t|u) &= 1 - \exp\Big(- \exp\Big(\frac{\log(t) - \mu}{\sigma}\Big)\Big) \\ &\quad + G(t|u)\int_t^\infty \frac{\exp(\frac{\log x - \mu}{\sigma} - \exp(\frac{\log x - \mu}{\sigma}))}{\sigma x G(x|u)}dx,\end{aligned} \tag{2.18}$$

$$\tilde{f}_{Z|u}(t|u) = g(t|u)\int_t^\infty \frac{\exp\Big(\frac{\log x - \mu}{\sigma} - \exp(\frac{\log x - \mu}{\sigma})\Big)}{\sigma x G(x|u)}dx, \tag{2.19}$$

$$F^*_{Z|u}(t|u) = q_u \tilde{F}_{Z|u}(t|u), \tag{2.20}$$

where $g(t|u) = \frac{\partial}{\partial t}G(t|u)$. For the conditional CRAF, let $G(t|u_l) = u_l t^k$ where k is a integer. From Equations (2.16)-(2.20), the likelihood contribution from an observation x_{li} for $l = 1 \ldots p$ and $i = 1, \ldots, m$ is

$$\begin{aligned}f^*_{X|u}(x_{li}|u_l) &= (1 - q_{u_l})\frac{1}{\sigma x_{li}} \exp\left\{\frac{\log x_{li} - \mu_{li}}{\sigma} - \exp\left\{\frac{\log x_{li} - \mu_{li}}{\sigma}\right\}\right\} \\ &= \frac{\exp(-\frac{\mu}{\sigma})(1 - q_{u_l})}{\sigma} x_{li}^{\frac{1}{\sigma}-1} \exp\left\{ - x_{li}^{\frac{1}{\sigma}}.\exp(-\frac{\mu_{li}}{\sigma})\right\},\end{aligned} \tag{2.21}$$

while the contribution from z_{lj} is

$$f^*_{Z|u}(z_{lj}|u_l) = kq_{u_l}u_l z_{lj}^{k-1}\int_{z_{lj}}^{\infty}\frac{\exp\left\{\frac{\log x-\mu_{lj}}{\sigma}-\exp\left\{\frac{\log x-\mu_{lj}}{\sigma}\right\}\right\}}{u_l x^k\sigma x}dx \tag{2.22}$$
$$= kq_{u_l}z_{lj}^{k-1}\exp(-\frac{\mu_{lj}}{\sigma})\frac{1}{\sigma}\int_{z_{lj}}^{\infty}x^{\frac{1}{\sigma}-(k+1)}\exp\left\{-x^{\frac{1}{\sigma}}.\exp(-\frac{\mu_{lj}}{\sigma})\right\}dx,$$

for $j = 1,\ldots,n$. If u_l is independent and identically distributed normal random variables with mean 0 and variance σ_u^2 then the associate LFF of the available RSCD is

$$\begin{aligned} l(\boldsymbol{\beta},\sigma,\sigma_u^2,q_u;\mathbf{x},\mathbf{z},\mathbf{w},\mathbf{u}) = &-p\log(\sqrt{2\pi}\sigma_u)+\sum_{l=1}^{p}n_l\log k-(m_l+n_l)\log\sigma \\ &+\sum_{l=1}^{p}\log\Bigg(\int_{-\infty}^{\infty}(1-q_{u_l})^{m_l}q_{u_l}^{n_l} \\ &\prod_{i=1}^{m_l}\exp(-\frac{\mu_{li}}{\sigma})x_{li}^{\frac{1}{\sigma}-1}\exp\left\{-x_{li}^{\frac{1}{\sigma}}\exp(-\frac{\mu_{li}}{\sigma})\right\} \\ &\prod_{j=1}^{n_l}z_{lj}^{k-1}\exp(-\frac{\mu_{lj}}{\sigma}) \\ &\int_{z_{lj}}^{\infty}x^{\frac{1}{\sigma}-(k+1)}\exp\left\{-x^{\frac{1}{\sigma}}.\exp(-\frac{\mu_{lj}}{\sigma})-\frac{u_l^2}{2\sigma_u^2}\right\}dxdu_l\Bigg). \end{aligned} \tag{2.23}$$

Note that the maximum of $(1-q_{u_l})^{m_l}q_{u_l}^{n_l}$ is $\hat{q}_{u_l} = \frac{n_l}{n_l+m_l}$. There is no explicit expression for the ML estimates and one may use numerical methods.

3 Transmission belt data set

A case study was carried out in one of the production lines in a processing industry in Malaysia (Amad et al 2011). This production line comprises a series of processing machines that perform different operations such as rewinding, embossing, cutting and packaging. In this serial configuration, a breakdown in at least in one of the machines will cause the production line to stop. Currently, the main problem with this setup is the high rate of machine breakdowns due to "unplanned maintenances". Many cases of unplanned maintenances are due to component failures. Ahmad et al. [1] reports that the highest record of component failures is on the transmission belt. Table 2 presents the failure data set of the transmission belt with censored and uncensored classification. This data set is analysed in this section to determine the failure distribution parameters and here to estimate the effects of some covariates on the transmission belt failure. Three covariates, namely dust factor (DUST), related component factor (RCOMP) and product type factor (PRODT), that may influence the transmission belt failure are identified

Table 2: Failure times of the transmission belt and the codified values of the influencing covariates

No. of failure	Belt lifetime	Censored = 0 and uncensored= 1	Dust	RCOMP	PRODT
1	28	1	−1	−1	0.6
2	52	1	+1	+1	0.4
3	42	0	+1	−1	0.4
4	8	1	−1	−1	0.3
5	14	1	−1	−1	0.6
6	13	1	−1	−1	0.6
7	47	1	+1	+1	0.5
8	38	1	−1	+1	0.6
9	25	1	−1	+1	0.4
10	12	1	−1	−1	0.8
11	50	0	−1	+1	0.6
12	42	1	−1	−1	0.6

Dust factor (DUST) with coding −1 (bad condition) means the component exposed to extreme dust condition without any preventive action (cleaning) and +1 means the component is recorded for cleaning activity. The related component factor (RCOMP) with coding −1 (bad condition) stands for the related component (e.g. the bearing in a pulley for supporting the belt operation) is NOT replace together with the component (transmission belt) and +1 denotes the related component (e.g. the bearing in a pulley for supporting the belt operation) is replace together with the component (transmission belt). The product type factor (PRODT) was formulated in continuous form and it is based on the product types which are hard and soft. For example, assume that the failure time of the component (transmission belt) is 10 days, then the types of product that produce at each day (failure time = 10 working days) is recorded. For example, 3 days have produced soft types of products and 7 days have produced hard types of product. Therefore, this covariate (PRODT) is codified as 0.7 (%), referred to hard type product. It is because, the transmission belt needs more force (more risk to fail) to process the hard type product compared to the soft type product.

In this example, "Dust" considered as a random effect and "PRODT " as a fixed effect. We consider the component follows the Weibull distribution and the conditional CRAF be $G(t|u_l) = u_l t^k$ as in Subsection 2.1. The ML estimates of the parameters, derived by maximizing the LLF (2.23), are given in Table 3. The Akaike's information criterion (AIC) is given by $AIC = -2\widehat{LLF} + 2M$, where $\widehat{LLF}$ is the LLF at the ML estimates and M is the number of the parameters of the model. From Table 3, the model with $k = 8$ has the lowest Akaike. So the RAMs with the conditional CRAF $G(t|u_l) = u_l t^8$ is the best among the considered models. On the other hand, it can be seen that the "PRODT" with the negative coefficient is effective for the lifetime of the belt. Therefore the harder product the less lifetime. Consider the following problem of hypothesis

Table 3: ML estimates for transmission belts data

k	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\sigma}$	$\hat{\sigma_u}$	LLF	AIC
1	2.51	−1.09285	0.483172	0.8915	−58.7881	127.576
2	3.0658	−1.35247	0.5158	3.5656	−58.7785	127.557
3	3.5108	−1.14828	0.501579	1.98	−58.2825	126.565
4	3.55803	−2.28736	0.49	3.65	−58.2865	126.573
5	3.47849	−1.10249	0.49	1.98	−57.8294	125.659
6	2.64467	−0.562965	0.49	3.60708	−57.8056	125.611
7	2.41858	−0.115864	0.458149	3.19141	−58.3604	126.721
8	**2.87199**	**−1.02092**	**0.4765**	**3.6889**	**−57.5749**	**125.15**
9	2.26947	0.336346	0.484392	3.47709	−58.0514	126.103
10	2.17317	0.229668	0.582577	3.32856	−58.21	126.42

testing

$$\begin{cases} H_0 : \beta_1 = 0, \\ H_0 : \beta_1 \neq 0. \end{cases}$$

The generalized likelihood ratio (GLR) test rejects H_0 at level α if $-2\log\frac{L_0^*}{L_1^*} > \chi^2_{1,1-\alpha}$ where $\chi^2_{\nu,1-\alpha}$ is $(1-\alpha)$-th quantile of the chi-square distribution with ν degrees of freedom. For $\beta_1 = 0$, we derived for Equation (2.23) $\log L_0^* = -63.6372$. Also, from Table 3 we get

$$-2\log\frac{L_0^*}{L_1^*} = -2(\log L_0^* - \log L_1^*) = -2 \times (-63.6372 + 57.5749) = 12.1246.$$

The p-value is $P(\chi_1^2 > 12.1246) = 0.0004976$ and therefore the null hypothesis H_0 is rejected at the significant level $\alpha = 0.05$. In summarize, "PRODT" has effect on the component lifetime as we expected.

4 Conclusions

This paper generalized the repair alert model when there are some additional environmental variables which may effect the system lifetime. Both random and fixed effect covariates were considered in the proposed model. The problem of estimating parameters was studied in details. It was assumed that the distribution of the system lifetime belongs to the log-location-scale family of distributions. The results of this paper may be extended in some directions. For example, one may consider other families of lifetime distributions such as log-Gamma distributions. Also, the problem of estimating parameters via a Bayesian approach is worth.

References

[1] Ahmad, R., Kamaruddin, S., Azid, I. A., Almanar, I. P. (2011). Preventive replacement schedule: A case study at a processing industry. *International Journal of Industrial and Systems Engineering*, **8**, 386-406.

[2] Atlekhani, M., Doostparast, M. (2016). Inference under random maintenance policies with a polynomial form as the repair alert function. *Journal of Statistical Computation and Simulation*, **80**, 1415-1429.

[3] Atlekhani, M., Doostparast, M. (2017). The repair alert models with fixed and random effects. Submitted.

[4] Christen, J. A., Ruggeri, F., Villa, E. (2011), Utility based maintenance analysis using a random sign censoring model. *Reliability Engineering and System Safety*, **96**, 425-431.

[5] Cook, R. M. (1993). The total time on test statistics and age-dependent censoring. *Statistics and Probability Letters*, **18**, 307-312.

[6] Cook, R. M. (1996). The design of reliability data bases, Part I and II: review of standard design concepts. *Reliability Engineering and System Safety*, **51**, 137-146, and 209-223.

[7] Lindqvist, B. H., Støve, B., Langseth, H. (2006). Modelling of dependence between critical failure and preventive maintenance: The repair alert model. *Journal of Statistical Planning and Inference*, **136**, 1701-1717.

[8] McCulloch, Ch.E., Searle, SH. R., Neuhaus, J. M. (2008). *Generalized linear and mixed models*. John Wiley, Canada.

[9] Meeker, W. O. and Escobar, L. A. (1998). *Statistical methods for reliability data*. John Wiley & Sons, Toronto.

[10] Nakagawa, T. (2014). *Random maintenance policies*. Springer-Verlag, London.

[11] Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences*, **72**, 20-22.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

New results on stochastic comparisons of series and parallel systems with generalized exponential components

Bashkar, E. [1] Torabi, H. [2] and Dolati, A. [3]

[1,2,3] Department of Statistics, Yazd University

Abstract

In this paper, we derive new results on stochastic comparisons of series and parallel systems with independent heterogeneous generalized exponential components. These comparisons are with respect to the usual stochastic ordering, the reversed failure rate ordering and the failure rate ordering. The result established here strengthens and generalizes some of the results of Balakrishnan et al. (2015).

Keywords: Generalized exponential distribution, Stochastic ordering, Majorization, Parallel system, Series system.

1 Introduction

Consider a distribution function F, and let $\alpha > 0$. We know that $G(x) = (F(x))^{\alpha}$ is also a distribution function and is known as exponentiated distribution. It is also known as proportional reversed hazard rate model (PRHRM). A flexible model which belongs to the exponentiation family, is the exponentiated Weibull (EW) distribution proposed by Mudholkar and Srivastava [9]. The EW distribution is quite adequate for modeling non-monotone failure rates, including the bathtub shaped hazard rate, which are quite common in reliability and biological studies.

[1]e.bashkar@stu.yazd.ac.ir
[2]htorabi@yazd.ac.ir
[3]adolati@yazd.ac.ir

Recently, a sub-model of the EW distribution, called the generalized exponential (GE) distribution, has been discussed extensively. Recall that, a random variable X is said to be distributed according the generalized exponential distribution, and will be denoted by $X \sim GE(\alpha, \lambda)$, if the distribution function is given by

$$G(x) = (1 - \exp\{-\lambda x\})^{\alpha}, x > 0,$$

where $\alpha > 0$ is a shape parameter and $\lambda > 0$ is a scale parameter. It can be shown that the GE has increasing hazard rate (IHR), if $\alpha > 1$, constant hazard rate (CHR), if $\alpha = 1$, and decreasing hazard rate (DHR) if $\alpha < 1$ [4].

To fit a model to lifetime data sets, it is observed in the literature that the generalized exponential distribution can be used as an alternative to the gamma, Weibull, and log-normal distributions. So, it is of interest to consider this distribution as the lifetimes of components of reliability systems. Balakrishnan et al. [1] examined the problem of the stochastic comparison of series and parallel systems with heterogeneous GE components. In this paper, we derive new results on stochastic comparisons of series and parallel systems with independent heterogeneous generalized exponential components. These comparisons are with respect to the usual stochastic, the reversed failure rate and the failure rate orderings. The result established here strengthens and generalizes some of the results of [1].

Let $X_{1:n} \leq \ldots \leq X_{n:n}$ denote the order statistics arising from random variables $X_1, \ldots, X_n$. Order statistics play important rules in statistical inference, reliability theory, life testing, operations research and other related areas. In reliability theory, the kth order statistic coresponds to the lifetime of a $(n - k + 1)$-out-of-n system. In particular, $X_{1:n}$ and $X_{n:n}$ correspond to the lifetimes of series and parallel systems, respectively. Various researchers have studied the stochastic comparisons for the lifetimes of the series and parallel systems, for example, see [3, 5, 6, 7, 11, 12] and references cited therein for details on these comparisons. To continue further we need some definitions which are standard in the literature. The terms such as increasing and decreasing will be used for non-decreasing and non-increasing, respectively throughout the paper.

Let X and Y be two univariate random variables with distribution functions F and G, density functions f and g, the survival functions $\bar{F} = 1 - F$ and $\bar{G} = 1 - G$, hazard rate functions $r_F = f/\bar{F}$ and $r_G = g/\bar{G}$, and reverse hazard rate functions $\tilde{r}_F = f/F$ and $\tilde{r}_G = g/G$, respectively. Random variable X is said to be smaller than Y in the

(i) hazard rate order, denoted by $X \leq_{\text{hr}} Y$, if $r_F(x) \geq r_G(x)$ for all x;

(ii) reversed hazard rate order, denoted by $X \leq_{\text{rh}} Y$, if $\tilde{r}_F(x) \leq \tilde{r}_G(x)$ for all x;

(iii) stochastic order, denoted by $X \leq_{st} Y$, if $\bar{F}(x) \leq \bar{G}(x)$ for x.

For a comprehensive discussion on various stochastic orders, one can see [10].

The concept of majorization is quite useful in dealing with various reliability related optimization problems. For preliminary notations and terminologies on majorization theory, we refer the reader to [8]. Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ be two real vectors and $x_{(1)} \leq \ldots \leq x_{(n)}$ be the increasing arrangement of the components of the vector $\mathbf{x}$.

Definition 1. The vector $\boldsymbol{x}$ is said to be

(i) weakly submajorized by the vector $\boldsymbol{y}$ (denoted by $\boldsymbol{x} \preceq_{\mathrm{w}} \boldsymbol{y}$) if $\sum_{i=j}^{n} x_{(i)} \leq \sum_{i=j}^{n} y_{(i)}$ for all $j = 1, \ldots, n$,

(ii) weakly supermajorized by the vector $\boldsymbol{y}$ (denoted by $\boldsymbol{x} \overset{\mathrm{w}}{\preceq} \boldsymbol{y}$) if $\sum_{i=1}^{j} x_{(i)} \geq \sum_{i=1}^{j} y_{(i)}$ for all $j = 1, \ldots, n$,

(iii) majorized by the vector $\boldsymbol{y}$ (denoted by $\boldsymbol{x} \overset{\mathrm{m}}{\preceq} \boldsymbol{y}$) if $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$ and $\sum_{i=1}^{j} x_{(i)} \geq \sum_{i=1}^{j} y_{(i)}$ for all $j = 1, \ldots, n-1$.

Clearly, $\boldsymbol{x} \overset{\mathrm{m}}{\preceq} \boldsymbol{y}$ implies both $\boldsymbol{x} \overset{\mathrm{w}}{\preceq} \boldsymbol{y}$ and $\boldsymbol{x} \preceq_{\mathrm{w}} \boldsymbol{y}$.

Another interesting weaker order regarding to the majorization order introduced by [2] is the p-larger order which is defined as follows:

Definition 2. A vector $\mathbf{x}$ in $\mathbb{R}_+^n$ is said to be p-larger than another vector $\mathbf{y}$ in $\mathbb{R}_+^n$ (denoted by $\mathbf{x} \overset{\mathrm{p}}{\succeq} \mathbf{y}$) if

$$\prod_{i=1}^{j} x_{(i)} \leq \prod_{i=1}^{j} y_{(i)}, \quad j = 1, \ldots, n,$$

where $\mathbb{R}_+^n = \{(x_1, ..., x_n) | x_i \geq 0,\ i = 1, ..., n\}$.

It is known that for any two non-negative vectors $\mathbf{x}$ and $\mathbf{y}$,

$$\mathbf{x} \overset{\mathrm{m}}{\preceq} \mathbf{y} \Longrightarrow \mathbf{x} \overset{\mathrm{p}}{\preceq} \mathbf{y}.$$

Balakrishnan et al. [1] had studied the stochastic comparisons of parallel (series) systems having heterogeneous GE components. In section 2, we derive new results on stochastic comparisons of series and parallel systems with independent heterogeneous generalized exponential components.

2 Main result

In the following theorem, we compare series systems with independent heterogeneous GE components when one of the parameters is fixed, and the results are then developed with respect to the other parameter. This result generalizes the corresponding result in Theorem 15 of [1]. In particular the majorization assumption is relaxed to the weak supermajorization.

Theorem 2.1. Let $X_1, \ldots, X_n$ $(X_1^*, \ldots, X_n^*)$ be independent random variables with $X_i \sim GE(\alpha_i, \lambda)$ $(X_i^* \sim GE(\alpha_i^*, \lambda))$, $i = 1, ..., n$. Then, for any $\lambda > 0$, we have

$$(\alpha_1, \ldots, \alpha_n) \overset{\mathrm{w}}{\succeq} (\alpha_1^*, \ldots, \alpha_n^*) \Longrightarrow X_{1:n} \leq_{\mathrm{hr}} X_{1:n}^*.$$

Proof. Fix $x > 0$. Then the hazard rate of $X_{1:n}$ is

$$r_{X_{1:n}}(x, \boldsymbol{\alpha}, \lambda) = \sum_{i=1}^{n} \frac{\alpha_i \lambda e^{-\lambda x}(1 - e^{-\lambda x})^{\alpha_i - 1}}{1 - (1 - e^{-\lambda x})^{\alpha_i}} = \lambda e^{-\lambda x} \sum_{i=1}^{n} \varphi(\alpha_i),$$

where $\varphi(\alpha) = \dfrac{\alpha(1 - e^{-\lambda x})^{\alpha - 1}}{1 - (1 - e^{-\lambda x})^{\alpha}}$, $x \geq 0, \lambda \geq 0$. From Theorem 3.A.8 of [8] it suffices to show that, for each $x > 0$, $r_{X_{1:n}}(x, \boldsymbol{\alpha}, \lambda)$ is Schur-convex and decreasing in α_i's. By taking $t = 1 - e^{-\lambda x}$ in Lemma 7 of [1], we see that $\varphi(\alpha)$ is convex in $\alpha \geq 0$ and it is easy to show that is a decreasing function, then the hazard rate function of $X_{1:n}$ is decreasing and convex in each α_i.

So, from Proposition 3.C.1 of [8], the Schur-convexity of $r_{X_{1:n}}(x, \boldsymbol{\alpha}, \lambda)$ follows from convexity of $\varphi(\alpha)$.This completes the proof of the required result. □

The following result considers the comparison on the lifetimes of series systems in terms of the usual stochastic order when two sets of scale parameters weakly majorize each other.

Theorem 2.2. Let $X_1, \ldots, X_n$ $(X_1^*, \ldots, X_n^*)$ be independent random variables with $X_i \sim GE(\alpha, \lambda_i)$ $(X_i^* \sim GE(\alpha, \lambda_i^*))$, $i = 1, ..., n$. If $0 < \alpha \leq 1 (\alpha \geq 1)$ and $(\lambda_1, \ldots, \lambda_n) \overset{\mathrm{w}}{\succeq} (\succeq_{\mathrm{w}})(\lambda_1^*, \ldots, \lambda_n^*)$, then $X_{1:n} \geq_{\mathrm{st}} (\leq_{\mathrm{st}}) X_{1:n}^*$.

Proof. For a fixed $x > 0$, the survival function of $X_{1:n}$ can be written as

$$\overline{F}_{X_{1:n}}(x, \boldsymbol{\lambda}) = \prod_{i=1}^{n} \left(1 - e^{-\lambda_i x}\right)^{\alpha}. \tag{2.1}$$

Now, using Theorem 3.A.8 of [8], it is enough to show that the function $\overline{F}_{X_{1:n}}(x, \boldsymbol{\lambda})$ given in (2.1) is Schur-convex (Schur-concave) and decreasing in λ_i's.

The partial derivatives of $\overline{F}_{X_{1:n}}(x, \boldsymbol{\lambda})$ with respect to λ_i is given by

$$\frac{\partial \overline{F}_{X_{1:n}}(x, \boldsymbol{\lambda})}{\partial \lambda_i} = -x \overline{F}_{X_{1:n}}(x, \boldsymbol{\lambda}) q(\alpha, 1 - e^{-\lambda_i x}),$$

where $q(\alpha, t) = \alpha \dfrac{(1 - t)t^{\alpha - 1}}{1 - t^{\alpha}} \geq 0, 0 < t < 1$. So, we have that $\overline{F}_{X_{1:n}}(x, \boldsymbol{\lambda})$ is decreasing in each λ_i.

From Theorem 3.A.4. in [8] the Schur-convexity (Schur-concavity) follows if we prove that, for any $i \neq j$,

$$(\lambda_i - \lambda_j)\left(\frac{\partial \overline{F}_{X_{1:n}}(x, \boldsymbol{\lambda})}{\partial \lambda_i} - \frac{\partial \overline{F}_{X_{1:n}}(x, \boldsymbol{\lambda})}{\partial \lambda_j}\right) \geq (\leq) 0,$$

that is, for $i \neq j$,

$$x \overline{F}_{X_{1:n}}(x, \boldsymbol{\lambda})(\lambda_i - \lambda_j)\left(q(\alpha, 1 - e^{-\lambda_j x}) - q(\alpha, 1 - e^{-\lambda_i x})\right) \geq (\leq) 0. \tag{2.2}$$

According to Lemma 3 of [1], $q(\alpha, t)$ is decreasing (increasing) in $0 < t < 1$ for any $0 < \alpha \leq 1$ $(\alpha \geq 1)$, which in turn implies that the function $q(\alpha, 1 - e^{-\lambda_i x})$ is decreasing (increasing) in $\lambda_i, i = 1, \ldots, n$ for any $0 < \alpha \leq 1$ $(\alpha \geq 1)$. This completes the proof of the required result. □

Note that $(\lambda_1,\ldots,\lambda_n) \overset{\text{m}}{\succeq} (\lambda_1^*,\ldots,\lambda_n^*)$ implies both $(\lambda_1,\ldots,\lambda_n) \overset{\text{w}}{\succeq} (\lambda_1^*,\ldots,\lambda_n^*)$ and $(\lambda_1,\ldots,\lambda_n) \succeq_{\text{w}} (\lambda_1^*,\ldots,\lambda_n^*)$, Theorem 2.2 substantially improves the corresponding ones provided by [1], in the sense that the majorization is relaxed to the weak majorization. Naturally, one may wonder whether the following statement are actually also true: For $0 < \alpha \leq 1$, $(\lambda_1,\ldots,\lambda_n) \overset{\text{p}}{\succeq} (\lambda_1^*,\ldots,\lambda_n^*)$ gives rise to the usual stochastic order between $X_{1:n}$ and $X_{1:n}^*$. The following example gives negative answer to this conjecture.

Example 2.3. Let (X_1, X_2) $((X_1^*, X_2^*))$ be a vector of independent heterogeneous GE random variables. Set $\alpha = 0.6$. For $(\lambda_1,\lambda_2) = (1, 5.5) \overset{\text{p}}{\succeq} (2, 3) = (\lambda_1^*,\lambda_2^*)$, $X_{1:2} \leq_{\text{st}} X_{1:2}^*$; however, for $(\lambda_1,\lambda_2) = (1, 2.25) \overset{\text{p}}{\succeq} (1.1, 2.14) = (\lambda_1^*,\lambda_2^*)$, $X_{1:2} \geq_{\text{st}} X_{1:2}^*$. So, $(\lambda_1,\lambda_2) \overset{\text{p}}{\succeq} (\lambda_1^*,\lambda_2^*)$ implies neither $X_{1:2} \leq_{\text{st}} X_{1:2}^*$ nor $X_{1:2} \geq_{\text{st}} X_{1:2}^*$ for $0 < \alpha \leq 1$.

In the next theorem, we compare parallel systems in the case when two sets of scale parameters weakly majorize each other that generalizes the corresponding result in Theorem 10 (ii) of [1]. In particular the majorization assumption is relaxed to the super-majorization.

Theorem 2.4. Let $X_1,\ldots,X_n$ be independent random variables with $X_i \sim GE(\alpha,\lambda_i)$ and $X_1^*,\ldots,X_n^*$ be another set of independent random variables with $X_i^* \sim GE(\alpha,\lambda_i^*), i = 1,...,n$. Then for any $\alpha > 0$,

$$(\lambda_1,\ldots,\lambda_n) \overset{\text{w}}{\succeq} (\lambda_1^*,\ldots,\lambda_n^*) \Longrightarrow X_{n:n} \geq_{\text{rh}} X_{n:n}^*. \tag{2.3}$$

Proof. Fix $x > 0$. Then the reverse hazard rate of $X_{n:n}$ is

$$\tilde{r}_{X_{n:n}}(x,\boldsymbol{\lambda}) = \sum_{i=1}^{n} \alpha\lambda_i \tilde{r}(\lambda_i x) = \frac{\alpha}{x}\sum_{i=1}^{n} \varphi(\lambda_i x),$$

where $\varphi(x) = x\tilde{r}(x)$, and $\tilde{r}(x) = \dfrac{e^{-x}}{1-e^{-x}}, x \geq 0$. From Theorem 3.A.8 of [8] it suffices to show that, for each $x > 0$, $\tilde{r}_{X_{n:n}}(x,\boldsymbol{\lambda})$ is Schur-convex and decreasing in λ_i's.

It is easy to see that, $\varphi(x)$ is decreasing and convex in x, then the reverse hazard rate function of $X_{n:n}$ is decreasing and convex in each λ_i.

Now, from Proposition 3.C.1 of [8], Schur-convexity of $\tilde{r}_{X_{n:n}}(x,\boldsymbol{\lambda})$ follows from convexity of $\varphi(x)$. This completes the proof of theorem. □

3 Conclusions

In this paper, we derived new results on stochastic comparisons of series and parallel systems with independent heterogeneous generalized exponential components. The result established here strengthens and generalizes some of the results of [1].

References

[1] Balakrishnan, N. , Haidari, A. and Masoumifard, K. (2015), Stochastic comparisons of series and parallel systems with generalized exponential components, IEEE Transactions on Reliability, 64, 333-348.

[2] Bon, J. L. and Păltănea, E. (1999), Ordering properties of convolutions of exponential random variables, *Lifetime Data Anal.* **5** 185-192.

[3] Fang, L. and Zhang, X. (2015), Stochastic comparisons of parallel systems with exponentiated weibull components, statistics & probability letters, 97, 25-31.

[4] Gupta, R. D., and Kundu, D. (1999), Generalized exponential distributions. Australian & New Zealand Journal of Statistics, 41(2), 173-188.

[5] Khaledi, B.-E., Farsinezhad, S. and Kochar, S.C. (2011), Stochastic comparisons of order statistics in the scale model, Journal of Statistical Planning and Inference, 141, 276-286.

[6] Kochar, S. C. and Torrado, N. (2015), On stochastic comparisons of largest order statistics in the scale model, Communications in Statistics - Theory and Methods, 44(19), 4132-4143.

[7] Kundu, A. and Chowdhury, S. (2016), Ordering properties of order statistics from heterogeneous exponentiated weibull models, statistics & probability letters, 114, 119-127.

[8] Marshall A.W., Olkin, I. and Arnold, B.C. (2011), Inequalities: Theory of Majorization and its Applications. second ed., Springer, New York .

[9] Mudholkar, G. S. and Srivastava, D. K. (1993), Exponentiated weibull family for analyzing bathtub failure rate data, IEEE Transactions on Reliability, 42, 299-302.

[10] Shaked, M. and Shanthikumar, J.G. (2007), Stochastic Orders, Springer, New York.

[11] Zhao, P. and Balakrishnan, N. (2014), A stochastic inequality for the largest order statistics from heterogeneous gamma variables, Journal of Multivariate Analysis, 129, 145-150.

[12] Zhao, P., Wang, L. and Zhang, Y. (2016), On extreme order statistics from heterogeneous beta distributions with applications, Communications in Statistics - Theory and Methods, just-accepted.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Optimal sample size based on prediction problem under progressively type-II censoring

Basiri, E. [1] and MirMostafaee, S.M.T.K. [2]

[1] Department of Statistics, Kosar University of Bojnord
[2] Department of Statistics, University of Mazandaran

Abstract

Choosing the sample size is a problem faced by anyone doing a survey of any type. In this paper, we study this problem by considering two criteria, the total cost of experiment and mean squared prediction error in prediction problem. Towards this end, we discuss the problem of Bayesian predicting a future progressive censored order statistic from an exponential distribution based on an observed progressive censored order statistics.

Keywords: Optimal sample size, Prediction, Progressive censoring.

1 Introduction

The scheme of progressive Type-II censoring is an important method of obtaining data in lifetime studies. Suppose n units are placed on a lifetime test. At the ith failure time, r_i surviving items are randomly withdrawn from the test, where $r_i = n - m - \sum_{j=0}^{i-1} r_j$, $i = 1, \cdots, m$, where $r_0 = 0$. Then the failure times $X_{1:m:n}, \cdots, X_{m:m:n}$ are called *progressively Type-II censored order statistics* (PCOs). In what follows we focus on finding optimal values for m based on a prediction problem for the exponential distribution under progressive Type-II censoring.

[1] elham_basiri2000@yahoo.com
[2] m.mirmostafaee@umz.ac.ir

2 Main results

Let $\tilde{x} = (x_{1:m_1:n_1}, \cdots, x_{m_1:m_1:n_1})$ be observed PCOs with censoring scheme $(r_1, \cdots, r_{m_1})$ from the one parameter exponential distribution with the probability density function (pdf)

$$f(x) = \theta e^{-\theta x}, \quad x > 0,\ \theta > 0. \tag{2.1}$$

Then, the likelihood function takes the form (see [3])

$$L(\theta, \tilde{x}) = C \prod_{i=1}^{m} (1 - F(x_i))^{r_i} f(x_i) = C\theta^{m_1} \exp(-\theta T),$$

where $C = \prod_{j=1}^{m}(n - \sum_{i=0}^{j-1} r_i - j + 1)$ and $T = \sum_{i=1}^{m_1}(1 + r_i)x_i$.

The conjugate prior for θ is considered as

$$\pi(\theta) = \frac{b^a}{\Gamma(a)}\theta^{a-1}e^{-b\theta}, \quad \theta > 0, \quad a, b > 0.$$

Therefore, the posterior distribution of θ is obtained to be

$$\pi(\theta|\tilde{x}) = \frac{(b+T)^{a+m_1}}{\Gamma(a+m_1)}\theta^{a+m_1-1}e^{-\theta(b+T)},$$

respectively, where $\Gamma(\cdot)$ is the complete gamma function.

Let $Y_{s:m_2:n_2}$ be the sth future PCO with censoring scheme $(r'_1, \cdots, r'_{m_2})$ from the same exponential distribution with the pdf given in (2.1). Then, the marginal pdf of $Y_{s:m_2:n_2}$, $(1 \le s \le m_2)$ is (see [3])

$$f_{Y_{s:m_2:n_2}}(y) = \theta c'_{s-1} \sum_{i=1}^{s} a'_{i,s} \exp\{-\theta\gamma'_i y\}, \quad y > 0,$$

where $\gamma'_i = n_2 - i + 1 - \sum_{j=1}^{i-1} r'_j$, $c'_{s-1} = \prod_{j=1}^{s} \gamma'_j$ and $a'_{i,s} = \prod_{j=1, j\neq i}^{s} \frac{1}{\gamma'_j - \gamma'_i}$, $1 \le i \le s \le m_2$. So, the predictive density function for $Y_{s:m_2:n_2}$ is

$$\begin{aligned} f^*_{Y_{s:m_2:n_2}}(y|\tilde{x}) &= c'_{s-1} \sum_{i=1}^{s} a'_{i,s} \frac{(b+T)^{a+m_1}}{\Gamma(a+m_1)} \int_0^\infty \theta^{a+m_1} e^{-\theta(b+T+\gamma'_i y)} \mathrm{d}\theta \\ &= c'_{s-1}(a+m_1)(b+T)^{a+m_1} \sum_{i=1}^{s} \frac{a'_{i,s}}{(b+T+\gamma'_i y)^{a+m_1+1}}. \end{aligned}$$

Define the functions g and h as

$$g(s, \tilde{r}_s, n) = \sum_{l=1}^{s} \frac{1}{n - \sum_{k=0}^{l-1} r_k - l + 1} \quad \text{and} \quad h(s, \tilde{r}_s, n) = \sum_{l=1}^{s} \frac{1}{\left(n - \sum_{k=0}^{l-1} r_k - l + 1\right)^2},$$

where $\tilde{r}_s = (r_0, r_1, \cdots, r_{s-1})$. Then, we can write (see [6])

$$E(Y_{s:m_2:n_2}) = \frac{1}{\theta} g(s, \tilde{r}'_s, n_2) \quad \text{and} \quad V(Y_{s:m_2:n_2}) = \frac{1}{\theta^2} h(s, \tilde{r}'_s, n_2),$$

where $\tilde{r}'_s = (r'_0, r'_1, \cdots, r'_{s-1})$ and $r'_0 = 0$.

The point predictor for $Y_{s:m_2:n_2}$, $1 \le s \le m_2$, under squared error loss (SEL) function is

$$\begin{aligned} \widehat{Y}_{s:m_2:n_2} &= c'_{s-1} \sum_{i=1}^{s} a'_{i,s} \frac{(b+T)^{a+m_1}}{\Gamma(a+m_1)} \int_0^\infty \int_0^\infty y \theta^{a+m_1} e^{-\theta(b+T+\gamma'_i y)} \mathrm{d}y \mathrm{d}\theta \\ &= \frac{(b+T)}{(a+m_1-1)} c'_{s-1} \sum_{i=1}^{s} \frac{a'_{i,s}}{\gamma'^2_i} = \frac{(b+T)}{(a+m_1-1)} g(s, \tilde{r}'_s, n_2), \end{aligned}$$

where the last equality is obtained by using the following identity

$$c'_{s-1} \sum_{i=1}^{s} \frac{a'_{i,s}}{\gamma'^2_i} = g(s, \tilde{r}'_s, n_2).$$

Since in many real applications, no prior knowledge is available about θ, we may take $a = b = 0$, i.e. the non-informative prior for θ. Therefore

$$E(\widehat{Y}_{s:m_2:n_2}) = \frac{1}{\theta} \frac{m_1}{m_1 - 1} g(s, \tilde{r}'_s, n_2) \quad \text{and} \quad V(\widehat{Y}_{s:m_2:n_2}) = \frac{1}{\theta^2} \frac{m_1}{(m_1-1)^2} g^2(s, \tilde{r}'_s, n_2).$$

We obtain the mean squared prediction error (MSPE) of $\widehat{Y}_{s:m_2:n_2}$ as

$$\mathrm{MSPE}(\widehat{Y}_{s:m_2:n_2}) = \frac{1}{\theta^2} \left\{ h(s, \tilde{r}'_s, n_2) + g^2(s, \tilde{r}'_s, n_2) \frac{m_1 + 1}{(m_1 - 1)^2} \right\}. \tag{2.2}$$

Another criterion considered in this paper is the total cost of test, which plays an important role in practice. The total cost associated with the information sample $\tilde{x} = (x_{1:m_1:n_1}, \cdots, x_{m_1:m_1:n_1})$, is given by

$$TC = c_0 + c_t X_{m_1:m_1:n_1} + c_u n_1,$$

where c_0, c_t and c_u are the sampling set-up cost or any other related cost involved in sampling, cost of total time on test and cost per unit, respectively. We consider the expected of cost function which is

$$\begin{aligned} E(TC) &= c_0 + c_t E(X_{m_1:m_1:n_1}) + c_u n_1 \\ &= c_0 + \frac{c_t}{\theta} g(m_1, \tilde{r}_{m_1}, n_1) + c_u n_1, \end{aligned} \tag{2.3}$$

which depends on the unknown parameter θ. One may replace θ with its preliminary estimate obtained based on past experiments.

We try to find optimal values for m_1 such that $\mathrm{MSPE}(\widehat{Y}_{s:m_2:n_2}) \le \kappa$ and $E(TC) \le c$, where κ and c are pre-fixed values. From (2.2), $\mathrm{MSPE}(\widehat{Y}_{s:m_2:n_2}) \le \kappa$, if and only if $\frac{m_1+1}{(m_1-1)^2} \le \frac{\theta^2 \kappa - h(s, \tilde{r}'_s, n_2)}{g^2(s, \tilde{r}'_s, n_2)}$.

Moreover, by using (2.3), $E(TC) \leq c$ is equivalent to $g(m_1, \tilde{r}_{m_1}, n_1) \leq \frac{(c-c_0-n_1c_u)\theta}{c_t}$. Note that we expect $g(m_1, \tilde{r}_{m_1}, n_1)$ to be an increasing function of m_1 when n_1 is kept fixed regardless of the set $\tilde{r}_{m_1}$. If so, then assuming $\theta^2\kappa > h(s, \tilde{r}'_s, n_2)$, we readily find that the optimal value for m_1, say $m_{1_{opt}}$, satisfies the following inequality

$$\max\{2, m_0\} \leq m_{1_{opt}} \leq m'_1, \tag{2.4}$$

where m'_1 is the largest integer value such that

$$g(m'_1, \tilde{r}_{m'_1}, n_1) \leq \frac{(c - c_0 - n_1c_u)\theta}{c_t},$$

and m_0 is the smallest integer value such that

$$m_0 \geq 1 + \frac{1 + \sqrt{8T^* + 1}}{2T^*} \quad \text{and} \quad T^* = \frac{\theta^2\kappa - h(s, \tilde{r}'_s, n_2)}{g^2(s, \tilde{r}'_s, n_2)}.$$

From the relation (2.4), we see that the upper bounds for $m_{1_{opt}}$ can be determined based on the total cost of time and the lower bounds can be obtained based on the MSPE. Let us consider the censoring schemes as $(r_1, r_2, \cdots, r_{m_j}) = (0, 0, \cdots, n_j - m_j)$ for $j = 1, 2$, then Table 1 consists of values of $m_{1_{opt}}$ for some selected values of s, c and κ when $m_2 = 10$, $n_1 = n_2 = 20$, $c_0 = c_t = 1$, $c_u = 0.2$ and $\theta = 1$. The computations were obtained by the codes written in R [3].

Table 1: Values of $m_{1_{opt}}$ for some selected values of s, c and κ when $m_2 = 10$, $n_1 = n_2 = 20$, $c_0 = c_t = 1$, $c_u = 0.2$ and $\theta = 1$.

		s	1	3	5	8	10
c	κ						
5.5	0.1		$\{2, \cdots, 8\}$	$\{2, \cdots, 8\}$	$\{3, \cdots, 8\}$	$\{6, 7, 8\}$	–
	0.2		$\{2, \cdots, 8\}$	$\{2, \cdots, 8\}$	$\{3, \cdots, 8\}$	$\{4, \cdots, 8\}$	$\{6, 7, 8\}$
	0.3		$\{2, \cdots, 8\}$	$\{2, \cdots, 8\}$	$\{2, \cdots, 8\}$	$\{3, \cdots, 8\}$	$\{4, \cdots, 8\}$
6	0.1		$\{2, \cdots, 12\}$	$\{2, \cdots, 12\}$	$\{3, \cdots, 12\}$	$\{6, \cdots, 12\}$	$\{12\}$
	0.2		$\{2, \cdots, 12\}$	$\{2, \cdots, 12\}$	$\{3, \cdots, 12\}$	$\{4, \cdots, 12\}$	$\{6, \cdots, 12\}$
	0.3		$\{2, \cdots, 12\}$	$\{2, \cdots, 12\}$	$\{2, \cdots, 12\}$	$\{3, \cdots, 12\}$	$\{4, \cdots, 12\}$
7	0.1		$\{2, \cdots, 17\}$	$\{2, \cdots, 17\}$	$\{3, \cdots, 17\}$	$\{6, \cdots, 17\}$	$\{12, \cdots, 17\}$
	0.2		$\{2, \cdots, 17\}$	$\{2, \cdots, 17\}$	$\{3, \cdots, 17\}$	$\{4, \cdots, 17\}$	$\{6, \cdots, 17\}$
	0.3		$\{2, \cdots, 17\}$	$\{2, \cdots, 17\}$	$\{2, \cdots, 17\}$	$\{3, \cdots, 17\}$	$\{4, \cdots, 17\}$

In Table 1, dash (-) means that there is no $m_{1_{opt}}$ which satisfies condition (2.4). Because in this case the lower bound in (2.4) is greater than the corresponding upper one. From Table 1, we can state the following findings:

- As one would expect, the lower bounds for $m_{1_{opt}}$ increases as s increases when all other components are held fixed.
- For fixed values of s and κ, the maximum value of $m_{1_{opt}}$ is an increasing function of c. Because enhancing the total cost of test leads to a greater value for $m_{1_{opt}}$.
- The MSPE is a decreasing function of $m_{1_{opt}}$. So, for fixed values of s and c, the minimum value of $m_{1_{opt}}$ is a decreasing function of κ.

- When c and κ are kept fixed, then the minimum values of $m_{1_{opt}}$ for the lower order statistics are smaller than those for the upper order statistics but they have the same maximum value of $m_{1_{opt}}$.

From Table 1, we can see that $m_{1_{opt}}$ is not unique. In practice, the engineer may request only one value of m_1. So, this question arises for a practitioner that which $m_{1_{opt}}$ must be chosen in a practical situation. We respond that it depends on which criterion is more important for that user. If the cost of an experiment is more important than the MSPE, then the lower bound for $m_{1_{opt}}$ can be considered. On the other hand, as expected intuitively, if we consider the MSPE as the most important criterion, then we must prefer the upper bound for $m_{1_{opt}}$.

References

[1] Balakrishnan, N. and Aggarwala, R. (2000), *Progressive Censoring: Theory, Methods, and Applications*, Birkhäuser, Boston.

[2] Balakrishnan, N. (2007), Progressive censoring methodology: An appraisal. *Test*, **16**, 211–259.

[3] R Development Core Team (2016), R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Some measures of income inequality and reliability of systems

Behdani, Z. [1] Mohtashami Borzadaran, G, R. [2]

and Sadeghpour Gildeh, B. [3]

[1,2,3] Department of Statistics, Ferdowsi University of Mashhad

Abstract

Reliability engineers often need to work with systems having elements connected in parallel and series, and to calculate their reliability. The concepts of inequality measures play an important role in economic, social sciences and other areas. In this paper, we study the relationship between a mean life time of system consisting of n identical and independents having parallel or series structure and some measures of income inequality. The Gini index and generalized Gini index are two particular member of general class of dispersion measures. It is natural to try to relate Gini indices to mean life time of system.

Keywords: Mean residual life function, Reliability, Lorenz curve, Gini index.

1 Introduction

Recently, reliability theoretic ideas and methods have been used successfully in several other areas of investigation with a view towards exploiting concepts and tools, such as demography, queueing theory and economics. Reliability engineers often need to work with systems having elements connected in parallel and series, and to calculate their reliability. In recent years, various inequality curves have been developed or investigated as the descriptors of income inequality. The Bonferroni

[1]zbehdani@yahoo.com
[2]grmohtashami@um.ac.ir
[3]sadeghpour@um.ac.ir

curve and the Zenga-2007 curve appear to be essentially the functions of the Lorenz curve. Arnold [2] showed, they each determine the parent distribution up to scale factor, and they each, yield an inequality partial order that is equivalent to the Lorenz order. The Lorenz curve is considered as a very useful tool of economic due to its significant role in the measurement of the inequality of income distributions and wealth. Several authors have discussed relationships between income and wealth inequality measures and some central notions used in reliability theory and survival analysis. Chandra and Singpurwalla [7] illustrated the connection between Lorenz curves and Gini index in economics and total time on test and mean residual life in reliability. Klefsjo [11] derived more results along the same lines and presented reliability interpretations of some more concepts from economics. Giorgi and Crescenzi [8] proved some relationships between the total time on test and Bonferroni curve and showed how the Bonferroni curve may be applied in reliability theory. Singpurwalla [16] noted the relationship between the survival function in reliability and the asset pricing formula of fixed income instrument such as a risk-free zero coupon bond.
The aim of this article is to answer the question What is the relationship between the lifetime of a system and inequality indices?

2 Definitions and notations

Throughout this paper, X and Y are two non-negative continuous random variables with finite and positive means. We consider F and G for the distribution functions and use the symbols f and g to denote respective probability (density) distributions, if they exist. The survival function of F is denoted by $\overline{F} = 1 - F$ and the quantile function will be denoted by $F^{-1}(u) = \inf\{x : F(x) \geq u,\, u \in (0,1)\}$, where $F^{-1}(0)$ and $F^{-1}(1)$ are the lower and upper bounds of the support of F (S_F) respectively, for G analogously. Here $m_F(x) = \int_x^\infty \overline{F}(t)dt/\overline{F}(x), x \geq 0$ stands for the mean residual life function of F. All other functions are notated by similar notations. Let us give some definitions that will be used in this article.
A system is a collection of components, subsystems and/assemblies arranged to a specific design in order to achieve the desired functions with acceptable performance and reliability. The types of components, their qualities, their quantities and the way in which they are set inside the system have direct effect on system reliability. Coherent systems are often considered in reliability theory to describe the structure and the performance of complex systems. For example, k-out-of-n systems, and series and parallel systems in particular, are coherent systems. Series and parallel structures are the basis for building more complicated structures which use redundancy to increase system reliability.

Definition 1. *A Series system is a configuration such that, if any one of the system components fails, the entire system fails i.e. the series system fails as long as any one of the components fails. Some applications of series system are in some computer networks, chains, multi-cell batteries and decorative tree-lights. When components fail independently the system reliability as*

$$\overline{F}_s(t) = \overline{F}_1(t)\overline{F}_2(t)\cdots\overline{F}_n(t) = \prod_{i=1}^{n}\overline{F}_i(t) \tag{2.1}$$

Its clear that the system reliability can not be greater than the smallest component reliability. It is important for all components to have a high reliability, especially when the system has a large number of components. If the times-to-failure of the components behave according to the exponential p.d.f, then the overall p.d.f of times-to-failure is also exponential. The mean residual life function of a series system is defined by the conditional expectation of residual life length

$$\varphi_n(t) = E(X_{1:n} - t | X_{1:n} > t), \tag{2.2}$$

Definition 2. *A Parallel system is a system in which components are connected in parallel and the system does not fail, even one component is in good working condition i.e the system fails only when all components have failed. A parallel system works as long as any one of the components works. Conceptually, in a parallel configuration the total system reliability is higher than the reliability of any single system component. Jet engines, Braking systems, Tires in trucks and Projector light bulbs are applications of parallel system. For parallel systems, it is easier to work with failure probability than reliability because*

$$F_s(t) = \prod_{i=1}^{n} F_i(t) \tag{2.3}$$

and therefore

$$\overline{F}_s(t) = 1 - \prod_{i=1}^{n} 1 - \overline{F}_i(t) \tag{2.4}$$

The system reliability can not be less than the largest component reliability. The conditional expectation of residual life length of the parallel system is called the mean residual life function of parallel system

$$\psi_n(t) = E(X_{n:n} - t | X_{1:n} > t), \tag{2.5}$$

Bairamov et al. [3] showed $\psi_n(t) = \frac{n}{\overline{F}^n(t)} \int_t^\infty y[F(y) - F(t)]^{n-1} f(y)dy - t$ *and they obtained*

$$\overline{F}(x) = exp\{\frac{-1}{n} \int_0^x \frac{\psi_n'(t) + 1}{\psi_n(t) - \psi_{n-1}(t)} dt\}.$$

Lorenz curve provides a useful graphical method to analyse income inequality for about one hundred years since they were designed. The Gini coefficient has been found useful to study the inequality of incomes. The value of Gini coefficient reveals the degree of income inequality. The Lorenz curve was first defined by Lorenz (1905).

Definition 3. *Let X be a non-negative random variable with finite and positive mean, the Lorenz curve of X is given by*

$$L_X(p) = \frac{1}{E(X)} \int_0^p F^{-1}(u)du, \quad 0 \le p \le 1. \tag{2.6}$$

It should be mentioned that L_X has the following properties:

- Lorenz curve is a distribution function, twice differentiable, convex (and hence also star-shaped), increasing and is continuous.

- $L(0) = 0$ and $L(1) = 1$ on $[0; 1]$.

- $lim_{p\to 1} L'(p)(1-p) = 0$, $L_X(p) \leq p$.

- Moreover function possessing these properties is the Lorenz curve of a certain statistical distribution.

The most famous measure of income inequality is the Gini coefficient. It is certainly a relative measure of income inequality, and is given as twice the area between the equality line and Lorenz curve:

$$G = 2\int_0^1 (p - L(p))dp = 1 - 2\int_0^1 L(p)dp.$$

Kakwani [10] proposed a one-parameter family of generalized Gini indices by introducing different weighting functions for the area under the Lorenz curve as,

$$G_n = 1 - n(n-1)\int_0^1 L(p)(1-p)^{n-2}dp, \quad n \geq 1,$$

can be written as

$$G_n = 1 - \frac{1}{\mu}\int_0^\infty \overline{F}^n(x)dx = 1 - \frac{E(X_{1:n})}{E(X)}. \tag{2.7}$$

The Gini coefficient is obtained for $n = 2$.
It may also be of interest to note that the piesch measure

$$P = \int_0^1 3t(t - L(t))dt \tag{2.8}$$

The Bonferroni curve $B_X(p)$ is a relatively minor modification of the Lorenz curve. It is defined by:

$$B_X(p) = \frac{L(p)}{p}, 0 < p < 1. \tag{2.9}$$

Consequently, we cannot say in general that Bonferroni curve starts from the original of the orthogonal plane, as it depends on the definition of X [8]. The Bonferroni curve is strictly increasing and could be convex in some parts and concave in the others.
One of the another measure of income inequality is Zenga index. The Zenga curve $Z(p)$ is the ratio of the mean income of the poorest $100p$ in the distribution to that of the rest of the distribution, namely the $100(1-p)$ richest. It is defined by:

$$Z_X(p) = 1 - \frac{L(p)}{p} \cdot \frac{1-p}{1-L(p)} \quad p \in (0, 1).$$

The Zenga index, can be written as:

$$Z = \int_0^1 Z(p)dp.$$

Inequality partial orders are defined for each of the four inequality curves.
We say that X is smaller than Y in the Lorenz order ($X \leq_L Y$), Bonferroni order ($X \leq_B Y$) or Zenga order ($X \leq_Z Y$) iff $L_Y(p) \leq L_X(p)$, $B_Y(p) \leq B_X(p)$ or $Z_Y(p) \leq Z_X(p)$ respectively. These orders are invariant with respect to scale transformation. The Lorenz order is the natural mathematical abstraction of Lorenz's (1905) comparison of income distributions via nested Lorenz curves [1]. From definitions of the Zenga and Bonferroni curves and definitions of the Zenga, Bonferroni and Lorenz orders immediately conclude that

$$X \leq_B Y \Longleftrightarrow X \leq_L Y \Longleftrightarrow X \leq_Z Y$$

3 Main results

In here, we present some of the important results about the relationship between mean life of a system and some measure of income inequality. As our first main result, the following theorems exhibit some new connections between mean residual life function of a parallel system and Lorenz order.

Theorem 3.1. *Let $\psi(t)$ be the mean residual life function of a parallel system consisting of 2 identical and independent components with continuous life distribution function F. Then the following identity holds*

$$\overline{L}(F(t)) = \frac{1}{\mu}\left(\overline{F}(t)(\psi(t)+t) - \frac{[\psi'(t)+1](\overline{F}(t))^2}{2f(t)}\right)$$

Proof. For simplest case $n = 2$ we have the equation

$$\psi(t) = \psi_2(t) = \frac{2\int_0^\infty x[F(x) - F(t)]f(x)dx}{\overline{F}^2(t)} - t,$$

differentiating with respect to t we have

$$\begin{aligned}[\psi'(t)+1]\overline{F}^2(t) - 2f(t)\overline{F}(t)(\psi(t)+t)\\ = -2f(t)\int_t^\infty xf(x)dx,\end{aligned}$$

after some derivation we have

$$\overline{L}(F(t)) = \frac{1}{\mu}\left(\overline{F}(t)(\psi(t)+t) - \frac{[\psi'(t)+1](\overline{F}(t))^2}{2f(t)}\right)$$

□

Remark 3.2. The integral in equation 2.7 can be interpreted as the mean life of a series system consisting of n independent components each of which has a time to failure distributed according

to F [11]. The mean life of a parallel system consisting of n independent component each of which has a time to failure with distribution function F can be denoted as

$$\int_0^\infty 1 - F^n(x)dx.$$

Which has closely related to the general family of the piesch measure in equation 2.8

The theorem below shows that, if $X \leq_L Y$, then a parallel system of n components having independent lifetimes which are copies of Y has a larger lifetime, in the sense of Lorenz order, than a similar system of n components having independent lifetime which are copies of X.

Theorem 3.3. *Let $X_1, X_2, \cdots, X_n$ be a collection of independent and identically distributed random variables, and let $Y_1, Y_2, \cdots, Y_n$ be another collection of i.i.d. random variables and these random variables have a common mean. If $X_1 \leq_L Y_1$, then $X_{n:n} \leq_L Y_{n:n}$ for $n \geq 1$.*

Proof. Assume that $X_1 \leq_L Y_1$. It suffices to consider only the case $n = 2$ for other cases are similar. Let the survival functions of X_1 and Y_1 denote by $\overline{F}$ and $\overline{G}$ respectively, now 2.4 that $\overline{F}_{max\{X_1,X_2\}}(t) = 1 - F^2(t)$ and $\overline{G}_{max\{Y_1,Y_2\}}(t) = 1 - G^2(t)$. Now, from the assumed $X_1 \leq_L Y_1$ it follows that

$$\int_0^u (F^{-1}(t) - G^{-1}(t))dt \geq 0,$$

since $F^{-1}_{max\{X_1,X_2\}}(p) = F^{-1}(\sqrt{p})$ and $G^{-1}_{max\{Y_1,Y_2\}}(p) = G^{-1}(\sqrt{p})$ for $p \in (0,1)$, it follows that

$$\int_0^u F^{-1}_{max\{X_1,X_2\}}(x)dx \geq \int_0^u G^{-1}_{max\{Y_1,Y_2\}}(x)dx,$$

that is, $X_{n:n} \leq_L Y_{n:n}$. □

Theorem 3.4. *Let $X_1, X_2, \cdots, X_n$ be a collection of independent and identically distributed random variables, and let $Y_1, Y_2, \cdots, Y_n$ be another collevtion of i.i.d. random variables and these random variables have a common mean. If $X_1 \leq_Z Y_1$, then $X_{n:n} \leq_Z Y_{n:n}$ for $n \geq 1$.*

Proof.

$$\begin{aligned}
X_1 \leq_Z Y_1 &\Longrightarrow Z_{X_1}(p) \leq Z_{Y_1}(p), \\
&\Longrightarrow \frac{\int_p^1 F^{-1}(t)dt - (1-p)E(X)}{p\int_p^1 F^{-1}(t)dt} \leq \frac{\int_p^1 G^{-1}(t)dt - (1-p)E(Y)}{p\int_p^1 G^{-1}(t)dt}, \\
&\Longrightarrow E(Y)\int_p^1 F^{-1}(t)dt \leq E(X)\int_p^1 G^{-1}(t)dt, \\
&\Longrightarrow \int_p^1 F^{-1}(\sqrt{t})dt \leq \int_p^1 G^{-1}(\sqrt{t})dt, \\
&\Longrightarrow \frac{\int_p^1 F^{-1}(\sqrt{t})dt - (1-p)E(X_{n:n})}{p\int_p^1 F^{-1}(\sqrt{t})dt} \leq \frac{\int_p^1 G^{-1}(\sqrt{t})dt - (1-p)E(Y_{n:n})}{p\int_p^1 G^{-1}(\sqrt{t})dt}, \\
&\Longrightarrow Z_{X_{n:n}}(p) \leq Z_{Y_{n:n}}(p), \\
&\Longrightarrow X_{n:n} \leq_Z Y_{n:n}.
\end{aligned}$$

□

Corollary 3.5. *Recall that, a random variable X with mean μ is HNBUE, iff*

$$X \leq_L Y \tag{3.1}$$

where Y denotes an exponential random variable with mean μ. Now consider a parallel system of n components having i.i.d. HNBUE lifetimes $X_1, X_2, \cdots, X_n$ with the common mean μ. Let $Y_1, Y_2, \cdots, Y_n$ be i.i.d. exponential random variables with mean μ. From Theorem 3.3 we obtain $X_{n:n} \leq_L Y_{n:n}$ and consequently the lifetime of parallel system is also HNBUE. As $X_{n:n} \leq_L Y_{n:n}$ we have $cv(X_{n:n}) \leq cv(Y_{n:n})$. Kochar [12] obtained upper bounds on the mean and on the variance of the lifetime of the parallel system then

$$cv(X_{n:n}) \leq \frac{\sqrt{\left\{2\sum_{k=1}^{n}\frac{(-1)^{k+1}}{k^2}\binom{n}{k} - \left(\sum_{k=1}^{n}\frac{1}{k}\right)^2\right\}}}{\sum_{k=1}^{n}\frac{1}{k}}$$

References

[1] Arnold, B.C. (1991). Preservation and attenuation of inequality as measured by Lorenz order. *Stochastic Orders and Decision under Risk, IMS Lectures Notes - Monograph* Series, 25-37.

[2] Arnold, B. C. (2015). On Zenga and Bonferroni curves. *Metron*, **73(1)**, 25-30.

[3] Bairamov, I., Ahsanullah, M. and Akhundov, I. (2002). A residual life function of a system having parallel or series structures. *Journal of Statistical Theory and Applications*, **1(2)**, 119-132.

[4] Banzhaf, J. F. (1965). Weighted voting doesnt work: A mathematical analysis. *Rutgers Law Review*, **19**, 317-343.

[5] Barlow, R. E., Proschan, F. (1975). Importance of system components and fault tree event. *Stochastic Processes and Their Applications*, **3**, 153-172.

[6] Birnbaum, Z.W. (1969). *On the importance of different components in a multicomponent system*, In: Multivariate Analysis II, Ed. P.R. Krishnaiah, New York: Academic Press, pp. 581-592.

[7] Chandra, M., Singpurwalla, N. D. (1981). Relationships between some notions which are common to reliability theory and economics. *Mathematics of Operations Research*, **6**, 113-121.

[8] Giorgi G.M., Crescenzi M. (2001) A look at the Bonferroni inequality measure in a reliability framework. *Statistica LXL*, **4**, 571583.

[9] Freixas, J., Puente, M. A. (2002). Reliability importance measures of the components in a system based on semivalues and probabilistic values. *Annals of Operations Research*, **109**:331-342.

[10] Kakwani, N. (1980). On a class of poverty measures. *Econometrica*, **48**, 437-446.

[11] Klefsjo, B. (1984). Reliability interpretations of some concepts from economics. *Naval Research Logistics Quarterly*, **31**:301-308.

[12] Kochar, S. C., Li, X. a nd Shaked, M. (2002). The total time on test transform and the excess wealth stochastic orders of distributions. *Advances in Applied Probability*, **34(04)**, 826-845.

[13] Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association*, **9(70)**, 209-219.

[14] Lynn, N. J. (2004). The price of failure. *In: Mathematical Reliability: An Expository Perspective*, Eds. R. Soyer, T. A. Mazzuchi, N. D. Singpurwalla. Norwell, MA: Kluwer, pp. 303-316.

[15] Shapley, L. S., Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American Political Science Review*, **48**:787-792.

[16] Singpurwalla, N. D. (2007). Reliability and survival in financial risk. *In: Advances in Statistical Modeling and Inference Essays in Honor of Kjell Doksum*, Ed. V. J. Nair. Singapore: World Scientific, pp. 93-114.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Some results on the residual entropy of coherent systems

Chahkandi, M. [1] and Toomaj, A. [2]

[1] Department of Statistics, University of Birjand
[2] Department of Statistics, Gonbad Kavous University

Abstract

The residual entropy measures the concentration of the conditional probability distribution and the uncertainty of the residual lifetime of a random phenomenon. In this paper, we find some expressions for the entropy of residual lifetime of a coherent system under different conditions on the status of the components. The obtained results can be applied to compare the predictability of the system lifetimes, when we have some information about the component lifetimes of the system at time t.

Keywords: Coherent system, Residual entropy, Signature vector.

1 Introduction

The concept of *Shannon entropy*, introduced by Shannon [1], is used for measuring of the uncertainty associated with its probability in behavior of a random variable. In fact, it is used to measure the predictability of a random phenomenon. Let X be a non-negative absolutely continuous random variable with probability density function (pdf) f. The Shannon entropy is defined as

$$H(X) = -\int_0^\infty f(x)\log(f(x))dx, \tag{1.1}$$

[1] mchahkandi@birjand.ac.ir
[2] ab.toomaj@gonbad.ac.ir

where $\log(\cdot)$ stands for the natural logarithm. This measure might be of interest to engineers and system operators for decision making, because a unit with great uncertainty is less reliable than a unit with low uncertainty.

Frequently, in survival analysis and life testing, one has information about the current age of the system. In such cases, the age must be taken into account to measure uncertainty. Obviously, the measure $H(X)$ in (1.1) is unsuitable in such situations and must be modified to take the age into account. Ebrahimi [2] proposed the residual entropy of the random variable X as follows:

$$H(X;t) = -\int_t^\infty \frac{f(x)}{\bar{F}(t)} \log \frac{f(x)}{\bar{F}(t)} dx. \tag{1.2}$$

Note that $H(X;t)$ measures the concentration of the conditional probability distribution i.e. the distribution of $X_t = (X - t|X > t)$.

In this paper, we consider a coherent system with lifetime T which is working at time t and we have some information about the lifetimes of its components. We consider three cases of conditional lifetimes and then we obtain explicit expressions for the residual entropy of the system in each cases.

2 Main results

Consider a coherent system with lifetime T, consisting of n i.i.d. component lifetimes $X_1, \ldots, X_n$ having a common cumulative distribution function (cdf) F. Also, suppose that $X_{1:n}, \ldots, X_{n:n}$ denote the associated order statistics of the component lifetimes. Let us assume that the system has the property that, with probability one, it is alive as long as at least $n-s+1$ $(s \leq n)$ components are alive. Such systems must have a signature vector of the form

$$\mathbf{p} = (0, 0, \ldots, 0, p_s, p_{s+1}, \ldots, p_n). \tag{2.1}$$

More details on the system signature is given in [3]. Some examples of coherent systems of order 4 with signature vector of the form (2.1) are given by:

$$\begin{aligned} T_1 &= \max(X1, \min(X2, X3, X4)), \\ T_2 &= \max(X1, \min(X2, X3), \min(X3, X4)), \\ T_3 &= \max(X_{2:3}, X_4), \\ T_4 &= \max(X_1, X_2, X_3, X_4), \end{aligned}$$

where their respective signature vectors are $\mathbf{p}_1 = (0, \frac{1}{2}, \frac{1}{4}, \frac{1}{4})$, $\mathbf{p}_2 = (0, \frac{1}{6}, \frac{7}{12}, \frac{1}{4})$, $\mathbf{p}_3 = (0, 0, \frac{3}{4}, \frac{1}{4})$ and $\mathbf{p}_4 = (0, 0, 0, 1)$.

Now, consider the situation that at time t, the system is working and we have some information about the number of failed components in the system. We are interested in studying the measure of uncertainty of the system residual lifetime. The following cases are considered.

Case I: Suppose that at time t the system is alive and we know that exactly ℓ $(\ell \leq s-1)$ components of the system are failed. The residual lifetime of this system may be presented as

$$T_{\ell,\ell+1}^t = (T - t|X_{\ell:n} \leq t < X_{\ell+1:n}),\ \ell = 1, 2, \ldots, s-1. \tag{2.2}$$

The survival function of $T^t_{\ell,\ell+1}$ can be found as

$$\begin{aligned}\bar{F}_{T^t_{\ell,\ell+1}}(x) &= \Pr(T > t+x | X_{\ell:n} \le t < X_{\ell+1:n}) \\ &= \sum_{k=s}^{n} p_k \Pr(X_{k:n} > t+x | X_{\ell:n} \le t < X_{\ell+1:n}) \\ &= \sum_{k=s}^{n} p_k \Pr(X_{k-\ell:n-\ell} > x+t | X_{1:n-\ell} > t), \end{aligned} \tag{2.3}$$

where the last equality follows from Remark 2.2 of Goliforushani and et al. [4]. In the next theorem, we find an expression for the entropy of $T^t_{\ell,\ell+1}$.

Theorem 2.1. *Let $T^t_{\ell,\ell+1}$ for $\ell = 1, 2, \ldots, s-1$ be the residual lifetime of a system with signature vector of the form (2.1). The entropy of $T^t_{\ell,\ell+1}$ can be found as*

$$H(T^t_{\ell,\ell+1}) = H(V_{\ell,\ell+1}) - \sum_{k=s}^{n} p_k E(\log(f_t(\bar{F}_t^{-1}(W_{k-\ell:n-\ell})))), \tag{2.4}$$

where $g_{V_{\ell,\ell+1}}(x) = \sum_{k=s}^{n} p_k g_{W_{k-\ell:n-\ell}}(x)$, and $W_{k-\ell:n-\ell}$ has a beta distribution with parameters $n-k+1$ and $k-\ell$.

Proof. The results can be found by implementing probability integral transformation $V_{\ell,\ell+1} = \bar{F}_t(T^t_{\ell,\ell+1})$, where $\bar{F}_t(x) = \frac{\bar{F}(t+x)}{\bar{F}(t)}$ and the entropy transformation formula given by Ebrahimi et al. [5]. □

Since entropy is a concave function of the density function, we can find a lower bound for the entropy of $T^t_{\ell,\ell+1}$ as

$$H(T^t_{\ell,\ell+1}) \ge \sum_{k=s}^{n} p_k H(X^t_{1,k-\ell,n-\ell}) = H_L(T^t_{\ell,\ell+1}), \tag{2.5}$$

where $X^t_{1,i,n} = (X_{i:n} - t | X_{1:n} > t)$. Using the results of Chahkandi and Toomaj [6], we have the next result.

Lemma 2.2. *If F is DFR (IFR), then $H(T^t_{\ell,\ell+1})$ is an increasing (decreasing) function of t, for a given ℓ and n.*

The next example is given to clarify the results of Theorem 2.1 and Lemma 2.2.

Example 2.3. Suppose that X follows a Weibull distribution with the survival function

$$\bar{F}(t) = e^{-t^\alpha}, \ t > 0, \ \alpha > 0. \tag{2.6}$$

Let us consider a coherent system with system signature $\mathbf{p} = (0, 0, 0.75, 0.25)$. It is easy to see that $H(V_{1,2}) = -0.8075$. It is not difficult to verify that

$$E(\log f_t(\bar{F}_t^{-1}(W_{i:n}))) = \log\alpha + E[\log W_{i:n}] + \frac{\alpha-1}{\alpha} E[\log(t^\alpha - \log W_{i:n})],$$

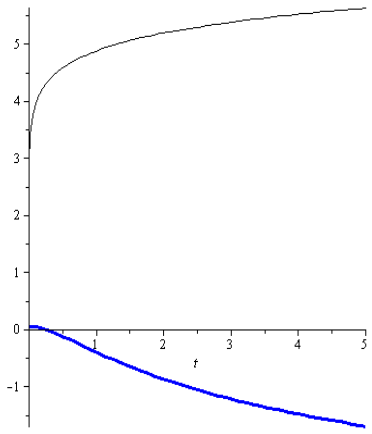


Figure 1: Exact values of $H(T_{1,2}^t)$ for $\alpha = 0.2$ (black line) and $\alpha = 2$ (blue line).

and hence

$$H(T_{1,2}^t) = H(V_{1,2}) - \log\alpha - \sum_{k=s}^{n} p_k E[\log W_{k-1:n-1}] - \frac{\alpha-1}{\alpha}\sum_{k=s}^{n} p_k E[\log(t^\alpha - \log W_{k-1:n-1})].$$

It is known that $E[\log W_{i:n}] = \psi(i) - \psi(n+1)$, where $\psi(x) = \frac{d\Gamma(x)}{dx}$. It is hard to obtain neat analytical results for $E[\log(t^\alpha - \log W_i)]$ and therefore we are forced to proceed via numerical computations. In Figure 1, we plotted the exact values of $H(T_{1,2}^t)$ with respect to time t for $\alpha = 0.2$ and $\alpha = 2$, respectively. It is well-known that when $0 < \alpha < 1$, then X is DFR and for $\alpha > 1$, X is IFR.

Case II) Now, consider the situation that we know at time t the system is working and the minimum and maximum number of failed components are i and j, respectively, where $1 \le i < j \le s$. Then the residual lifetime of the system is

$$T_{i,j}^t = (T - t | X_{i:n} \le t < X_{j:n}) \; ; \qquad 1 \le i < j \le s.$$

Theorem 2.4. *Under the condition that at time t at least i and at most j components of the system are failed, where $1 \le i < j \le s$, the entropy of the system residual lifetime is given by*

$$H(T_{i,j}^t) = H(V_{i,j}^t) - \sum_{\ell=i}^{j-1} b_{\ell,i,j}(t) E(\log(f_t(\bar{F}^{-1}(V_{\ell,\ell+1})))),$$

where $g_{V_{i,j}^t}(x) = \sum_{\ell=i}^{j-1} b_{\ell,i,j}(t) g_{V_{\ell,\ell+1}}(x)$, $b_{\ell,i,j}(t) = \frac{\binom{n}{\ell}(\phi(t))^\ell}{\sum_{m=i}^{j-1}\binom{n}{m}(\phi(t))^m}$, $1 \le i \le \ell < j \le n$ *and* $\phi(t) = \frac{F(t)}{\bar{F}(t)}$ *for* $t > 0$.

The following lemma can be found immediately by implementing the concave property of entropy.

Lemma 2.5. *$H(T_{i,j}^t)$ and $H(T_{\ell,\ell+1}^t)$ satisfy in the following inequality*

$$H(T_{i,j}^t) \geq H_L(T_{i,j}^t) = \sum_{\ell=i}^{j-1} b_{\ell,i,j}(t) H(T_{\ell,\ell+1}^t).$$

Case III) Under the condition that at time t at least $n-r+1$; $r \leq s$ components of the system operate, the residual lifetime of the system is given by

$$T_r^t = (T - t | X_{r:n} \geq t); \quad r = 1, 2, \ldots, s, \quad s = 1, 2, \ldots, n. \tag{2.7}$$

Khaledi and Shaked [7] showed that the survival function of T_r^t can be expressed as

$$\begin{aligned} \Pr(T_r^t > x) &= \Pr(T > t + x | X_{r:n} > t) \\ &= \sum_{k=s}^{n} p_k \Pr(X_{k:n} > t + x | X_{r:n} > t). \end{aligned} \tag{2.8}$$

If we assume that $X_{i:n} = 0$ for $i = 0,\ n \geq 1$, then $b_{\ell,0,r}(t) = \frac{\Pr(X_{\ell:n} \leq t < X_{\ell+1:n})}{\Pr(X_{r:n} > t)}$. Therefore, the entropy of T_r^t can be expressed as

$$H(T_r^t) = H(V_r^t) - \sum_{\ell=0}^{r-1} b_{\ell,0,r}(t) E(\log(f_t(\bar{F}^{-1}(V_{\ell,\ell+1})))),$$

where $V_r^t = \bar{F}(T_r^t)$.
In each cases, we can study the properties of the obtained residual entropy and find some bounds for it. We can also compare the predictability of two systems' residual lifetimes and extend the results for the general case $(T - t | T > t)$.

References

[1] Shannon, C. E. (1948). A mathematical theory of communication. *Bell. Sys. Tech. J.* **27**, pp. 379-423.

[2] Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution. *Sankhy A.* **58**, pp. 48-57.

[3] Samaniego, F. J. (2007). *System Signatures and their Applications in Engineering Reliability.* Springer: New York.

[4] Goliforushani, S., Asadi, M. and Balakrishnan, M. (2012). On the residual and inactivity times of the components of used coherent systems. *J. Appl. Prob.*, **49**, 385-404.

[5] Ebrahimi, N, Soofi, E. S. and Soyer, R. (2010). Information measures in perspective. *Int. Stat. Rev.* **78**, pp. 383-412.

[6] Chahkandi, M. and Toomaj, S. (2016). Measure of uncertainty in the residual lifetime of a coherent system. *2nd Seminar on Reliability Theory and its Applications*, 53-58.

[7] Khaledi B. E. and Shaked, M. (2007). Ordering conditional lifetimes of coherent systems. *J. Stat. Plan. Inference.* **137**, 1173-1184.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Best linear unbaised estimator based on k-record values from half-normal distribution

Dastbaravarde, A. [1] Zakerzadeh, H. [2] and Manouchehri, F. [3]

[1,2,3] Department of Statistics, Yazd University

Abstract

The article is discussed the k-record values from the half-normal distribution. The mean, variance and covariance of k-record values are computed. The best linear unbiased estimators for the location and scale parameters of the half-normal distribution are determined based on k-record values.

Keywords: k-record values, Half-normal distribution, Best linear unbiased estimator.

1 Introduction

A random variable X has the half-normal distribution, with the location parameter μ and the scale parameter σ, denoted by $X \sim HN(\mu, \sigma)$, it's cumulative distribution function (cdf) is given by

$$F_X(x;\mu,\sigma) = \begin{cases} 2\Phi(\frac{x-\mu}{\sigma}) - 1 & x \geq \mu \ , \\ 0 & x < \mu \end{cases}$$

[1]dastbaravarde@yazd.ac.ir
[2]h.zaker@yazd.ac.ir
[3]fatemeh.shurije@gmail.com

Where Φ is the cdf of standard normal distribution. So, the probability density function (pdf) of X is given by

$$f_X(x;\mu,\sigma) \;=\; 2\;\varphi(\frac{x-\mu}{\sigma})\;; \qquad x \geq \mu \;\;,$$

where φ is the pdf of the standard normal distribution. The half-normal distribution is the left truncated normal distribution at μ (see [6] and [7]).

Let $\{X_i, i \geq 1\}$ be a sequence of independent and identically distributed (iid) random variables with cdf $F_X(x)$. A k-record is defined based on the kth largest X yet seen. For a definition of k-record, let the nth (k-record) times $T_{n(k)}$ are defined by

$$T_{0(k)} = k \qquad with \;\; probability \;\; 1$$

and, for $n \geq 1$

$$T_{n(k)} = min\left\{j : j > T_{n-1(k)}, X_j > X_{T_{n-1(k)}-k+1:T_{n-1(k)}}\right\}$$

Where $X_{i:m}$ denote the ith order statistic in a sample of size m. The nth k-records are then defined by

$$R_{n(k)} = X_{T_{n(k)}-k+1:T_{n(k)}} \quad ; \;\; n \geq 0 \;,\; k \geq 1.$$

For $k = 1$, the k-records are the ordinary records. So, the results for ordinary records can be obtained as special case. Statistical inference problems based on k-records have been considered by several authors, see, [5],[1], and [4].

The joint pdf of mth and nth k-record values for $m < n$ is given by,

$$f_{R_m(k),R_n(k)}\left(r_m, r_n;\theta\right) = \frac{k^{n+1}\left[-log\left(\bar{F}_X\left(r_m;\theta\right)\right)\right]^m}{m!(n-m-1)!}\left[log\left(\frac{\bar{F}_X\left(r_m;\theta\right)}{\bar{F}_X\left(r_n;\theta\right)}\right)\right]^{n-m-1}$$
$$\times \frac{\left[\bar{F}_X\left(r_n;\theta\right)\right]^{k-1}}{\bar{F}_X\left(r_m;\theta\right)} f_X\left(r_m;\theta\right) f_X\left(r_n;\theta\right); -\infty < r_m < r_n < \infty,$$

and the pdf of $R_{n(k)}$, for $n \geq 0$ is given by (see, [1])

$$f_{R_n(k)}\left(r_n;\theta\right) = \frac{k^{n+1}}{n!}\left[-log\left(\bar{F}_X\left(r_n;\theta\right)\right)\right]^n\left[\bar{F}_X\left(r_n;\theta\right)\right]^{k-1} f_X\left(r_n;\theta\right); \quad r_n \in \mathbb{R}\;.$$

Records theory is a relatively new branch is growing in recent decades. This theory, in addition to comments feature has important practical applications. Climate change, some traffic issues, natural hazards such as wind and issues to determine the strength of materials and calculate the probability of failure and so it is important applications.

In this paper, we consider k-record values from half-normal distribution. In Section 2, we compute the means, variances and covariances of k-record values from half-normal distribution. In Section 3, we determine the best linear unbiased estimators (BLUEs) of the location and scale parameters of half-normal distribution based on k-record values.

2 Mean, variance and covariance of the k-record values

Note that if $R_{n_{(k)}}$ is the nth k-record from the half-nornal distribution with the location μ and scale σ, then

$$R_{n_{(k)}} = \mu + \sigma {R^*}_{n_{(k)}}$$

Where $R^*_{n_{(k)}}$ is the nth k-record from the standard half-normal distribution, $HN(0,1)$. Hence, it's enough we compute the mean, variance and covariance of the standard half-normal distribution.

Let $R^*_{0(k)}, R^*_{1(k)}, ..., R^*_{n(k)}$ be the first $(n+1)$ upper k-record values arising from a sequence of iid standard half-normal random variables. The joint pdf of mth and nth k-record values, $R^*_{m(k)}$ and $R^*_{n(k)}$ for $m < n$ is given by

$$f^*_{R^*_m(k),R^*_n(k)}\left(r^*_m, r^*_n\right) = \frac{2^k k^{n+1}}{m!(n-m-1)!}\left[-log\left(2\bar{\Phi}\left(r^*_m\right)\right)\right]^m \left[log\left(\frac{\bar{\Phi}\left(r^*_m\right)}{\bar{\Phi}\left(r^*_n\right)}\right)\right]^{n-m-1}$$
$$\times \frac{\left[\bar{\Phi}\left(r^*_n\right)\right]^{k-1}}{\bar{\Phi}\left(r^*_m\right)} \varphi\left(r^*_m\right)\varphi\left(r^*_n\right); \quad 0 < r^*_m < r^*_n \ ,$$

where $\varphi(.)$ is density function of the standard normal distribution and $\overline{\Phi} = 1 - \Phi$. The pdf of nth upper k-record value $R^*_{n(k)}$ is given by

$$f^*_{R^*_n(k)}\left(r^*_n\right) = \frac{2^k k^{n+1}}{n!}\left[-log\left(2\bar{\Phi}\left(r^*_n\right)\right)\right]^n \left[\bar{\Phi}\left(r^*_n\right)\right]^{k-1} \varphi\left(r^*_n\right); \quad r^*_n > 0 \ .$$

We define $\alpha_{n(k)} := E(R^*_{n(k)})$, $\alpha_{m,n(k)} := E(R^*_{m(k)}R^*_{n(k)})$, $\beta_{n,n(k)} := Var(R^*_{n(k)})$ and $\beta_{m,n(k)} := Cov(R^*_{m(k)}, R^*_{n(k)})$.

We have computed numerically the values of $\alpha_{n(k)}$ and $\alpha_{m,n(k)}$ for sample size up to 10. The values of means, $\alpha_{n(k)}$ for $n = 0, ..., 9$ and $k = 1, ..., 5$ are given in Table 1. From Table 1, we see that the mean of k-record increases as n increases. But, it decreases as k increases, for fix n. Also, we have computed the values of variances and covariances, $\beta_{m,n(k)}$ for $0 \le m \le n \le 9$ and $k = 1, ..., 5$ and are given in Table 1. From Table 1, we see that the variance of k-record decreases as n increases. Also, we see that there is a positive corrolation between $R_{n_{(k)}}$ and $R_{m_{(k)}}$, for all values of m and n, and it decreases as $|n - m|$ increases.

Table 1: Means of the upper k-record values from $HN(0,1)$.

	n									
k	0	1	2	3	4	5	6	7	8	9
1	0.798	1.389	1.864	2.271	2.628	2.950	3.246	3.520	3.777	4.018
2	0.467	0.843	1.160	1.437	1.686	1.911	2.120	2.314	2.496	2.668
3	0.335	0.616	0.861	1.078	1.275	1.456	1.623	1.780	1.928	2.068
4	0.262	0.488	0.688	0.870	1.035	1.188	1.331	1.465	1.592	1.712
5	0.216	0.406	0.577	0.732	0.875	1.009	1.134	1.253	1.364	1.470

3 Best linear unbiased estimation

Let $R_{0(k)}, R_{1(k)}, ..., R_{n(k)}$ be the first $(n+1)$ k-record values from a half-normal distribution with the location and scale parameters μ and σ, respectively. Then, we have

$$\begin{aligned} E\left(R_{n(k)}\right) &= \mu + \sigma\ \alpha_{n(k)}, \\ Var\left(R_{n(k)}\right) &= \sigma^2\ \beta_{n,n(k)}, \\ Cov\left(R_{m(k)}, R_{n(k)}\right) &= \sigma^2\ \beta_{m,n(k)}. \end{aligned}$$

Suppose $\mathbf{R}_{n(k)} = [R_{0(k)}, ..., R_{n(k)}]'$ denote the vector of k-record values. Then,

$$E(\mathbf{R}_{n(k)}) = \mu\mathbf{1} + \sigma\boldsymbol{\alpha},$$

where $\boldsymbol{\alpha} = \left(\alpha_{0(k)}, \alpha_{1(k)}, ..., \alpha_{n(k)}\right)'$ and $\mathbf{1}$ is a column vector of $(n+1)$ ones. The variance-covariance matrix of $\mathbf{R}_{n(k)}$ is given by $V\left(\mathbf{R}_{n(k)}\right) = \mathbf{B}\sigma^2$, where

$$\mathbf{B} = \left[\beta_{i,j(k)}\right] = \begin{bmatrix} \beta_{00(k)} & \beta_{01(k)} & \cdots & \beta_{0n(k)} \\ \beta_{10(k)} & \beta_{11(k)} & \cdots & \beta_{1n(k)} \\ \vdots & \vdots & \ldots & \vdots \\ \beta_{n0(k)} & \beta_{n1(k)} & \cdots & \beta_{nn(k)} \end{bmatrix}$$

Following the generalized least-squares approach, the BLUEs of μ and σ, denoted by $\hat{\mu}_B$ and $\hat{\sigma}_B$, are given, respectively, by (see, [3], pp. 80-81)

$$\begin{aligned} \hat{\mu}_B &= \left\{ \frac{\boldsymbol{\alpha}'\ \boldsymbol{B}^{-1}\ \boldsymbol{\alpha}\ \mathbf{1}'\ \boldsymbol{B}^{-1} - \boldsymbol{\alpha}'\ \boldsymbol{B}^{-1}\ \mathbf{1}\ \boldsymbol{\alpha}'\ \boldsymbol{B}^{-1}}{(\boldsymbol{\alpha}'\ \boldsymbol{B}^{-1}\ \boldsymbol{\alpha}')\ (\mathbf{1}'\ \boldsymbol{B}^{-1}\ \mathbf{1}) - (\boldsymbol{\alpha}'\ \boldsymbol{B}^{-1}\ \mathbf{1})^2} \right\} \mathbf{R}_{n(k)} \\ &= \boldsymbol{\alpha}'\boldsymbol{\Delta}'\mathbf{R}_{n(k)} \\ &= \sum_{i=0}^{n} c_i R_{i(k)}, \end{aligned}$$

and

$$\begin{aligned} \hat{\sigma}_B &= \left\{ \frac{\mathbf{1}'\ \mathbf{B}^{-1}\ \mathbf{1}\ \boldsymbol{\alpha}'\ \mathbf{B}^{-1} - \mathbf{1}'\ \mathbf{B}^{-1}\ \boldsymbol{\alpha}\ \mathbf{1}'\ \mathbf{B}^{-1}}{(\boldsymbol{\alpha}'\ \mathbf{B}^{-1}\ \boldsymbol{\alpha}')\ (\mathbf{1}'\ \mathbf{B}^{-1}\ \mathbf{1}) - (\boldsymbol{\alpha}'\ \mathbf{B}^{-1}\ \mathbf{1})^2} \right\} \mathbf{R}_{n(k)} \\ &= \mathbf{1}'\boldsymbol{\Delta}\mathbf{R}_{n(k)} \\ &= \sum_{i=0}^{n} d_i R_{i(k)}, \end{aligned}$$

Where $\boldsymbol{\Delta}$ is given by

$$\boldsymbol{\Delta} = \left\{ \frac{\mathbf{B}^{-1}\left(\mathbf{1}\ \boldsymbol{\alpha}' - \boldsymbol{\alpha}\ \mathbf{1}'\right)\mathbf{B}^{-1}}{(\boldsymbol{\alpha}'\ \mathbf{B}^{-1}\ \boldsymbol{\alpha})\ (\mathbf{1}'\ \mathbf{B}^{-1}\ \mathbf{1}) - (\boldsymbol{\alpha}'\ \mathbf{B}^{-1}\ \mathbf{1})^2} \right\}.$$

Furthermore, the variances and covariance of the above estimators are given by (see, [3], pp. 80-81)

$$Var\left(\hat{\mu}_B\right) = \sigma^2 \left\{ \frac{\boldsymbol{\alpha}' \, \mathbf{B}^{-1} \, \boldsymbol{\alpha}}{\left(\boldsymbol{\alpha}' \, \mathbf{B}^{-1} \, \boldsymbol{\alpha}'\right) \left(\mathbf{1}' \, \mathbf{B}^{-1} \, \mathbf{1}\right) - \left(\boldsymbol{\alpha}' \, \mathbf{B}^{-1} \, \mathbf{1}\right)^2} \right\},$$

$$Var\left(\hat{\sigma}_B\right) = \sigma^2 \left\{ \frac{\mathbf{1}' \, \mathbf{B}^{-1}}{\left(\boldsymbol{\alpha}' \, \mathbf{B}^{-1} \, \boldsymbol{\alpha}'\right) \left(\mathbf{1}' \, \mathbf{B}^{-1} \, \mathbf{1}\right) - \left(\boldsymbol{\alpha}' \, \mathbf{B}^{-1} \, \mathbf{1}\right)^2} \right\},$$

and

$$Cov\left(\hat{\mu}_B, \hat{\sigma}_B\right) = \sigma^2 \left\{ \frac{-\boldsymbol{\alpha}' \, \mathbf{B}^{-1} \, \mathbf{1}}{\left(\boldsymbol{\alpha}' \, \mathbf{B}^{-1} \, \boldsymbol{\alpha}'\right) \left(\mathbf{1}' \, \mathbf{B}^{-1} \, \mathbf{1}\right) - \left(\boldsymbol{\alpha}' \, \mathbf{B}^{-1} \, \mathbf{1}\right)^2} \right\}.$$

By using of the values of means, variances and covariances presented in Tables 1 and Table 1, we can compute the coefficients a_i and b_i, $i = 0, 1, ..., n$ of BLUEs of μ and σ for $n = 1, ..., 9$; $k = 1, ..., 5$; The coefficients of BLUEs of μ and σ are presented in Table 3 and Table 4, respectively. From Table 3 and Table 4 , we see that as n increases the variances of $\hat{\mu}_B$ and $\hat{\sigma}_B$ decreases.

4 Conclusion

In this paper, we discussed the k-record values arising from half-normal distribution. We provided the best linear unbiased estimators for the location and scale parameters of half-normal distiribution based on k-record values.

References

[1] Ahmadi, J., Doostparast, M. and Parsian, A. (2005). Estimation and prediction in a two-parameter exponential distribution based on k-record values under LINEX loss function, *Communications in StatisticsTheory and Methods,* **34**, 795–805.

[2] Arnold, B.C., Balakrishnan, N. and Negaraja, H.N. (1998), *Records*, A Wiley-Interscience Publication, New York.

[3] Balakrishnan, N. and Cohen, A.C. (1991). *Order Statistics and Inference Estimation Model*, Academic Press, San Diego.

[4] Chacko, M. and Shy Mary, M. (2013). Estimation avd prediction based on k-record values from normal distribution, *Statistica, anno LXXIII*, **4**, 505–516.

[5] Malinowska, I. and Szynal, D. (2004). on a family of bayesian estimators and predictors for Gumbel model based on the k-th lower record, *Applicationes Mathematicae*, **33**, 107–115.

[6] Pewsey, A. (2002). Large sample inference for general half-normal distribution, *Communication in Statistics-Theory and Methods,* **31**(7), 1045–1054.

[7] Pewsey, A. (2004). Improved likelihood based inference for the general half-normal distribution , *Scandinavian Journal of Statistics*, **12**(2), 171–178.

Table 2: The variances and covariances of the upper k-record values from $HN(0,1)$.

		m									
k	n	0	1	2	3	4	5	6	7	8	9
1	0	0.363									
	1	0.291	0.471								
	2	0.247	0.397	0.510							
	3	0.215	0.348	0.448	0.526						
	4	0.195	0.311	0.398	0.470	0.534					
	5	0.176	0.287	0.366	0.429	0.490	0.535				
	6	0.165	0.265	0.342	0.401	0.450	0.493	0.536			
	7	0.153	0.250	0.314	0.376	0.421	0.458	0.490	0.536		
	8	0.144	0.233	0.300	0.350	0.398	0.437	0.469	0.500	0.535	
	9	0.136	0.222	0.282	0.334	0.375	0.410	0.445	0.478	0.503	0.534
2	0	0.145									
	1	0.122	0.207								
	2	0.106	0.179	0.236							
	3	0.094	0.161	0.210	0.251						
	4	0.086	0.146	0.190	0.229	0.260					
	5	0.079	0.133	0.175	0.210	0.239	0.265				
	6	0.073	0.124	0.164	0.192	0.221	0.246	0.267			
	7	0.070	0.117	0.153	0.185	0.207	0.233	0.252	0.269		
	8	0.066	0.110	0.145	0.174	0.193	0.216	0.233	0.253	0.270	
	9	0.062	0.103	0.136	0.164	0.189	0.207	0.224	0.243	0.253	0.271
3	0	0.080									
	1	0.070	0.122								
	2	0.062	0.108	0.144							
	3	0.056	0.098	0.130	0.158						
	4	0.052	0.090	0.120	0.144	0.166					
	5	0.047	0.083	0.111	0.134	0.155	0.171				
	6	0.044	0.076	0.105	0.125	0.145	0.160	0.175			
	7	0.042	0.073	0.097	0.118	0.135	0.151	0.165	0.177		
	8	0.040	0.070	0.092	0.111	0.127	0.144	0.154	0.168	0.179	
	9	0.038	0.066	0.087	0.106	0.122	0.135	0.148	0.161	0.170	0.179
4	0	0.052									
	1	0.046	0.082								
	2	0.042	0.074	0.099							
	3	0.038	0.067	0.091	0.111						
	4	0.035	0.062	0.084	0.102	0.119					
	5	0.033	0.058	0.078	0.096	0.111	0.124				
	6	0.030	0.054	0.073	0.090	0.105	0.117	0.127			
	7	0.029	0.051	0.070	0.84	0.098	0.110	0.119	0.130		
	8	0.027	0.049	0.065	0.080	0.094	0.104	0.116	0.123	0.132	
	9	0.026	0.046	0.063	0.076	0.088	0.099	0.109	0.116	0.124	0.133
5	0	0.037									
	1	0.033	0.059								
	2	0.030	0.054	0.074							
	3	0.027	0.049	0.068	0.084						
	4	0.026	0.046	0.063	0.078	0.090					
	5	0.024	0.043	0.059	0.073	.085	.095				
	6	0.023	0.041	0.055	0.068	0.080	0.089	0.098			
	7	0.021	0.039	0.053	0.065	0.075	0.085	0.093	0.101		
	8	0.020	0.037	0.049	0.061	0.073	0.080	0.088	0.097	0.103	
	9	0.019	0.035	0.048	0.060	0.068	0.077	0.083	0.091	0.098	0.105

Table 3: Coefficients of k-record values for $\hat{\mu}_B$.

		c_i									
k	n	0	1	2	3	4	5	6	7	8	9
1	1	2.352	−1.352								
	2	1.674	0.170	−0.843							
	3	1.445	0.122	0.091	−0.658						
	4	1.313	0.115	0.060	0.106	−0.594					
	5	1.247	0.067	0.088	0.098	−0.001	−0.499				
	6	1.178	0.092	0.049	0.093	0.028	0.055	−0.495			
	7	1.136	0.084	0.045	0.080	0.050	0.026	0.055	−0.477		
	8	1.091	0.094	0.051	0.078	−0.006	0.069	0.054	0.039	−0.469	
	9	1.072	0.067	0.064	0.063	0.022	0.056	0.043	−0.015	0.111	−0.484
2	1	2.244	−1.244								
	2	1.622	0.113	−0.735							
	3	1.424	0.042	0.115	−0.580						
	4	1.310	0.046	0.068	0.061	−0.485					
	5	1.241	0.049	0.053	0.044	0.027	−0.415				
	6	1.192	0.024	0.086	0.007	0.030	0.089	−0.429			
	7	1.157	0.026	0.059	0.039	−0.019	0.131	−0.022	−0.372		
	8	1.123	0.024	0.061	0.037	−0.008	0.097	0.043	−0.046	−0.330	
	9	1.097	0.028	0.059	0.038	−0.009	0.102	−0.023	0.034	0.007	−0.331
3	1	2.193	−1.193								
	2	1.604	0.073	−0.677							
	3	1.402	0.056	0.049	−0.507						
	4	1.299	0.035	0.069	0.011	−0.414					
	5	1.232	0.035	0.052	0.014	0.056	−0.389				
	6	1.188	0.033	0.029	0.042	0.030	0.024	−0.346			
	7	1.154	0.022	0.037	0.043	0.023	−0.022	0.118	−0.375		
	8	1.130	0.024	0.032	0.021	0.026	0.036	0.037	−0.009	−0.298	
	9	1.111	0.018	0.027	0.038	0.013	0.032	0.019	0.030	−0.005	−0.283
4	1	2.159	−1.159								
	2	1.590	0.054	−0.644							
	3	1.390	0.045	0.045	−0.481						
	4	1.291	0.029	0.044	0.040	−0.403					
	5	1.230	0.012	0.047	0.036	0.038	−0.364				
	6	1.186	0.024	0.024	0.054	0.024	−0.027	−0.285			
	7	1.155	0.024	0.023	0.047	−0.004	0.034	−0.009	−0.270		
	8	1.133	0.005	0.044	0.015	0.032	−0.016	0.045	0.043	−0.302	
	9	1.114	0.009	0.025	0.036	0.027	−0.005	0.021	0.047	−0.039	−0.234
5	1	2.133	−1.133								
	2	1.559	0.079	−0.637							
	3	1.375	0.072	−0.011	−0.436						
	4	1.272	0.054	0.013	0.047	−0.386					
	5	1.218	0.054	−0.020	0.058	0.012	−0.324				
	6	1.176	0.037	−0.006	0.065	−0.007	0.046	−0.311			
	7	1.145	0.044	−0.014	0.048	−0.001	0.045	0.017	−0.283		
	8	1.124	0.039	−0.012	0.046	0.013	0.008	0.030	−0.007	−0.241	
	9	1.102	0.039	−0.018	0.063	−0.006	0.018	−0.015	0.055	0.031	−0.270

Table 4: Coefficients of k-record values for $\hat{\sigma}_B$.

		d_i									
k	n	0	1	2	3	4	5	6	7	8	9
1	1	−1.693	1.693								
	2	−0.880	−0.130	1.010							
	3	−0.622	−0.076	−0.048	0.745						
	4	−0.482	−0.068	−0.016	−0.063	0.628					
	5	−0.414	−0.019	−0.044	−0.054	0.016	0.515				
	6	−0.348	−0.042	−0.007	−0.049	−0.011	−0.013	0.471			
	7	−0.310	−0.036	−0.004	−0.038	−0.031	0.014	−0.023	0.429		
	8	−0.270	−0.044	−0.009	−0.036	0.018	−0.024	−0.022	−0.026	0.414	
	9	−0.254	−0.022	−0.020	−0.024	−0.005	−0.013	−0.013	0.019	−0.071	0.404
2	1	−2.663	2.663								
	2	−1.357	−0.187	1.544							
	3	−0.961	−0.046	−0.147	1.154						
	4	−0.743	−0.054	−0.057	−0.078	0.932					
	5	−0.614	−0.059	−0.029	−0.048	−0.020	0.771				
	6	−0.534	−0.019	−0.083	0.014	−0.025	−0.061	0.708			
	7	−0.476	−0.022	−0.038	−0.040	0.058	−0.131	0.028	0.620		
	8	−0.420	−0.019	−0.041	−0.037	0.041	−0.076	−0.075	0.095	0.532	
	9	−0.379	−0.024	−0.038	−0.038	0.042	−0.084	0.028	−0.029	0.006	0.518
3	1	−3.562	3.562								
	2	−1.827	−0.165	1.991							
	3	−1.255	−0.117	−0.065	1.437						
	4	−0.971	−0.059	−0.119	0.003	1.146					
	5	−0.796	−0.058	−0.076	−0.005	−0.071	1.006				
	6	−0.683	−0.054	−0.016	−0.077	−0.005	−0.056	0.892			
	7	−0.602	−0.028	−0.035	−0.080	0.013	0.052	−0.213	0.892		
	8	−0.548	−0.032	−0.024	−0.030	0.006	−0.080	−0.028	0.059	0.678	
	9	−0.504	−0.017	−0.013	−0.069	0.036	−0.070	0.014	−0.033	−0.011	0.665
4	1	−4.417	4.417								
	2	−2.266	−0.168	2.435							
	3	−1.534	−0.137	−0.088	1.759						
	4	−1.192	−0.079	−0.082	−0.043	1.397					
	5	−0.983	−0.021	−0.095	−0.030	−0.122	1.250				
	6	−0.830	−0.063	−0.013	−0.093	−0.074	0.089	0.983			
	7	−0.729	−0.063	−0.010	−0.069	0.018	−0.112	0.073	0.891		
	8	−0.665	−0.008	−0.072	0.023	−0.085	0.031	−0.084	−0.017	0.877	
	9	−0.606	−0.020	−0.011	−0.041	−0.071	−0.001	−0.007	−0.031	0.052	0.735
5	1	−5.257	5.257								
	2	−2.662	−0.218	2.880							
	3	−1.813	−0.187	−0.011	2.012						
	4	−1.373	−0.108	−0.115	−0.061	1.657					
	5	−1.145	−0.111	0.025	−0.109	−0.034	1.375				
	6	−0.973	−0.039	−0.030	−0.137	0.045	−0.137	1.271			
	7	−0.847	−0.069	0.002	−0.067	0.019	−0.131	−0.045	1.137		
	8	−0.768	−0.049	−0.008	−0.060	−0.033	0.010	−0.097	0.079	0.927	
	9	−0.690	−0.050	0.015	−0.121	0.032	−0.025	0.062	−0.141	−0.029	0.946

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Comparison of nonparametric and parametric estimations of a failure rate function for exact data

Fani, S. [1] and Ghoreyshi, A. S. [2]

[1,2] Department of Statistics, Allameh Tabataba'i University

Abstract

The study of reliability or hazard functions is one of the major topics of interest in biomedical studies and reliability engineering. Many parametric models have received considerable attention in this regard, mainly for reasons of straightforward implementation and ease of analysis. The model misspecification problem of parametric methods motivated many researchers to develop nonparametric approaches in recent years, which demand fewer restrictive assumptions for modeling lifetime data. Although some nonparametric approaches, such as Kaplan-Meier and kernel-based methods, are popular tools for solving function estimation problems, they suffer from some non-trivial issues like restrictions, difficulties with bandwidth or tuning parameter selection. In contrast, one can avoid these issues at the cost of enforcing some qualitative shape constraints. The current study compares the performance of a nonparametric shape-constrained approach with some well-known parametric ones. Empirical studies using simulated and real data sets indicate that the nonparametric shape-constrained method outperforms the parametric counterparts in the case of exact data.

Keywords: Lifetime analysis, Nonparametric shape-constrained approach, Failure rate function, Exact data.

[1] s.fani@atu.ac.ir
[2] at.sa.gh.13700507@gmail.com

1 Introduction

The primary concern in the analysis of lifetime and reliability data is to estimate the distribution of the event times of interest. Many statistical approaches study the direct estimation of the density or the reliability function. However, assessing the risk of an item or subject at certain time, which is called the hazard rate or failure rate function, has received considerable attention in recent years. Fully parametric models with monotone hazard rate, in particular exponential, Weibull, and Gompertz distributions, are commonly used to model lifetimes. For review of such distributions, see, e.g., [6].

In many reliability engineering applications, the lifetime distribution of many mechanical and electrical components has non-monotone hazard rate functions such as bathtub or U-shaped. A bathtub-shaped failure rate function starts at optimum burn-in time at the beginning of a product cycle and then remarkably decreases in the early stage toward an approximately constant hazard rate during the useful life stage. Afterwards, the failure rate starts to increase with the onset of wear-out.

The earliest parametric approach for modeling bathtub-shaped failure rates can be traced back to Smith and Bain who proposed a new probability distribution function, known as the exponential power, to fit real life data with bathtub-shaped failure rates. A comprehensive review of such distributions that accommodate some of the forms of the risk function were studied in [8]. Recently, a new five-parameter distribution, called the beta modified Weibull (BMW) distribution, was introduced by [9], which includes several important submodels, such as the exponentiated Weibull (EW) distribution, the exponentiated exponential (EE) distribution, the modified Weibull (MW) distribution, the beta Weibull (BW) distribution and the generalized modified Weibull (GMW), among several others. Although the advantages of parametric models include the ease in their computation, interpretation and prediction, the main limitation of parametric methods is the necessity of the strict model assumptions. In order to avoid the essential assumptions with these methods, one could employ nonparametric approaches to estimate a failure rate function.

In practice, some prior knowledge may often be available regarding the shape of the underlying failure rate function; therefore, it may be reasonable to make use of this information in estimation. In such situations, this former knowledge can be turned into restrictions on the shape of the true underlying function. Therefore, shape-constrained nonparametric failure rate function estimator create a balance between limited flexible and high restrictive parametric estimators and extremely flexible unconstrained nonparametric estimators. Thus, the non-trivial issues of parametric and nonparametric approaches like, respectively, restrictions and smoothing parameter selection can be eliminated at the price of imposing some qualitative shape constraints on the true failure rate function.

Despite the high incidence of uncensored lifetime data in reliability engineering, literature on nonparametric estimation of a failure rate function under shape constraints is relatively sparse. The nonparametric maximum likelihood estimator (NPMLE) and the least squares estimator of a convex hazard function was studied by [5]. The maximum likelihood estimator (MLE) of a convex hazard function was proved to be piecewise linear function with three segments as shown in [5].

Due to lack of manageable closed-form analytical solution for the NPMLE of a bathtub-shaped

failure rate function, one has to use iterative algorithm. An iterative two-step optimization method was proposed by [4]. More recenetly, [10] developed a fast and simple computational algorithm for computing a failure rate function subject to shape constraints that can be adjusted to accommodate for the case of exact data. This research assess and compare the performance of the nonaparametric shape-constrained approach to some well-known parametric ones on some simulated and real data sets.

The remainder of the paper is structured as follows. In Section 2, we lay out a framework for nonparametric failure rate function estimation through imposing bathtub-shaped restriction in the case of exact observations . Parametric and the shape-constrained nonparametric estimators of a U-shaped failure rate function are compared via a simulation study that is reported in Section 3, and a real-world application that is studied in Section 4.

2 Nonparametric MLE

Let T denote a non-negative continuous random variable with density $f(t)$ representing the time until some specified event occurs. The failure rate function of T is given by

$$h(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{\exp\{-H(t)\}}$$

where $R(t) = 1 - F(t)$ is the reliability function and $H(t)$ denotes the cumulative hazard rate function. Therefore,

$$f(t) = h(t)\exp\{-H(t)\} = h(t)\exp\left\{-\int_0^t h(u)\,\mathrm{d}u\right\}.$$

The main objective of this research is to regulate and compare the nonparametric shape-constrained estimator proposed by [10] with the other existing parametric ones for the case of uncensored observations.

In our general framework, the observed data for each item under study is uncensored (exact). By uncensoring, we mean that the event time of interest is being observed exactly on all items under study. Consider an independent and identically distributed (i.i.d.) random sample of items or subjects $T_1, \cdots, T_n$ from a homogeneous population with a continuous distribution F and density f. Hence, the likelihood function in the case with exact observations can be written as

$$\mathcal{L}(h) = \prod_{i=1}^{n} h(T_i)\exp\{-H(T_i)\}. \tag{2.1}$$

As suggested by [2] in the case when all the observations are uncensored, $\ell(h)$ can be made arbitrarily large by increasing the value of h at the largest observed value $T_{(n)}$. Therefore, one can simply maximize the following modified log-likelihood function

$$\tilde{\ell}(h) = \sum_{\substack{i=1 \\ T_i \neq T_{(n)}}}^{n} \log h(T_i) - \sum_{i=1}^{n} H(T_i), \tag{2.2}$$

where $T_{(n)} = \max\{T_1, \cdots, T_n\}$. In the end, the full NPMLE of a U-shaped hazard function is obtained by additionally setting $\hat{h}(T_{(n)}) = \infty$. [5] showed that the NPMLE of a bathtub-shaped failure rate function is piecewise linear. It can hence be expressed as a piecewise linear function with three parts as follows,

$$h(t) = \alpha + \sum_{j=1}^{k} \nu_j (\tau_j - t)_+ + \sum_{j=1}^{m} \mu_j (t - \eta_j)_+, \tag{2.3}$$

where, for example,

$$(\tau_j - t)_+ = \begin{cases} (\tau_j - t), & \text{if } t < \tau_j, \\ 0, & \text{if } t \geq \tau_j. \end{cases}$$

Note that $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_k)^\top$ and $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_m)^\top$ are the vector of non-negative masses changes at $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_k)^\top$ and $\boldsymbol{\eta} = (\eta_1, \cdots, \eta_m)^\top$ for the decreasing and increasing segments $[0, \tau_k)$ and $[\eta_1, T_{(n)})$, respectively.

This problem has no closed-form solution, and hence iterative methods must be used. [10] proposed a new algorithm for computing a bathtub-shaped failure rate function that can be adjusted to accommodate for the case of exact observations. The main idea of their solution is to always maintain a constant hazard segment, even if its length is zero. Particularly, they consider two situations, namely positive and zero length constant part, which are computationally interchangeable during the computation of the algorithm. By applying this new idea, they reduce the inevitable double looping method to a single loop and hence save a remarkable computation cost.

3 Simulation studies

A simulation study is carried out to compare the performance of nonparametric bathtub-shaped hazard (CNMBH) estimator with that of the parametric BMW, MW, EW, GMW, and BW models in this section. To generate random samples from a bathtub-shaped failure rate function in the simulation study, we applied the bathtub (BT) distribution, which was introduced by [3]. The BT distribution has density function $f(t) = \frac{1+2\beta}{2\alpha\sqrt{\beta^2+(1+2\beta)t/\alpha}}$ for $0 \leq t \leq \alpha$. It is worth pointing out that the BT distribution has bathtub-shaped failure rate function for $-1/3<\beta<1$. In our implementation, we choose $\beta = 0.3$ and $\alpha = 100$ to simulate a dataset similar to real human mortality data.

Now, we first present a brief review of parametric methods which is used for this simulation study. Let G be the cumulative distribution function (cdf) of a random variable. Hence, a generalized class of distributions which known as the BMW distribution is as follows:

$$F(x) = I_{G(x)}(a, b) = \frac{1}{B(a, b)} \int_0^{G(x)} w^{a-1}(1-w)^{b-1}\, \mathrm{d}w, \quad a > 0, b > 0, \tag{3.1}$$

where $I_y(a, b) = B_y(a, b)/B(a, b)$ is the incomplete beta function ratio and $B_y(a, b) = \int_0^y w^{a-1}(1 - w)^{b-1}$ denotes the incomplete beta function GI2010. The corresponding density of (3.1) can be

expressed as $f(x) = \frac{1}{B(a,b)} G(x)^{a-1}[1 - G(x)]^{b-1} g(x)$ where $g(x) = \mathrm{d}G(x)/\mathrm{d}x$ is the density of the baseline distribution. By replacing the probability density function (pdf) and the cdf of the three parameters MW distribution that was proposed by [7], the general form of the BMW cdf and pdf can be obtained.

Several well-known distributions can be considered as special cases of the BMW distribution, e.g., it reduces to the MW distribution when $a = 1$ and $b = 1$, it becomes the EW distribution if $b = 1$ and $\lambda = 0$, it simplifies to the case of the GMW distribution when $b = 1$, and it reduces to the BW distribution for $\lambda = 0$.

In order to assess the performance of a density estimator, two loss functions, namely, the integrated squared error (ISE) and the Hellinger distance HD are used, which are given by, respectively,

$$\mathrm{ISE}(f, \widehat{f}\,) = \int_{\mathbb{R}} \left\{ f(x) - \widehat{f}(x) \right\}^2 \mathrm{d}x, \quad \text{and} \quad \mathrm{HD}(f, \widehat{f}\,) = \int_{\mathbb{R}} \left\{ f(x)^{\frac{1}{2}} - \widehat{f}(x)^{\frac{1}{2}} \right\}^2 \mathrm{d}x,$$

where $\widehat{f}$ is an estimator of the true density f. The expectation of a loss function with regard to the true density f is then applied to evaluate the performance of an estimator. For each entry in the Table 1, we compute the ISE and HD values respectively for each combination of estimator and replication. The mean integrated squared error (MISE) and mean Hellinger distance (MHD) are eventually estimated by the average of these ISE and HD values.

For the BT distribution, the results of the simulation study based on 100 replications with sample sizes 200, 500, and 1000 are summarized in Table 1. Each entry in the table is one of the empirical MISE and MHD values with their corresponding standard errors in parentheses. Also, the best expected loss value among density estimators for a given density is highlighted in boldface.

It can be seen from the results that the performance of the CNMBH estimator dominates the other parametric ones. Broadly speaking, the nonparametric shape-constrained estimator is superior to its parametric density estimators' competitors in terms of the MISE and MHD. Furthermore, one could also notice that the difference between nonparametric shape-constrained estimator and the parametric estimators is more tangible in terms of both loss functions as the sample size increases.

4 Real data example

A real data set named "Aarset Data" is now considered in our numerical study. The data in Table 2 given by [1] contains of lifetimes of 50 devices (in weeks) and the superscripts represent the frequency of duplicate values. In many reliability applications, an appealing graphical technique called total time on test (TTT) transform is used to identify the failure rate shape appropriate to a given data. It has been proved that the failure rate function has a bathtub-shaped if the scaled-TTT curve changes once from convex to concave in $(0, 1)$; see e.g., [1]. The scaled TTT-transform plot indicates a U-shaped hazard rate in Figure 1 (a).

In the empirical studies based on real-world data, we do not have the knowledge of the true underlying density function, and thus we can not exactly utilize the loss functions given in Section 3. As a substitute, we replace the true density f with the empirical probability mass function $\widehat{f}_n$ based

Table 1: Simulation results for the BT distribution in terms of the MISE and MHD.

Estimator	Density	
	BT$(0.2, 100)$	
$n = 200$		
	MISE	MHD
CNMBH	**0.0007** (0.0001)	**0.0145** (0.0008)
BMW	0.0011 (0.0001)	0.0195 (0.0005)
MW	0.0009 (0)	0.0189 (0.0004)
EW	0.0008 (0)	0.0275 (0.0005)
GMW	0.0010 (0.0001)	0.0190 (0.0004)
BW	0.0008 (0)	0.0222 (0.0007)
$n = 500$		
CNMBH	**0.0003** (0)	**0.0061** (0.0003)
BMW	0.0014 (0.0002)	0.0191 (0.0005)
MW	0.0008 (0)	0.0173 (0.0002)
EW	0.0007 (0)	0.0267 (0.0002)
GMW	0.0009 (0)	0.0173 (0.0002)
BW	0.0007 (0)	0.0210 (0.0006)
$n = 1000$		
CNMBH	**0.0002** (0)	**0.0034** (0.0001)
MBW	0.0012 (0.0002)	0.0187 (0.0004)
MW	0.0007 (0)	0.0167 (0.0001)
EW	0.0007 (0)	0.0260 (0.0002)
GMW	0.0008 (0)	0.0166 (0.0001)
BW	0.0006 (0)	0.0200 (0.0005)

Table 2: Lifetimes of 50 devices.

0.1	0.2	$1^{(5)}$	2	3	6	7	11	12	$18^{(5)}$	21	32	36	40	45
46	47	50	55	60	$63^{(2)}$	$67^{(4)}$	72	79	$82^{(2)}$		83	$84^{(3)}$	$85^{(5)}$	$86^{(2)}$

on a test set of size n. Two loss functions, the ISE and the Kullback-Leibler (KL) divergence, are computed for the case of uncensored data as below, respectively,

$$\mathrm{ISE}(\widehat{f}_n, \widehat{f}\,) = \int_{\mathbb{R}} \left\{\widehat{f}(x)\right\}^2 \mathrm{d}x - \frac{2}{n}\sum_{i=1}^{n} \widehat{f}(x_i), \quad \text{and} \quad \mathrm{KL}(\widehat{f}_n, \widehat{f}\,) \qquad = -\frac{1}{n}\sum_{i=1}^{n} \log\left\{\widehat{f}(x_i)\right\}.$$

where $\widehat{f}_n$ denotes the empirical mass function from a test set of size n and $\widehat{f}$ the density estimate is obtained from a training set. In addition, additive constants are excluded from the above formulas.

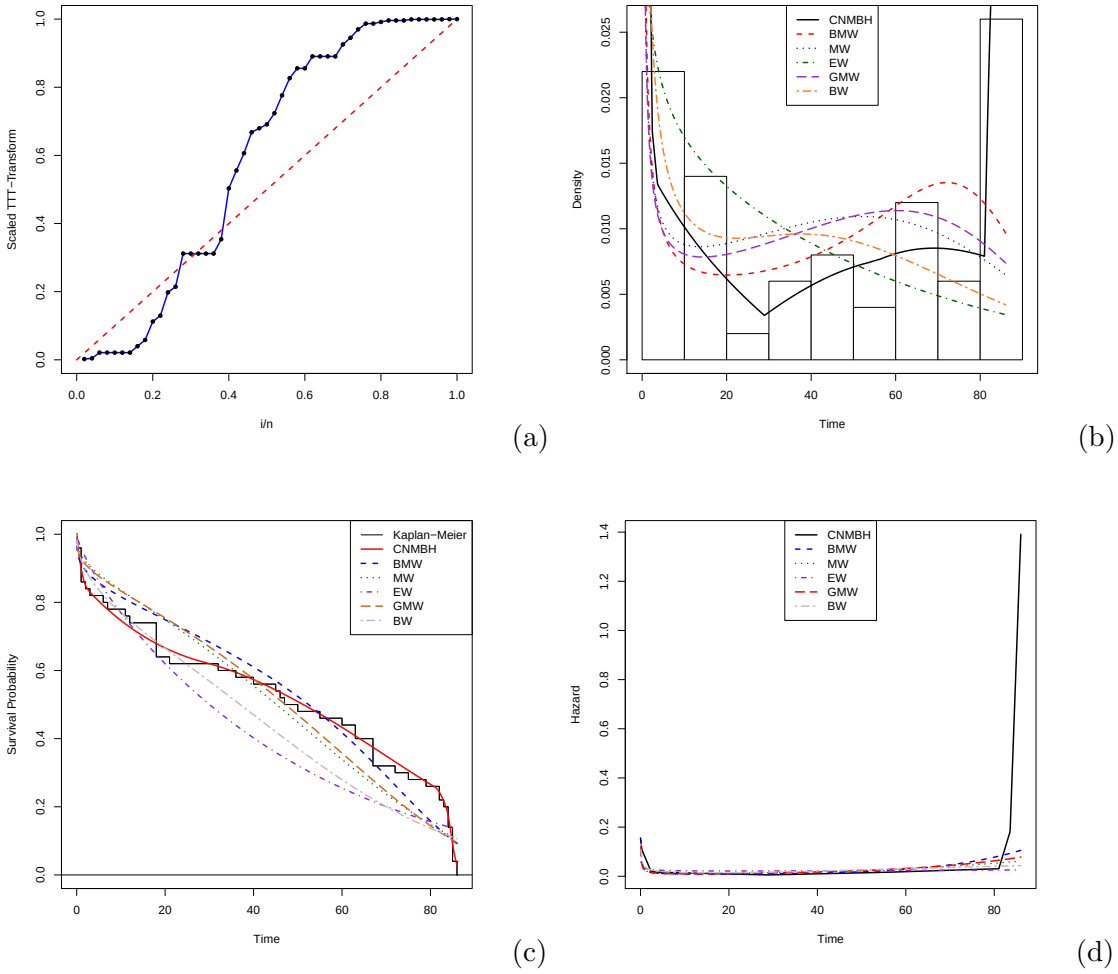


Figure 1: For Aarst data: (a) Scaled TTT-transform; (b) Histogram and fitted nonparametric and parametric pdfs; (c) Empirical and fitted survival functions; (d) Fitted nonparametric and parametric hazard rate functions.

To assess the performance, we ran 2-fold cross-validation, with results produced by averaging over 20 replications. Table 3 provides a brief summary of the estimation results for our proposed nonparametric approach and the parametric ones. In addition, our shape-constrained nonparametric and parametric density, survival and hazard rate estimates are displayed in Figure 1. Overall, the CNMBH estimator provides the best fit to this data set by having the smallest values in terms of both criteria: MISE and MKL. The results reveal that the shape-constrained estimator is more accurate and outperforms the parametric ones for estimation of the failure rate function in

Table 3: Cross-validation results for the nonparametric density estimator and five different parametric estimators for the Aarset data in terms of the MISE and MKL.

Method	MKL	MISE
CNMBH	**4.2661** (0.0812)	**-0.0176** (0.0059)
MBW	4.565 (0.0451)	−0.0031 (0.0071)
BW	4.6397 (0.0142)	−0.0114 (0.0004)
GMW	4.5307 (0.035)	−0.069 (0.0022)
MW	4.5753 (0.0379)	−0.0089 (0.0020)
EW	4.8052 (0.0657)	−0.0102 (0.0027)

the case of exact observations.

References

[1] Aarset, M. V. (1987). How to identify bathtub hazard rate. *IEEE Transactions on Reliability 36*, 106–108.

[2] Grenander, U. (1956). On the theory of mortality measurement. II. *Skand. Aktuarietidskr. 39*, 125–153.

[3] Haupt, E. and H. Schäbe (1997). The TTT transformation and a new bathtub distribution model. *Journal of Statistical Planning and Inference 60*(2), 229–240.

[4] Jankowski, H. K. and J. A. Wellner (2009a). Computation of nonparametric convex hazard estimators via profile methods. *Journal of Nonparametric Statistics 21*(4), 505–518.

[5] Jankowski, H. K. and J. A. Wellner (2009b). Nonparametric estimation of a convex bathtub-shaped hazard function. *Bernoulli 15*(4), 1010–1035.

[6] Kalbfleisch, J. D. and R. L. Prentice (2002). *The statistical analysis of failure time data* (2nd ed.). John Wiley, New York.

[7] Lai, C. D., M. Xie, and D. N. P. Murthy (2003). A modified Weibull distribution. *IEEE Transactions on Reliability 52*, 33–37.

[8] Pham, H. and C. D. Lai (2007). On recent generalizations of the Weibull distribution. *IEEE Transactions on Reliability 56*, 454–458.

[9] Silva, G. O., E. M. Ortega, and G. M. Cordeiro (2010). The beta modified Weibull distribution. *Lifetime Data Analysis 16*, 409–430.

[10] Wang, Y. and S. Fani (2017). Nonparametric maximum likelihood computation of a U-shaped hazard function. *Statistics and Computing*.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Bayesian nonparametric goodness of fit test for survival data

Fazeli Asl, F. [1] and Saberi, Z. [2]

[1,2] Department of Mathematical Sciences, Isfahan University of Technology

Abstract

The paper revisits two-sample hypothesis testing problems. We consider the problem in a fully Bayesian nonparametric framework. Our results are expected to be useful in life testing and survival analysis where classical methods are not quite suitable. Our purpose is to compare two different Bayesian nonparametric methods, Dirichlet process and Pólya tree prior for these problems. Infact, for given two sets of samples $\mathbf{X}$ and $\mathbf{Y}$, with unknown cumulative distribution F_1 and F_2, we show that when one of the two samples of life time data had heavier tail, then the Pólya tree prior are better than Dirichlet process prior.

Keywords: Life testing, Bayesian nonparametric, Dirichlet process, Pólya tree.

1 Introduction

One of the important area of research, that has recently received considerable attention, is two-sample Bayesian nonparametric hypothesis testing with various aplicability. For example, consider $\mathbf{X} = (X_1, ..., X_{m_1})$ and $\mathbf{Y} = (Y_1, ..., Y_{m_2})$ be two samples from F_1 and F_2 respectively, where F_1 and F_2 being unknown continuous cumulative distribution function. The problem is to evaluate the evidence for testing hypothesis:

$$H_0 : F_1 \equiv F_2 \quad versuse \quad H_1 : F_1 \neq F_2. \tag{1.1}$$

[1] forough.fazeli@math.iut.ac.ir
[2] z_saberi@cc.iut.ac.ir

Such tests are encountered in various disciplines, for example they can useful in life testing and survival analysis.

Bayesian nonparametric is a fast developing area in statistics but there has been little written on two-sample Bayesian nonparametric hypothesis testing. For two-sample nonparametric hypothesis test, we can point to the following works. Ma and Wong [8] propose the Coupling Optional Pólya tree prior and used it in two-sample problems. Huang and Ghosh [5] used lehman alternative and Pólya tree prior for lifetime testing. Two of the newest work has been done as follows: Al Labadi et al. [1] proposed a way that is based on the Kolmogorov distance and approximate samples from the Dirichlet process. As for the second way, Holmes et al. [6] used a Pólya tree prior centered on some distribution G and calculated Bayes factor. In practice, we would like to know which one is better in two-sample problems. In this article, our purpose is to compare the power of these two methods. The outline of the remaining sections is as follows. The definition and basic properties of Dirichlet process and Polya tree priors are reviewed in Section 2. Section 3 deals with the introduction of two Bayesian nonparametric methods for two-sapmple problems that have been recently presented. In section 4, we empirically compare the power of these two methods with each other. We illustrate the importance of these Bayesian nonparametric hypothesis testing by means of ovarian cancer data sets in section 5.

2 Prior distribution in Bayesian nonparametric

Putting an appropriate prior on spaces of probability measure with easy and simple calculations, is the most important issue in Bayesian nonparametric problems. Two of these priors that will be used in this article, are Dirichlet process and Pólya tree prior. In this section, Dirichlet process prior and Pólya tree prior are presented for two sample Bayesian nonparametric testing.

2.1 Dirichlet Process priors

Dirichlet process is the most popular prior in Bayesian nonparametric problems with wide applications that introduced by Ferguson [3]. let $\mathcal{X}$ be a space and $\mathcal{A}$ a $\sigma - field$ of a subsets of $\mathcal{X}$. suppose $\mathbf{P}_0$ be a fixed probability measure on ($\mathcal{X}$,$\mathcal{A}$) and α be apositive number. A random probability $\mathbf{P}$ is said to be a Dirichlet process on ($\mathcal{X}$,$\mathcal{A}$) with parameter $\mathbf{P}_0$ and α, if for any finite measurable partion $(A_1, ..., A_k)$ of $\mathcal{X}$, the random vector ($\mathbf{P}(A_1)$,...,$\mathbf{P}(A_k)$) has a Dirichlet distribution with parameter ($\alpha\mathbf{P}_0(A_1)$,...,$\alpha\mathbf{P}_0(A_k)$), where $K \geq 2$. We say $\mathbf{P}$ has a Dirichlet process distribution with base distribution $\mathbf{P_0}$ and concentrate parameter α, and write $\mathbf{P} \sim DP(\alpha, \mathbf{P}_0)$. If $X_1, ..., X_m$ is a sample from $\mathbf{P}$ such that $\mathbf{P} \sim DP(\alpha, \mathbf{P}_0)$, then the posterior distribution of $\mathbf{P}$ given $X_1, ..., X_m$ is also a Dirichlet process on ($\mathcal{X}$,$\mathcal{A}$) with parameter $\alpha' = \alpha + m$ and $\mathbf{P}'_0 = \frac{\alpha}{\alpha+m}\mathbf{P}_0 + \frac{m}{\alpha+m}\frac{\sum_{i=1}^{m}\delta_{X_i}}{m}$, where δ_X denotes the measure giving mass one to the point X. with the increase in α, the posterior base distribution $\mathbf{P}'_0$ gets close to the prior base measure $\mathbf{P}_0$. An alternative definition of Dirichlet process is as follows [3]. For $1 \leq i \leq n+1$, consider ν_i be a sequence of independent and identically distributed (i.i.d.) random variables from an exponential distribution with rate 1 and $\eta_i = \sum_{j=1}^{i}\nu_j$. Also, for $1 \leq i \leq n+1$, suppose that $\gamma_i i.i.d. \mathbf{P}_0$ such that γ_i's are independent of η_i for each i. Ferguson [3], showed that the Dirichlet process with

parameter α and $\mathbf{P}_0$ can be expressed as a normalized series representation

$$\mathbf{P}(.) = \sum_{i=1}^{\infty} \frac{N^{-1}(\eta_i)}{\sum_{i=1}^{\infty} N^{-1}(\eta_i)} \delta_{\gamma_i}(.), \tag{2.1}$$

where,

$$N(x) = \alpha \int_x^{\infty} e^{-t} t^{-1}\, dt, \quad for\ x > 0. \tag{2.2}$$

Working with (2.1) is difficult in practice because no closed form for the inverse of (2.2) exists. One way to approximate (2.2) is as follows [9].

Consider X_n be a random variable with distribution Gamma$(\frac{\alpha}{n}, 1)$. Define

$$G_n(x) = Pr(X_n > x) = \int_x^{\infty} \frac{1}{\Gamma(\frac{\alpha}{n})} e^{-t} t^{\frac{\alpha}{n}-1}\, dt. \tag{2.3}$$

and

$$G_n^{-1} = inf\{x: \quad G_n(x) \geq y\}, \quad 0 < y < 1. \tag{2.4}$$

Consider γ_i and η_i for each $1 \leq i \leq n+1$ as before, then,

$$\mathbf{P}_n = \sum_{i=1}^{n} \frac{G_n^{-1}(\eta_i)}{\sum_{i=1}^{n} G_n^{-1}(\eta_i)} \delta_{\gamma_i}, \tag{2.5}$$

converges almost surely to $\mathbf{P}$ as $n \to \infty$. One way to generate a sample from (A.2) is given in Al labadi et al. [1]. Because of (2.1) Dirichlet process is a disceret measure with probability 1 [3].

2.2 Pólya Tree priors

Pólya trees form a class of distributions for a random probability measure $\mathbf{P}$ intermediate between Dirichlrt process and tail-free processes [3]. Pólya tree can be constructed to give probability 1 to the set of continuous praobability measures. Thus they are much more flexible than the Dirichlet process. This prior is defined as follows.

Let $E = \{0, 1\}$, $E^0 = \phi$, let E^m be the m-fold product $E \times E \times ... \times E$ and $E^* = \bigcup_{m=0}^{\infty} E^m$. Let $\mathcal{X}$ be a separable measurable space, let $\pi_0 = \{\mathcal{X}\}$ and let $\Pi = \{\pi_m : \quad m = 0, 1, ...\}$ be a separating binary tree of partions of $\mathcal{X}$; that is, let $\pi_0, \pi_1, ...$ be sequence of partions such that $\bigcup_{m=0}^{\infty} E^m$ generates the measurable sets. Let $B_\phi = \mathcal{X}$ and for all $\underline{\epsilon}_m = \epsilon_1 ... \epsilon_m \in E^*$, let $B_{\underline{\epsilon}_m 0}$ and $B_{\underline{\epsilon}_m 1}$ be the two pieces into which $B_{\underline{\epsilon}_m}$ is split. From Lavine [7], a random probability measure $\mathbf{P}$ on $\mathcal{X}$ is said to have a Pólya tree distribution with parameter $(\Pi, \mathcal{A})$, written $\mathbf{P} \sim PT(\Pi, \mathcal{A})$, if there exist nonnegative numbers $\mathcal{A} = \{\alpha_{\underline{\epsilon}_m} : \quad \underline{\epsilon}_m \in E^*\}$ and random variables $\mathcal{Y} = \{\beta_{\underline{\epsilon}_m} : \quad \underline{\epsilon}_m \in E^*\}$ such that the following hold:

1)All the random variables in $\mathcal{Y}$ are independent,

2)For every $\underline{\epsilon}_m \in E^*$, $\beta_{\underline{\epsilon}_m}$ has a Beta distribution with parameters $\alpha_{\underline{\epsilon}0}$ and $\alpha_{\underline{\epsilon}1}$,

3)For every m=1,2,... and every $\underline{\epsilon}_m \in E^*$,

$$\mathbf{P}(B_{\underline{\epsilon}_m}) = \Big(\prod_{j=1;\epsilon_j=0}^{m} \beta_{\underline{\epsilon}_{j-1}}\Big)\Big(\prod_{j=1;\epsilon_j=1}^{m} (1 - \beta_{\underline{\epsilon}_{j-1}})\Big), \tag{2.6}$$

Lavine [7] presented a canonical construction of a Pólya tree. This construction is based on some chosen distribution $\mathbf{G}$ so that $E(\mathbf{P}) = \mathbf{G}$. In this case we can choose $\mathcal{A}$ and Π such that $\alpha_{\underline{\epsilon}_m 0} = \alpha_{\underline{\epsilon}_m 1} = c_m$ for every m=0,1,... and

$$B_{\underline{\epsilon}_m} = (\mathbf{G}^{-1}(\frac{k}{2^m}), \mathbf{G}^{-1}(\frac{k+1}{2^m})], \quad k = 0, ..., 2^m - 1. \tag{2.7}$$

For $c_m = m^2$, $\mathbf{P}$ is absolutely continuous with probability one [4]. If $\mathbf{X} = (X_1, ..., X_n)$ is a sample from $\mathbf{P}$ such that $\mathbf{P} \sim PT(\Pi, \mathcal{A})$, then the posterior distribution of $\mathbf{P}$ given $\mathbf{X}$ is Pólya tree with parameter Π and $\mathcal{A}^*$ such that,

$$\mathcal{A}^* = \{\alpha^*_{\underline{\epsilon}_m} : \quad \underline{\epsilon}_m \in E^*\} \quad and \quad \alpha^*_{\underline{\epsilon}_m} = \alpha_{\underline{\epsilon}_m} + n_{\underline{\epsilon}_m}, \tag{2.8}$$

where $n_{\underline{\epsilon}_m}$ denotes the number of observations in $\mathbf{X}$ that lie in the partition $B_{\underline{\epsilon}_m}$.

3 Two procedures for two-sample Bayesian nonparametric hypothesis testing

3.1 Two-sample Bayesian nonparametric hypothesis by using Dirichlet process prior

Let $\mathbf{P}^*_{n1,m1}$ and $\mathbf{Q}^*_{n2,m2}$ are approximations of the posterior distribution of the Dirichlet process given the first and second sample that calculated by (A.2). Also let d:=d($\mathbf{P}^*_{n1,m1}$,$\mathbf{Q}^*_{n2,m2}$) be the kolmogrov distance between $\mathbf{P}^*_{n1,m1}$ and $\mathbf{Q}^*_{n2,m2}$ [1]. This method is defined as follows:

Theorem 3.1. ***Dirichlet process test (DP test)***

1) Set $\boldsymbol{P}_0$ as Normal(0,1) and α=1 in Dirichlet process.

2) Generate a random sample from $\boldsymbol{P}^*_{n1,m1}$ and $\boldsymbol{Q}^*_{n2,m2}$ by (A.2).

3) Compute kolmogorov distance d($\boldsymbol{P}^*_{n1,m1}$,$\boldsymbol{Q}^*_{n2,m2}$).

4) Repead steps (1) − (3) for r times.

5) Repead steps (1) − (4) to calculate the distance d_0 between prior distributions of the Dirichlet process, where α for the first and second sample are $1 + m_1$ and $1 + m_2$ respectively. Use 0.975 quantile of d_0 as a critical value.

6) If the mean of d is greater than d_0, we reject the null hypothesis in (1.1).

3.2 Two-sample Bayesian nonparametric hypothesis by using Pólya tree prior

In this method, we assume that under H_0, $(\mathbf{X}, \mathbf{Y}) \sim PT(\Pi, \mathcal{A})$ and under H_1, $\mathbf{X}$ and $\mathbf{Y}$ independently draws from $\mathbf{P} \sim PT(\Pi, \mathcal{A})$. Holmes et al. [6] calculated Bayes Factor base on the marginal distribution under the two hypothesis. This Bayes factor is as follows:

$$\mathbf{BF} = \frac{Pr((\mathbf{X}, \mathbf{Y})|\Pi, H_0)}{Pr(\mathbf{X}, \mathbf{Y}|\Pi, H_0)}, \tag{3.1}$$

where the marginal distribution in fraction (3.1) is computed in Holmes et al. [6].

Theorem 3.2. ***Pólya tree test (PT test)***

1) Set G as Normal(0,1) in costruction of pólya tree.

2) Compute log of $\boldsymbol{BF}$ *in* (3.1) *as LBF.*

3) Calculate $Pr(H_0|(\boldsymbol{X}, \boldsymbol{Y})) = \dfrac{1}{1+e^{-LBF}}$.

4) If $Pr(H_0|(\boldsymbol{X}, \boldsymbol{Y})) < 0.5$, *we reject* H_0.

4 Power comparison

In this section, we present some simulation studies to compare the performance of two procedure that describe in the previous section. We consider the following experiments designed to explore various canonical depratures from the null such as mean and variance shift and tails:

Case 1: $\mathbf{X} \sim N(0,1)$ and $\mathbf{Y} \sim N(1,1)$

Case 2: $\mathbf{X} \sim N(0,1)$ and $\mathbf{Y} \sim N(1,2)$

Case 3: $\mathbf{X} \sim N(0,1)$ and $\mathbf{Y} \sim t_{0.5}$

Case 4: $log\mathbf{X} \sim N(0,1)$ and $log\mathbf{Y} \sim N(1,1)$.

In each of the above cases, we apply DP and PT tests for $m = m_1 = m_2 = 5, 15, 30, 50, 100$ in 1000 replications and calculate the percentage of rejecting the null hypothesis as the power of test. Also, one may be interested to know, what is the advantage of bayesian nonparametric test than the classical nonparametric test? In order to answer this question, we compare the power of the Bayesian nonparametric tests with the Kolmogorov-Smirnof (K-S) test. The results of the power of the hypothesis testing (1.1) for DP, PT and K-S test are presented in Table 1.

Table 1: Power of the DP, PT and K-S test.

Case	**m**	*DP*	*PT*	*K-S*	**Case**	**m**	*DP*	*PT*	*K-S*
Case1	5	0.433	0.135	0.066	Case2	5	0.119	0.09	0.011
	15	0.684	0.498	0.480		15	0.128	0.241	0.075
	30	0.856	0.806	0.881		30	0.338	0.569	0.211
	50	0.974	0.972	0.984		50	0.650	0.858	0.397
	100	1	0.999	1		100	0.862	0.997	0.81
Case3	5	0.100	0.095	0.008	Case4	5	0.465	0.124	0.072
	15	0.146	0.373	0.07		15	0.696	0.499	0.543
	30	0.435	0.758	0.209		30	0.861	0.758	0.878
	50	0.711	0.978	0.504		50	0.979	0.947	0.981
	100	0.940	1	0.954		100	1	0.997	1

The results of Table 1 show that when the difference between two samples are because of mean shift, then DP test is more efficient than the others but, if the difference between two samples are because of shift variances or tails, then PT test is more better than the others. Also, for small samples sizes (m=5, 15), the Bayesian nonparametric approaches, especially DP test, are more powerful than classical aproach.

5 Real aplication

We use the ovarian cancer data set as data example which was collected by Edmunson et al. [2]. The study include $n = 26$ patients with advanced ovarian carcinoma. Treatment of patients using either cyclophosphamide alone (treatment1 with $m_1 = 13$), or cyclophosphamide plus adriamycin (treatment2 with $m_2 = 13$), by i.v. injection every 3 weeks produced improvement in approximately one third of the patient. The objective of the trial was to see if the two treatments diffrentiate in prolonging the time of survival. Because the sample size is small, it is better to use Bayesian nonparametric approaches to answer this problem. Since the measurments are positive, we use weibull distribution with shape parameter $a = 1.8501$ and scale parameter $b = 674.1266$ as P_0 and G in DP and PT prior. Note that, a and b are the maximum likelihood estimation of shape and scale parameter based on the combined sample. The results of method A and B are presented in Table 1. Figure 2, shows probability density curve for this data set.The results of DP and PT test are presented in Table 1.

Table 2: Results of method A and B for ovarian data set

	d	d_0		$Pr(H_0\|(\underline{X},\underline{Y}))$
Method A	0.39	0.50	Method B	0.62

According to the results of Table 1, since $d < d_0$ and $Pr(H_0|(\underline{X},\underline{Y})) > 0.5$, for this particular data set, there appears to be no evidence against H_0 by using DP and PT test respectively, it means that we can accept the impact of cyclophosphamide plus adriamycin on treatment of ovarian cancer is not different with the impact of cyclophosphamide alone.

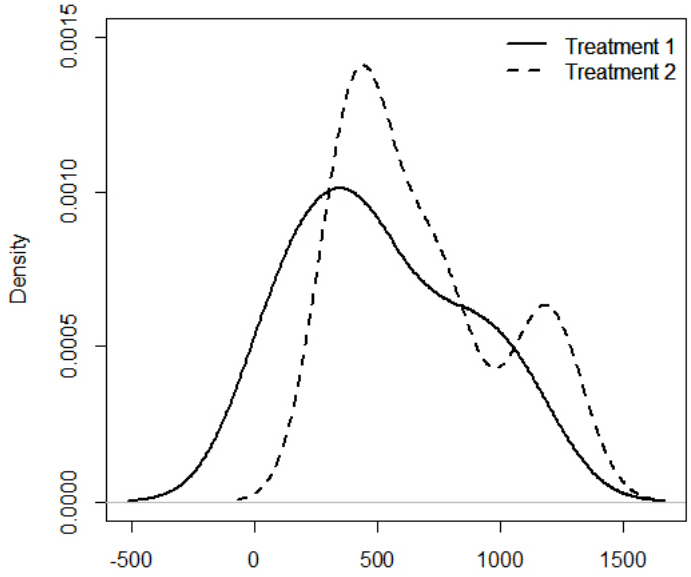


Figure 1: Probability density curve of treatment1 and treatment2.

References

[1] Al Labadi, L., Masuadi, E. and Zarepour, M. (2015), Two-sample Bayesian nonparametric goodness-of-fit test, arXiv:1411.3427v2 [math.ST].

[2] Edmunson, J. H., Fleming, T. R., Decker, G. D., Malkasam, G. D. and Kvols, L. K. (1979), Different chemotherapeutic sensitivities and host factors affecting prognosis in advanced ovarian carcinoma versus minimal disease residual. *Cancer Treatment Reports*, **63**, 241-247.

[3] Ferguson, T. S. (1973), A Bayesian analysis of some nonparametric problems, *Annals of Statistics*, **1**, 209-230.

[4] Ferguson, T. S. (1974), Prior distributions on spaces of probability measures, *Annals of Statistics*, **2**, 615-629.

[5] Huang, L. and Ghosh, M. (2014), Two-sample hypothesis testing under lehmann alternatives and Pólya tree priors, *Statistica Sinica*, **24**, 1717-1733.

[6] Holmes, C. C., Caron, F., Griffin, J. E., and Stephens, D. A. (2015), Two-sample Bayesian nonparametric hypothesis testing, *Bayesian Analysis*, **2**, 297-320.

[7] Lavine, M. (1992), Some aspects of Pólya tree distributions for statistical modelling, *The Annals of Statistics*, **20**, 1222-1235.

[8] Ma, L. and Wong, W.H. (2011), Coupling optional Pólya trees and the two sample problem, *Journal of the American Statistical Association*, **106**, 1553-1565.

[9] Zarepour, M. and Al Labadi, L. (2012), On a rapid simulation of the Dirichlet process, *Statistics and Probability Letters*, **82**, 916-924.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Optimizing the step stress accelerated life test with two stress variables under type-I progressive censoring

Hakamipour, N. [1]

Department of Mathematics, Buein Zahra Technical University

Abstract

Due to cost and time consideration, it is difficult to observe all of the products lifetime within a reasonable time period. Hence, censored lifetime data is usually collected in real applications. Even when accelerated life tests are used, censoring is usually inevitable. The main purpose of this paper is to consider the optimal design for simple step stress accelerated life testing with Type-I progressive censoring. Sometimes, one stress variable does not yield enough failure data. Thus, two stress variables are considered. The lifetime of the items follows the exponentiated exponential distribution and a cumulative exposure model is considered. The problem of choosing the optimal times is developed to minimize the asymptotic variance of the reliability estimate at normal stress condition.

Keywords: Exponentiated exponential distribution, Step stress accelerated life test, Type-I progressive censoring.

1 Introduction

Accelerated life testing (ALT) is a quick way to obtain information about the life distribution of a material, component or product. In Accelerated life testing items are subjected to conditions that are more severe than the normal ones, which yields shorter life but, hopefully, does not change the failure mechanisms. Some assumptions are needed in order to relate the life at high stress levels to

[1] nooshin.hakami@aut.ac.ir

life at normal stress levels in use. Based on these assumptions, the life distribution under normal stress levels can be estimated. Such way of testing reduces both time and cost [6].

Step stress accelerated life testing (SSALT) is an advanced case of ALT. In contrast, step stress accelerated life testing allows the stress level of a test unit to be changed at prespecified times or upon the occurrence of a fixed number of failures. This type of test has advantages of yielding more failure data in a short period of time without necessarily applying a high stress to all test units. The life data from the test are extrapolated through a life-stress relationship and a model that relates the life distribution under a step stress to that under a constant stress. In this paper, a log-linear relationship between the scale parameter of lifetime distribution and the stress are assumed. Furthermore, we model the effects of changing stress as a cumulative exposure (CE) function. The step stress procedure was first introduced, with the cumulative exposure model, by Nelson [8]. Arefi and Razmkhah [1] considered the simple SSALT plan in a discrete set-up for which the failure times at each level of stress are geometrically distributed. Hakamipour and Rezaei [4] proposed an optimal design of simple SSALT with Type-I censored Frechet data.

Since accelerating just one variable does not result enough failure data, it is desirable to include more stress variables. Li and Fard [7] studied SSALT for two stress variables with Weibull failure times under Type-I censoring. Hakamipour and Rezaei [3] studied SSALT for two stress variables with Gompertz failure times under Type-II censoring.

The main purpose of this study is to consider exponentiated exponential distribution for the step stress model under the CE model formulation. Note that the exponentiated exponential distribution is a particular member of the exponentiated Weibull distribution. It is observed in [2] that the two-parameter exponentiated exponential distribution can be used quite effectively in analyzing several lifetime data, particularly in place of two-parameter gamma or two parameter Weibull distribution. If the shape parameter is one, then all the three distributions coincide with the one parameter exponential distribution.

During the past two decades, the problem of optimal scheduling of the step stress test has attracted great attention in the reliability literature. Optimal design of step stress tests could also be based on other criteria chosen by the experimenter. The main focus of this paper is to investigate the choice of optimal change points of the stress levels. The optimization criterion is to minimize the asymptotic variance of the reliability estimate at a specified time ξ under a typical operating condition. The rest of this paper is organized as follows. In Section 2, we describe the model and some necessary assumptions. In Section 3, the problem of choosing the optimal change times will be obtained by using the optimization criterion. In Section 4, some numerical studies were conducted. Finally, some concluding remarks are made in Section 5.

2 Model and Assumption

We consider the bivariate SSALT, so that each stress variable has two levels. Let S_{lk} be the kth stress level of variable l, where $l = 1, 2$ and $k = 0, 1, 2$. The S_{10}, S_{20} are stress levels at typical operating conditions. Let all n units of experiment are initially placed at first step with stress levels (S_{11}, S_{21}) for certain time, τ_1, during which n_1 failures will be observed. At time τ_1, c_1 units are randomly removed from the remaining $n - n_1$ surviving units and the first stress variable is

increased from S_{11} to S_{12}. The second step is continued until the predetermined time τ_2, which n_2 failures will be occured (units fail). In this time, c_2 units are randomly removed from the remaining $n - n_1 - c_1 - n_2$. At the end of the second step, the other stress variable is increased from S_{21} to S_{22}. The test continues until the predetermined time of T, which n_3 units have failed in this step. At time T, the remaining surviving units $c_3 = n - n_1 - c_1 - n_2 - c_2 - n_3$ are all removed from the test.

Notice that we could have also changed the stress level of the second variable, while keeping the first variables stress level fixed. This would lead to a different result, but has no consequences on the model.

The decision to choose one of the stress variables for a stress level change, while the other stress variable remains constant, will lead to a different experimental result, but has no effect on the model. The purpose of the experiment is to obtain failure test data at different stress level combinations [7].

The cumulative distribution function of two parameters exponentiated exponential distribution with shape parameter α and scale parameter λ given respectively by

$$F(t) = (1 - e^{-\frac{t}{\lambda}})^{\alpha}, \quad t, \alpha, \lambda \geq 0$$

We proposed, the scale parameter λ_i at test step i, is a log-linear function of stress levels, for $i = 1, 2, 3$:

$$\begin{aligned} Step1 : \log(\theta_1) &= \beta_0 + \beta_1 S_{11} + \beta_2 S_{21}, \\ Step2 : \log(\theta_2) &= \beta_0 + \beta_1 S_{12} + \beta_2 S_{21}, \\ Step3 : \log(\theta_3) &= \beta_0 + \beta_1 S_{12} + \beta_2 S_{22}. \end{aligned} \tag{2.1}$$

where β_0, β_1 , and β_2 are unknown parameters depending on the nature of the product, and the method of test.

The lifetime of a unit under SSALT, using CE model, can be written as

$$G(t) = \begin{cases} \left(1 - e^{-\frac{t}{\lambda_1}}\right)^{\alpha}, & 0 \leq t < \tau_1, \\ \left(1 - e^{-\frac{\tau_1}{\lambda_1} - \frac{t-\tau_1}{\lambda_2}}\right)^{\alpha}, & \tau_1 \leq t < \tau_2, \\ \left(1 - e^{-\frac{\tau_1}{\lambda_1} - \frac{\tau_2-\tau_1}{\lambda_2} - \frac{t-\tau_2}{\lambda_3}}\right)^{\alpha}, & \tau_2 \leq t < T. \end{cases} \tag{2.2}$$

The corresponding probability distribution function (PDF) can be given as:

$$g(t) = \begin{cases} \dfrac{\alpha}{\lambda_1}\left(1 - e^{-\frac{t}{\lambda_1}}\right)^{\alpha-1} e^{-\frac{t}{\lambda_1}}, & 0 \leq t < \tau_1, \\ \dfrac{\alpha}{\lambda_2}\left(1 - e^{-\frac{\tau_1}{\lambda_1} - \frac{t-\tau_1}{\lambda_2}}\right)^{\alpha-1} e^{-\frac{\tau_1}{\lambda_1} - \frac{t-\tau_1}{\lambda_2}}, & \tau_1 \leq t < \tau_2, \\ \dfrac{\alpha}{\lambda_3}\left(1 - e^{-\frac{\tau_1}{\lambda_1} - \frac{\tau_2-\tau_1}{\lambda_2} - \frac{t-\tau_2}{\lambda_3}}\right)^{\alpha-1} e^{-\frac{\tau_1}{\lambda_1} - \frac{\tau_2-\tau_1}{\lambda_2} - \frac{t-\tau_2}{\lambda_3}}, & \tau_2 \leq t < T. \end{cases} \tag{2.3}$$

Let t_{ij}, $i = 1, 2, 3$, $j = 1, 2, \ldots, n_i$ denote the observation obtained form a Type-I progressively censored sample with random removals in a bivariate SSALT. The number of units removed from the test at each time follows a binomial distribution and any individual unit being removed is with the same probability p.

$$\begin{aligned} P(C_1 = c_1) &= \binom{n - n_1}{c_1} p^{c_1}(1-p)^{n-n_1-c_1} \\ P(C_2 = c_2 \mid C_1 = c_1) &= \binom{n - n_1 - n_2 - c_1}{c_2} p^{c_2}(1-p)^{n-n_1-n_2-c_1-c_2} \end{aligned}$$

Furthermore, suppose that C_i is independent of t_i for all i. Then the joint likelihood function can be found as

$$L(t_i; \lambda_1, \lambda_2, \lambda_3, \alpha, p) = L_1(t_i; \lambda_1, \lambda_2, \lambda_3, \alpha, p \mid \mathbf{C} = \mathbf{c}).P(\mathbf{C} = \mathbf{c}), \tag{2.4}$$

Here $\mathbf{C} = (C_1, C_2)$ and $\mathbf{c} = (c_1, c_2)$, and

$$P(\mathbf{C} = \mathbf{c}) = P(C_2 = c_2 \mid C_1 = c_1)P(C_1 = c_1),$$

and

$$L_1(t_i; \lambda_1, \lambda_2, \lambda_3, \alpha, p \mid \mathbf{C} = \mathbf{c}) = \prod_{i=1}^{3} g(t_i) \prod_{i=1}^{3} (1 - G(t_i))^{c_i},$$

where $G(t_i)$ and $g(t_i)$ will be replaced from equations (2.2) and (2.3). Finally, the likelihood function is obtained with the placement of $L_1(t_i; \lambda_1, \lambda_2, \lambda_3, \alpha, p \mid \mathbf{C} = \mathbf{c})$ and $P(\mathbf{C} = \mathbf{c})$ in equation (2.4). It is usually easier to maximize the logarithm of the likelihood function rather than the likelihood function itself. The logarithm of the likelihood function is

$$
\begin{aligned}
\ell\left(t;\lambda_1,\lambda_2,\lambda_3,\alpha,p\right) \propto\ & (n_1+n_2+n_3)\log\alpha - n_1\log\lambda_1 - n_2\log\lambda_2 - n_3\log\lambda_3 \\
& -\frac{1}{\lambda_1}\sum_{i=1}^{n_1} t_i - \frac{1}{\lambda_2}\sum_{i=1}^{n_2}(t_i-\tau_1) - \frac{1}{\lambda_3}\sum_{i=1}^{n_3}(t_i-\tau_2) \\
& -\frac{n_2\tau_1}{\lambda_1} - \frac{n_3\tau_2}{\lambda_1} - \frac{n_3(\tau_2-\tau_1)}{\lambda_2} + c_1\log\left(1-\left(1-e^{-\frac{\tau_1}{\lambda_1}}\right)^{\alpha}\right) \\
& +c_2\log\left(1-\left(1-e^{-\frac{\tau_1}{\lambda_1}-\frac{\tau_2-\tau_1}{\lambda_2}}\right)^{\alpha}\right) + c_3\log\left(1-\left(1-e^{-\frac{\tau_1}{\lambda_1}-\frac{\tau_2-\tau_1}{\lambda_2}-\frac{T-\tau_2}{\lambda_3}}\right)^{\alpha}\right) \\
& +(\alpha-1)\sum_{i=1}^{n_1}\log(1-e^{-\frac{t_i}{\lambda_1}}) + (\alpha-1)\sum_{i=1}^{n_2}\log(1-e^{-\frac{\tau_1}{\lambda_1}-\frac{t_i-\tau_1}{\lambda_2}}) \\
& +(\alpha-1)\sum_{i=1}^{n_3}\log(1-e^{-\frac{\tau_1}{\lambda_1}-\frac{\tau_2-\tau_1}{\lambda_2}-\frac{t_i-\tau_2}{\lambda_3}}) + (c_1+c_2)\log p \\
& +(2n-2n_1-2c_1-n_2-c_2)\log(1-p) + \log\left(\begin{array}{c} n-n_1 \\ c_1 \end{array}\right) \\
& +\log\left(\begin{array}{c} n-n_1-c_1-n_2 \\ c_2 \end{array}\right).
\end{aligned}
$$

where

$$
\begin{aligned}
A &= 1-e^{-\frac{\tau_1}{\lambda_1}}, \\
B &= 1-e^{-\frac{\tau_1}{\lambda_1}-\frac{\tau_2-\tau_1}{\lambda_2}},
\end{aligned}
$$

and

$$
\begin{aligned}
C &= 1-e^{-\frac{\tau_1}{\lambda_1}-\frac{\tau_2-\tau_1}{\lambda_2}-\frac{T-\tau_2}{\lambda_3}}, \\
D_i &= e^{-\frac{t_i}{\lambda_1}}, \\
E_i &= e^{-\frac{\tau_1}{\lambda_1}-\frac{t_i-\tau_1}{\lambda_2}}, \\
F_i &= e^{-\frac{\tau_1}{\lambda_1}-\frac{\tau_2-\tau_1}{\lambda_2}-\frac{t_i-\tau_2}{\lambda_3}}.
\end{aligned}
$$

Maximum likelihood estimators (MLEs) of the parameters can be derived by maximizing the above equation. Thus, the MLEs of λ_1, λ_2, λ_3, α and p can be found by solving the equations of the first order partial derivatives of the log likelihood function with respect to λ_1, λ_2, λ_3, α and p.

3 Optimization Criterion

In this section, optimality criterion presented. In the SSALT plans, minimizing the asymptotic variance (AV) of the MLE of any parameter of interest such as the mean life or some percentile life

at a specified level of stress may be considered as a commonly used optimization criterion. Since the reliability function and the mean time to failure (MTTF) are related together as $MTTF = \int_0^\infty R(t)dt$, where $R(\xi)$ stands for the reliability function at ξ, the optimization criterion could also be defined as a function of reliability. Here, we minimize the AV of the reliability estimate at time ξ under normal operating conditions to get the optimal values of the change times.

Let $x_i = \frac{S_{i1}-S_{i0}}{S_{i2}-S_{i0}}$, $i = 1, 2$, then $S_{i0} = \frac{S_{i1}-x_i S_{i2}}{1-x_i}$, $i = 1, 2$. From log-linear relationship between the scale parameter of lifetime distribution and the stress at typical operating condition, we can obtain $\log(\lambda_0)$ as follows:

$$\begin{aligned}
\log \lambda_0 &= \beta_0 + \beta_1 S_{10} + \beta_2 S_{20}, \\
&= \beta_0 + \beta_1 \frac{S_{11} - x_1 S_{12}}{1 - x_1} + \beta_2 \frac{S_{21} - x_2 S_{22}}{1 - x_2}, \\
&= \frac{1}{(1 - x_1)} \log \lambda_1 + \frac{(x_2 - x_1)}{(1 - x_1)(1 - x_2)} \log \lambda_2 - \frac{x_2}{(1 - x_2)} \log \lambda_3.
\end{aligned}$$

Thus, we have

$$\lambda_0 = \lambda_1^{\frac{1}{(1-x_1)}} \lambda_2^{\frac{(x_2-x_1)}{(1-x_1)(1-x_2)}} \lambda_3^{-\frac{x_2}{(1-x_2)}}. \tag{3.1}$$

Thus, the reliability under typical operating conditions at time ξ is

$$R_{(S_{10},S_{20})}(\xi) = 1 - \left(1 - e^{-\frac{t}{\lambda_0}}\right)^\alpha \tag{3.2}$$

with the placement of (3.1) in (3.2), MLE of $R_{(S_{10},S_{20})}(\xi)$ is

$$\widehat{R}_{(S_{10},S_{20})}(\xi) = 1 - \left(1 - \exp\left\{ - t\lambda_1^{-\frac{1}{(1-x_1)}} \lambda_2^{\frac{(x_1-x_2)}{(1-x_1)(1-x_2)}} - \lambda_3^{-\frac{x_2}{(1-x_2)}} \right\}\right)^\alpha.$$

The AV of the reliability estimate at time ξ under typical operating conditions can be obtained as follows, using delta theorem:

$$\mathrm{AV}(\widehat{R}_{(S_{10},S_{20})}(\xi)) = H' F^{-1} H, \tag{3.3}$$

where F is Fisher information matrix, which is explained in section 3.1; and H is the row vector of the first derivative of $\widehat{R}_{(S_{10},S_{20})}(\xi)$ with respect to $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$, $\hat{\alpha}_3$ and $\hat{p}$. i.e.,

$$H = \left[\frac{\partial \widehat{R}_{(S_{10},S_{20})}(\xi)}{\partial \hat{\lambda}_1}, \frac{\partial \widehat{R}_{(S_{10},S_{20})}(\xi)}{\partial \hat{\lambda}_2}, \frac{\partial \widehat{R}_{(S_{10},S_{20})}(\xi)}{\partial \hat{\lambda}_3}, \frac{\partial \widehat{R}_{(S_{10},S_{20})}(\xi)}{\partial \hat{\alpha}}, \frac{\partial \widehat{R}_{(S_{10},S_{20})}(\xi)}{\partial \hat{p}}\right]$$

where

$$\begin{aligned}
\frac{\partial \widehat{R}_{(S_{10},S_{20})}(\xi)}{\partial \hat{\lambda}_1} &= \frac{w(1-w)^{\alpha-1}\xi\alpha\lambda_1^{-\frac{1}{(1-x_1)}-1}\lambda_2^{\frac{(x_1-x_2)}{(1-x_1)(1-x_2)}}\lambda_3^{-\frac{x_2}{(1-x_2)}}}{(1-x_1)},\\
\frac{\partial \widehat{R}_{(S_{10},S_{20})}(\xi)}{\partial \hat{\lambda}_2} &= \frac{w(1-w)^{\alpha-1}\xi\alpha(x_1-x_2)\lambda_1^{-\frac{1}{(1-x_1)}}\lambda_2^{\frac{(x_1-x_2)}{(1-x_1)(1-x_2)}-1}\lambda_3^{-\frac{x_2}{(1-x_2)}}}{(1-x_1)(1-x_2)},\\
\frac{\partial \widehat{R}_{(S_{10},S_{20})}(\xi)}{\partial \hat{\lambda}_3} &= \frac{w(1-w)^{\alpha-1}\xi\alpha x_2\lambda_1^{-\frac{1}{(1-x_1)}}\lambda_2^{\frac{(x_1-x_2)}{(1-x_1)(1-x_2)}}\lambda_3^{-\frac{x_2}{(1-x_2)}-1}}{(1-x_2)},\\
\frac{\partial \widehat{R}_{(S_{10},S_{20})}(\xi)}{\partial \hat{\alpha}} &= -(1-w)^{\alpha}\log(1-w),\\
\frac{\partial \widehat{R}_{(S_{10},S_{20})}(\xi)}{\partial \hat{p}} &= 0.
\end{aligned}$$

where, $w = \exp\{-\xi\lambda_1^{-\frac{1}{(1-x_1)}}\lambda_2^{\frac{(x_1-x_2)}{(1-x_1)(1-x_2)}}\lambda_3^{-\frac{x_2}{(1-x_2)}}\}$.

The values τ_1^* and τ_2^* that minimizes $\mathrm{AV}\big[\widehat{R}_{S_{10},S_{20}}(\xi)\big]$, given by equation (3.3), leads to the optimal SSALT plan.

3.1 Fisher Information Matrix

As mentioned above, to estimate the $\mathrm{AV}\big[\widehat{R}_{S_{10},S_{20}}(\xi)\big]$, the expected Fisher information matrix F must be obtained. The Fisher information matrix plays a key role in the parameter estimation. Its elements are obtained by taking the expected values of the negative second partial derivatives of log likelihood function with respect to λ_1, λ_2, λ_3, α and p. Then the expected Fisher information matrix F is obtained as follows:

$$F = \begin{bmatrix}
-E(\frac{\partial^2\ell}{\partial\lambda_1^2}) & -E(\frac{\partial^2\ell}{\partial\lambda_2\partial\lambda_1}) & -E(\frac{\partial^2\ell}{\partial\lambda_3\partial\lambda_1}) & -E(\frac{\partial^2\ell}{\partial\alpha\partial\lambda_1}) & -E(\frac{\partial^2\ell}{\partial p\partial\lambda_1})\\
-E(\frac{\partial^2\ell}{\partial\lambda_2\partial\lambda_1}) & -E(\frac{\partial^2\ell}{\partial\lambda_2^2}) & -E(\frac{\partial^2\ell}{\partial\lambda_3\partial\lambda_2}) & -E(\frac{\partial^2\ell}{\partial\alpha\partial\lambda_2}) & -E(\frac{\partial^2\ell}{\partial p\partial\lambda_2})\\
-E(\frac{\partial^2\ell}{\partial\lambda_3\partial\lambda_1}) & -E(\frac{\partial^2\ell}{\partial\lambda_3\partial\lambda_2}) & -E(\frac{\partial^2\ell}{\partial\lambda_3^2}) & -E(\frac{\partial^2\ell}{\partial\alpha\partial\lambda_3}) & -E(\frac{\partial^2\ell}{\partial p\partial\lambda_3})\\
-E(\frac{\partial^2\ell}{\partial\alpha\partial\lambda_1}) & -E(\frac{\partial^2\ell}{\partial\alpha\partial\lambda_2}) & -E(\frac{\partial^2\ell}{\partial\alpha\partial\lambda_3}) & -E(\frac{\partial^2\ell}{\partial\alpha^2}) & -E(\frac{\partial^2\ell}{\partial p\partial\alpha})\\
-E(\frac{\partial^2\ell}{\partial p\partial\lambda_1}) & -E(\frac{\partial^2\ell}{\partial p\partial\lambda_2}) & -E(\frac{\partial^2\ell}{\partial p\partial\lambda_3}) & -E(\frac{\partial^2\ell}{\partial p\partial\alpha}) & -E(\frac{\partial^2\ell}{\partial p^2})
\end{bmatrix}$$

The above calculations have been done, but are not listed here because of the page restrictions.

4 Simulation Study

In this section, we first present an artificial example to illustrate the proposed procedure in obtaining the optimal SSALT plan with two stress variables. Then, to examine the effect of changes in the initial parameters λ_1, λ_2, λ_3, α and p on the optimal values of τ_1 and τ_2, a sensitivity analysis is performed. Its objective is to identify the sensitive parameters, which need to be estimated with special care to minimize the risk of obtaining an erroneous optimal solution. Toward this end, the values of τ_1^* and τ_2^* are derived when one of the objectives λ_1, λ_2, λ_3, α and p changes and the others are fixed. To determine the values of τ_1^* and τ_2^* in this analysis, we assume that $\lambda_1 = 0.2$, $\lambda_2 = 0.4$, $\lambda_3 = 0.6$, $\alpha = 0.5$, $p = 0.5$, $n = 50$, $T = 0.3$ and $\xi = 0.5$. For the initial values τ_1^* and τ_2^* are calculated 0.2409 and 0.2758, respectively.

Tables 1-5 present the sensitivity analysis for the different values of the parameters, λ_1, λ_2, λ_3, α and p, respectively. From these tables, we can see that as these parameters increase, the optimal stress change times not change or very slightly increase. We see, these parameters have a too small effect on the hold times. According to this criterion, the optimal hold times are not too sensitive.

Table 1: Optimal times τ_1^* and τ_2^* versus changes in λ_1 with $(\lambda_2 = 0.4, \lambda_3 = 0.6, \alpha = 0.5, p = 0.5)$.

λ_1	0.1	0.15	0.2	0.25	0.3
$\tau_1^\star$	0.2403	0.2407	0.2409	0.2413	0.2416
$\tau_2^\star$	0.2753	0.2754	0.2758	0.2761	0.2762

Table 2: Optimal times τ_1^* and τ_2^* versus changes in λ_2 with $(\lambda_1 = 0.2, \lambda_3 = 0.6, \alpha = 0.5, p = 0.5)$.

λ_2	0.3	0.35	0.4	0.45	0.5
$\tau_1^\star$	0.2408	0.2409	0.2409	0.2411	0.2415
$\tau_2^\star$	0.2753	0.2755	0.2758	0.2759	0.2760

Table 3: Optimal times τ_1^* and τ_2^* versus changes in λ_3 with $(\lambda_1 = 0.2, \lambda_2 = 0.4, \alpha = 0.5, p = 0.5)$.

λ_3	0.5	0.55	0.6	0.65	0.7
$\tau_1^\star$	0.2409	0.2409	0.2409	0.2411	0.2415
$\tau_2^\star$	0.2754	0.2755	0.2758	0.2760	0.2760

5 Conclusion

In this paper, we have proposed an optimum design for SSALT with two variables for exponentiated exponential distribution based on Type-I progressive censoring. The optimization criterion is based

Table 4: Optimal times τ_1^* and τ_2^* versus changes in α with $(\lambda_1 = 0.2, \lambda_2 = 0.4, \lambda_3 = 0.6, p = 0.5)$.

α	0.4	0.45	0.5	0.55	0.6
$\tau_1^\star$	0.2407	0.2408	0.2409	0.2413	0.2415
$\tau_2^\star$	0.2754	0.2756	0.2758	0.2759	0.2761

Table 5: Optimal times τ_1^* and τ_2^* versus changes in p with $(\lambda_1 = 0.2, \lambda_2 = 0.4, \lambda_3 = 0.6, \alpha = 0.5)$.

p	0.4	0.45	0.5	0.55	0.6
$\tau_1^\star$	0.2407	0.2409	0.2409	0.2413	0.2416
$\tau_2^\star$	0.2753	0.2756	0.2758	0.2760	0.2760

on minimizing the AV of the reliability estimation. Furthermore, according to simulation studies, we have found that since the optimal hold times are not too sensitive to the model parameters, thus we anticipate that the proposed design is robust.

References

[1] Arefi, A. and Razmkhah, M. (2013), Optimal simple step-stress plan for Type-I censored data from Geometric distribution, *Journal of Iranian Statistical Society* **12** (2), 193–210.

[2] Gupta, R. D. and Kundu, D. (2001), Exponentiated exponential family; an alternative to gamma and Weibull, *Biometrical Journal* **43**, 117–130.

[3] Hakamipour, N. and Rezaei, S. (2015), Optimal design for a bivariate simple step-stress accelerated life testing model with Type-II censoring and gompertz distribution, *International Journal of Information Technicaly & Decision Making* **14**(6), 1243–1262.

[4] Hakamipour, N. and Rezaei, S. (2017), Optimizing the simple step stress accelerated life test with Fréchet data, *REVSTAT- Statistical Journal* **15**(1), 1–23.

[5] Ismail, A. A., and Sarhan, A. M. (2009), Optimal design of step-stress life test with progressively Type-II censored exponential data, *In International Mathematical Forum*, **4** (40), 1963-1976.

[6] Kamal, M. and Zarrin, S. (2013), Step stress accelerated life testing plan for two parameter Pareto distribution, *Theory & Applications*, **30**, 30–40.

[7] Li, C. and Fard, N. (2007), Optimum bivariate step-stress accelerated life test for censored data, *Reliability, IEEE Transactions*, **56**(1), 77–84.

[8] Nelson, W. B. (1980), Accelerated life testing step-stress models and data analyses, *IEEE Transactions on Reliability*, **29**, 103–108.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

On invariant test with sequential order statistics: An open problem and some suggestions

Hashempour, M. [1] and Doostparast, M. [2]

[1] Department of Statistics, School of Science, University of Hormozgan
[2] Department of Statistics, Ferdowsi University of Mashhad

Abstract

In this paper, sequential order statistics (SOS) coming from heterogeneous exponential populations are considered. On the basis of the observed multiple SOS samples the generalized likelihood ratio test (GLRT) is derived for testing the homogeneity of the populations. It is shown that the GLRT in this case is scale invariant. Some guidelines for deriving the uniformly most powerful scale-invariant test (if exists) are also given.

Keywords: Hypotheses testing, Invariant test, Sequential order statistics.

1 Introduction

Let $X_1, \cdots, X_n$ be independent and identically distributed (i.i.d.) random variables with a common distribution function (DF), say F, and denoted by $X_1, \cdots, X_n \overset{i.i.d.}{\sim} F$. Denote in magnitude order of $X_1, \cdots, X_n$ by $X_{1:n} \leq \cdots \leq X_{n:n}$, which are called order statistics (OSs). In system reliability analyses, lifetimes of r-out-of-n systems, say T, coincide to $X_{r:n}$ in which $X_1, \cdots, X_n$ stand for component lifetimes. When the component lifetimes $X_1, \cdots, X_n \overset{i.i.d.}{\sim} F$, the OSs are used for describing the system lifetime. Notice that failing a component does not change the lifetimes of

[1] ma.hashempour@hormozgan.ac.ir
[2] doustparast@um.ac.ir

the surviving components. Motivated by Cramer and Kamps [3, 4], the failure of a component may results in a higher load on the remaining components and hence causes the distributions of the surviving components change. Examples of such phenomena include automobile industries, gas and oil transmission pipelines, etc. In these cases, the system lifetimes may be modelled by SOSs. To see this, suppose that F_j, for $j = 1, \cdots n$, denotes the common DF of the lifetime components when $n-j+1$ components are working. The components begin to work at time $t = 0$ independently with the common DF F_1. When at time x_1, the first failure occurs, the remaining $n-1$ components work independently with the common DF F_2. This process continues to $n-r+1$ components independently with the common DF 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 is then r-th component failure time, denoted by $X^{\star}_{(r)}$. In the literature, $(X^{\star}_{(1)}, \cdots, X^{\star}_{(n)})$ is called SOSs; See, Kamps [8, 9]. The problem of estimating parameters on the basis of SOS has been considered in the literature. For example, Cramer and Kamps [3] considered the problem of estimating the parameters on the basis of several independent SOS samples under a conditional proportional hazard rates (CPHR) model. They considered defined by $\bar{F}_j(t) = \bar{F}_0^{\alpha_j}(t)$ for $j = 1, \cdots, r$, where the baseline DF $F_0(t)$ is the exponential distribution, i.e.

$$F_0(x;\sigma) = 1 - \exp\left\{-\left(\frac{x}{\sigma}\right)\right\}, \quad x > 0, \quad \sigma > 0. \tag{1.1}$$

In this case, the hazard rate function of the DF F_j, defined by $h_j(t) = f_j(t)/\bar{F}_j(t)$ for $t > 0$ and $j = 1, \cdots, n$, is proportional to the hazard rate function of the baseline DF F_0, i.e. $h_j(t) = \alpha_j h_0(t)$ where $\alpha_j > 0$; See Kamps [8, 9], Cramer and Kamps [3, 4, 5]. Statistical inferences on the basis of multiply SOS from homogeneous populations have been considered in the literature; See, e.g., Hashempour and Doostparast [7], Esmailian and Doostparast [6], Beutner and Kamps [2], Bedbur [1] and references therein.

Cramer and Kamps [5] considered independent SOS samples from different populations and derived the maximum likelihood (ML) estimates of the parameters of interest under the CPHR model. Moreover, hypotheses testing on the basis of multiple SOS arising from heterogeneous populations have not been studied by the researchers in details. In this paper, we consider the problem of comparing heterogeneous exponential populations on the basis of independent multiply SOS samples coming from heterogeneous exponential populations under the CPHR model.

2 SOS-based likelihood analysis

We here assume that ($s \geq 2$) independent SOS samples of equal size r from s heterogeneous populations are available. The data may be represented by

$$\mathbf{x} = \begin{bmatrix} x_{11} & \dots & x_{1r} \\ \vdots & \ddots & \vdots \\ x_{s1} & \dots & x_{sr} \end{bmatrix}, \tag{2.1}$$

where the i-th row of the matrix $\mathbf{x}$ in (2) denotes the SOS sample coming from the i-th population. The LF of the available data given by (2.1) is then

$$L(\mathcal{F};\mathbf{x}) = A^s \prod_{i=1}^{s} \left(\prod_{j=1}^{r-1} \left[f_j^{[i]}(x_{ij}) \left(\frac{\bar{F}_j^{[i]}(x_{ij})}{\bar{F}_{j+1}^{[i]}(x_{ij})} \right)^{n-j} \right] f_r^{[i]}(x_{ir}) \bar{F}_r^{[i]}(x_{ir})^{n-r} \right), \tag{2.2}$$

where $A = n!/(n-r)!$, $\mathcal{F} = \{F_j^{[i]}, i = 1, \cdots, s, j = 1, \cdots, r\}$ and for $i = 1, \cdots, s$, $j = 1, \cdots, r$, $\bar{F}_j^{[i]}(x) = 1 - F_j^{[i]}(x)$, and $F_j^{[i]}$ calls for the common DF of component lifetimes of the $i-$th sequential r-out-of-n system. For more details, see Cramer and Kamps [4, 5].
Upon substituting Equation (1.1) into Equation (2.2), under the earlier mentioned CPHR model, when the baseline DF of the i-th population ($i = 1, \cdots, s$) follows the one-parameter exponential distribution with the mean σ_i, the LF of the available data reduces to

$$L(\sigma_1, \cdots, \sigma_s, \boldsymbol{\alpha}; \mathbf{x}) = A^s \left(\prod_{j=1}^{r} \alpha_j \right)^s \left(\prod_{i=1}^{s} \frac{1}{\sigma_i} \right)^r \exp\Big\{ - \sum_{i=1}^{s} \sum_{j=1}^{r} \Big(\frac{x_{ij} m_j}{\sigma_i} \Big) \Big\}, \tag{2.3}$$

where $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_r)$, and for $j = 1, \cdots, r$, $\alpha_j > 0$, and $m_j = (n-j+1)\alpha_j - (n-j)\alpha_{j+1}$ with convention $\alpha_{r+1} \equiv 0$. The problem of non-Bayesian estimating the parameters of interest in this case has been considered by Cramer and Kamps [4, 5]. We consider the problem of homogeneity testing on the basis of independent SOS samples from different exponential populations, i.e.,

$$H_0 : \sigma_1 = \cdots = \sigma_s \quad v.s \quad H_1 : \sigma_i \neq \sigma_j \quad \exists i \neq j. \tag{2.4}$$

Remark 2.1. Notice that Bedbur [1] obtained the uniformly most powerful unbiased (UMPU) tests for some elements of the parameter vector $\boldsymbol{\alpha}$ under the CPHR model based on multiply homogeneous SOS samples from a common exponential distribution.

Suppose that the vector parameter $\boldsymbol{\alpha}$ in Equation (2.3) is known. By Theorem 8.1 in Cramer and Kamps [5] and under the null hypothesis H_0 in (2.4), the unique ML estimate of the common mean of the s exponential populations, say σ_0, is

$$\hat{\sigma}_0 = \frac{\sum_{i=1}^{s} \sum_{j=1}^{r} x_{ij} m_j}{rs} = \frac{\sum_{i=1}^{s} \sum_{j=1}^{r} (n-j+1)\alpha_j D_{ij}}{rs}, \tag{2.5}$$

where $D_{ij} = x_{ij} - x_{i,j-1}$, for $j = 1, \cdots, r$. Here $x_{i0} := 0$ for $i = 1, \cdots, s$. When the baseline exponential populations are heterogeneous, Equation (2.5) yields the unique ML estimate of σ_i ($i = 1, \cdots, s$) as

$$\hat{\sigma}_i = \frac{\sum_{j=1}^{r} x_{ij} m_j}{r} = \frac{\sum_{j=1}^{r} (n-j+1)\alpha_j D_{ij}}{r}. \tag{2.6}$$

For more details, see Cramer and Kamps [4, 5].

Corollary 2.2. *Under the CPHR with the one-parameter exponential baseline CDF,*

$$T_i = \sum_{j=1}^{r} (n-j+1)\alpha_j D_{ij} \sim \Gamma(r, \sigma_i), \quad i = 1, \cdots, s, \tag{2.7}$$

where $\Gamma(a, b)$ calls for the gamma distribution with shape and scale parameters a and b, respectively.

Remark 2.3. Notice that

$\sum_{j=1}^{r} x_{ij} m_j = \sum_{j=1}^{r} (n-j+1)\alpha_j D_{ij}$, for $i = 1, \cdots, s$.

Now consider the problem of hypotheses testing (2.4). The generalized likelihood ratio test (GLRT) statistic for testing the problem (2.4) is (Lehman and Romano [10])

$$\begin{aligned} \Lambda_1 &= \frac{\sup_{\Omega_0} L(\sigma_1, \cdots, \sigma_s; \mathbf{x})}{\sup_{\Omega} L(\sigma_1, \cdots, \sigma_s; \mathbf{x})} \\ &= \prod_{i=1}^{s} \left(\frac{\hat{\sigma}_i}{\hat{\sigma}_0} \right)^r \exp\left\{ \sum_{i=1}^{s} \sum_{j=1}^{r} \left(\frac{1}{\hat{\sigma}_i} - \frac{1}{\hat{\sigma}_0} \right) m_j x_{ij} \right\}, \end{aligned} \tag{2.8}$$

where $\Omega = \{(\sigma_1, \cdots, \sigma_s) : \sigma_i > 0, i = 1, \cdots, s\} =: \mathbb{R}^{+s}$ is the whole parameter space while $\Omega_0 = \{(\sigma_1, \ldots, \sigma_s) : \sigma_1 = \cdots = \sigma_s\}$ denotes the parameter space under the null hypothesis H_0. After some algebraic manipulations, the logarithm of the GLRT statistic Λ_1 given by Equation (2.8) simplifies to

$$\log \Lambda_1 = r \sum_{i=1}^{s} \log \left(\frac{s T_i}{\sum_{j=1}^{s} T_j} \right), \tag{2.9}$$

where T_i is defined by Equation (2.7) and "log" stands for the natural logarithm. Then, the null hypothesis H_0 in (2.4) is rejected if

$$A(\boldsymbol{T}, \boldsymbol{\alpha}) > c, \tag{2.10}$$

where $\boldsymbol{T} = (T_1, \cdots, T_s)$ and

$$A(\boldsymbol{T}, \boldsymbol{\alpha}) = -\sum_{i=1}^{s} \log \left(\frac{T_i}{\sum_{j=1}^{s} T_j} \right).$$

The constant c in Equation (2.10) is obtained subject to the level of the test, say γ. To derive the constant c, we need the joint distribution of the statistic $\boldsymbol{T}$ under the null hypothesis H_0 in the problem (2.4). To do this, let $W_i = T_i / \sum_{j=1}^{s} T_j$, for $i = 1, \cdots, s-1$ and $W_s = -\sum_{i=1}^{s} \log \left(T_i / \sum_{j=1}^{s} T_j \right)$. The Jacobian transformation is then

$$J = \exp\{-w_s\} / \prod_{i=1}^{s-1} w_i. \tag{2.11}$$

The joint probability density function (PDF) of $(W_1, \cdots, W_s)$ under the homogeneity hypothesis H_0 is derived from Equations (2.7), (2.11) as

$$\begin{aligned} f_{W_1, \cdots, W_s}(w_1, \cdots, w_s) &= \frac{\Gamma(sr)}{\Gamma(r)^s} \left(\prod_{i=1}^{s-1} w_i^{r-1} \right) \frac{\exp\{-(r-1)w_s\}}{\prod_{i=1}^{s-1} w_i^{r-1}} \frac{\exp\{-w_s\}}{\prod_{i=1}^{s-1} w_i} \\ &= \frac{\Gamma(sr)}{\Gamma(r)^s} \frac{\exp\{-r w_s\}}{\prod_{i=1}^{s-1} w_i}, \end{aligned} \tag{2.12}$$

for $w_i \geq 0$, and $\sum_{i=1}^{s-1} w_i + \frac{\exp\{-w_s\}}{\prod_{i=1}^{s-1} w_i} = 1. Therefore, the marginal PDF of \mathrm{W}_s$ is readily obtained from Equation (2.12) as

$$f_{W_s}(w_s) = \overbrace{\int\int\cdots\int}^{A} \frac{\Gamma(sr)}{\Gamma(r)^s} \frac{\exp\{-rw_s\}}{\prod_{i=1}^{s-1} w_i} dw_1 \cdots dw_{s-1}, \tag{2.13}$$

where $A = \left\{(w_1, \cdots, w_{s-1}) | w_i \geq 0, \sum_{i=1}^{s-1} w_i + \exp\{-w_s\}/\prod_{i=1}^{s-1} w_i = 1\right\}$.

In practice, one may use numerical methods such as Bootstrap and Monte Carlo simulation to derive the threshold c in the rejection region (2.10).

3 Uniformly most powerful invariant test

It is easy to verify that the family of distribution (2.3) is invariant with respect to the group of the scale transformations

$$\mathcal{G} = \{g_a : g_a(\mathbf{x}) = a\mathbf{x} = \{ax_{ij}^\star\}_{1\leq i\leq s, 1\leq j\leq r}, a > 0\}. \tag{3.1}$$

Also, the problem of hypotheses testing (2.4) remains invariant under $\mathcal{G}$ since $\bar{G}(\Omega) = \Omega$ and $\bar{G}(\Omega_0) = \Omega_0$ where $\Omega = \{(\sigma_1, \cdots, \sigma_s) : \sigma_i > 0, i = 1, \cdots, s\} = \mathbb{R}^{+s}$, $\Omega_0 = \{(\sigma_1, \ldots, \sigma_s) : \sigma_1 = \cdots = \sigma_s\}$ and $\bar{G} = \bar{g}_a(\sigma_1, \cdots, \sigma_s) = a(\sigma_1, \cdots, \sigma_s)$ is the induced group of transformations on the parameter space Ω by the group of transformations $\mathcal{G}$ in Equation (3.1).

Lemma 3.1. *Let $X_1^\star, \cdots, X_r^\star$ be a SOS sample under the CPHR model with the baseline standard exponential distribution and $U_i = X_i^\star/X_r^\star$, for $i = 1, \cdots, r-1$. The joint PDF of $U_i = X_i^\star/X_r^\star$, for $i = 1, \cdots, r-1$, is*

$$f(u_1, \cdots, u_{r-1}) = \left(\prod_{j=1}^{r} \alpha_j\right) \frac{\Gamma(r)}{\left(\Sigma_{j=1}^{r-1} u_j m_j + m_r\right)^r}, \quad 0 < u_1 < \cdots < u_{r-1} < 1.$$

The problem of obtaining uniformly most powerful test among invariant tests, known as *uniformly most powerful invariant* (UMPI) test, remains as an open problem.

3.1 First Suggest

Notice that, the maximal invariant statistic is $M_1 = [[u_{ij}]]$ where $u_{ij} = x_{ij}^\star/x_{ir}^\star$ for $i = 1, \cdots, s$ and $j = 1, \cdots, r-1$. The PDF of the statistic M_1 is easily derived from Lemma 3.1 which is complicated for derivation of the UMPI test (if exists).

3.2 Second Suggest

Another approach is first applying the concept of sufficiency and then deriving the UMPI test based on the sufficient statistics. More precisely, from Equation (2.3), the (complete) sufficient statistics in this case is $\boldsymbol{T} = (T_1, \cdots, T_s)$. The maximal invariant test with respect to the group of scale transformations on $\boldsymbol{T}$-space induced by $\mathcal{G}$ is

$$M_2 = \left(\frac{T_1}{T_s}, \cdots, \frac{T_{s-1}}{T_s} \right).$$

Since $T_1, \cdots, T_s$ are independent, by Equation (2.7), the PDF of the statistic M_2 is easily derived as

$$\begin{aligned} f(m_1, \cdots, m_{s-1}) &= \frac{1}{\Gamma(r)^s} \left(\prod_{i=1}^{s} \frac{1}{\sigma_i} \right)^r \left(\prod_{i=1}^{s-1} m_i \right)^{r-1} \\ &\times \frac{\Gamma(sr)}{[m_1/\sigma_1 + \cdots + m_{s-1}/\sigma_{s-1} + 1/\sigma_s]^{sr}}, \end{aligned} \tag{3.2}$$

for $m_1 > 0, \cdots, m_{s-1} > 0$. By Theorem 6.5.3 of Lehmann and Romano [10], the UMPI test based on the sufficient statistic M_2 is also UMP test among all invariant test based on the original data $\mathbf{x}$ given by (2.1).

Remark 3.2. Note that the observed Fisher Information (FI), denoted by $i(\hat{\sigma}_1, \cdots, \hat{\sigma}_s)$, on the basis of available SOSs data is equal to minus of the Hessian matrix (HM) evaluated at the MLEs of the parameters, i.e.

$$i(\hat{\sigma}_1, \cdots, \hat{\sigma}_s) = [[(-\partial^2 \log(L)/\partial\sigma_i \partial\sigma_j)_{1 \le i,j \le s}]]|_{\sigma_1 = \hat{\sigma}_1, \cdots, \sigma_s = \hat{\sigma}_s}.$$

It is well known that the unique MLEs have asymptotically the multivariate normal distribution with mean vector $(\sigma_1, \cdots, \sigma_s)$ and the variance-covariance matrix $[i(\hat{\sigma}_1, \cdots, \hat{\sigma}_s)]^{-1}$ (see, e.g., Lehmann and Romano, [10], 2005, Chap. 14). Therefore, an approximate equi-tailed confidence interval for σ_i is

$$\left(\hat{\sigma}_i - z_{\gamma/2} \sqrt{\frac{\hat{\sigma}_i^2}{r}} \ , \ \hat{\sigma}_i + z_{\gamma/2} \sqrt{\frac{\hat{\sigma}_i^2}{r}} \right), \tag{3.3}$$

where z_γ stands for the γ-percentile of the standard normal distribution.

References

[1] Bedbur, S. (2010), UMPU Tests based on Sequential order statistics. *Journal of Statistical Planning and Inference*, **140**, 2520-2530.

[2] Beutner, E. and Kamps, U. (2009), Order restricted statistical inference for scale parameters based on sequential order statistics. *Journal of Statistical Planning and Inference*, **139**, 2963-2969.

[3] 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.

[4] 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.

[5] Cramer, E. and Kamps, U. (2001b), Sequential k-out-of-n systems. In Balakrishnan, N. and Rao, E., editors, *Handbook of Statistics, Advances in Reliability*, volume *20*, Chapter 12 , 301-372. .

[6] Esmailian, M and Doostparast, M. (2014), Estimation based on sequential order statistics with random removals. *Probability and Mathematical Statistics*, **34**(1) , 81-95.

[7] 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*, accepted.

[8] Kamps, U. (1995a), *A Concept of Generalized Order Statistics*. Teubner.

[9] Kamps, U. (1995b), A concept of generalized order statistics. *Journal of Statistical Planning and Inference*, *48*, 1-23.

[10] Lehmann, E. L. and Romano, J. (2005), *Testing Statistical Hypothesis*. 3-th edition, Springer, New York.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

The family of lifetime and power-series distributions

Hassantabar Darzi, F. [1] and Eslami, S. M. [2]

[1] Faculty of Mathematics and Statistics , University of Sistan and Baluchestan
[2] Faculty of Mathematics and Statistics, Velayat University

Abstract

In this study,we introduce the family of lifetime distributions by compounding any proper continuous lifetime distribution and truncated power-series (PS) distribution, named the family of lifetime-power series distributions (FPS). This new family of distributions contains several lifetime distributions such as generalized exponential (GE), generalized Weibull (GW), generalized Rayleigh (GR), inverse Weibull (IW) and many continuous distributions by compounding of PS distributions. The FPS distributions have decreasing, increasing, decreasing-increasing failure rate. The properties of the proposed family are discussed. A simple EM-type algorithm for iteratively computing maximum likelihood estimates is presented. A formal equation for Fisher information matrix is derived in order to obtaining the asymptotic covariance matrix. The experimental results are illustrated based on real data sets.

Keywords: Power-series distribution, EM algorithm, Hazard rate, Maximum likelihood estimation.

1 Introduction

Multi-parameter distributions to model lifetime data have been introduced by compounding a continuous lifetime and powerseries distributions. The Exponential Geometric (EG), Exponential Poisson (EP) and exponential logarithmic distributions were introduced and studied by Adamidis

[1]hasantabar_f@math.usb.ac.ir
[2]m.siavashi@cse.shirazu.ac.ir

and Loukas (1998), Kus (2007) and Tahmasbi and Rezaei (2008), respectively. Recently, Chahkandi and Ganjali (2009) introduced the Exponential Power Series (EPS) distributions, which contain these distributions. Lately, Alkarni and Oraby (2012), Alkarni (2012) obtained a class of truncated Poisson and logarithmic distribution with any continuous lifetime distribution. In this article we introduce a new lifetime distributions family, by combining a truncated at zero power series with a class of several lifetime continuous distributions such as exponential, weibull, pareto, generalized exponential and generalized weibull distributions and hence any mixture of continuous lifetime with truncated Power Series distribution becomes a special case of this class. This family includes special cases of lifetime distributions presented by Tahmasbi and Rezaei (2008), Chahkandi and Ganjali (2009), Alkarni and Oraby (2012) and Alkarni (2012). This study is organized as follow. The new class of Power Series lifetime (PSL) distributions with its properties are introduced in section 2.In Section 3 an estimation of the parameters with maximum likelihood via the EM algorithm is presented.The information matrix of the parameters are also obtained.

2 The Family of FPS distribution and its properties

Let $Y_1, \ldots, Y_Z$ be independent and identically distributed (iid) random variables with probability density function (pdf) given by:

$$f_Y(y;\beta) \quad ;\beta = (\beta_1, \ldots, \beta_k), k \geq 1$$

where Z is a random variable from power series distribution and is independent from Z. Power series distribution are the family of distribution with probability function $P_t(z;\theta) = \frac{a_z\theta^z}{A(\theta)}, z \in N \cup \{0\}, \theta > 0, A(\theta) = \sum_z a_z\theta^z$. This family of distributions includes Binomial, Poisson,Geometric, Negative binomial and Logarithmic distribution (vide, Johnson et al., 1993). We can define truncated at zero power series distribution with probability function $P_t(z;\theta) = c\frac{a_z\theta^z}{A(\theta)}I_{\{1,2,\ldots\}}, \theta > 0$ where $c = \frac{A(\theta)}{A(\theta)-a_0}$. Thus $P_t(z;\theta) = \frac{a_z\theta^z}{A(\theta)-a_0}, z \in N, \theta > 0$ and $A(\theta) - a_0 = \sum_{z\in N} a_z\theta^z$.
Let use $X = min\{Y_1, \ldots, Y_z\}$ then $f(x|z;\beta) = zf_\beta(x)[1 - F_\beta(x)]^{z-1}$and the marginal probability density function of X, if Z has truncated at zero power series distribution, is:

$$f(x;\xi) = \frac{-\frac{\partial}{\partial x}A[\theta(1 - F_\beta(x))]}{A(\theta) - a_0} \tag{2.1}$$

where $\xi = (\theta, \beta)$. Hence the cumulative distribution function(cdf) of X is:

$$F(x;\xi) = 1 - \frac{(A[\theta(1 - F_\beta(x))] - a_0)}{A(\theta) - a_0} \tag{2.2}$$

In the sequel we will be referring to the unconditional distribution of X as Family of the Power Series lifetime(FPS) distribution.This new class of distributions generalizes several distributions which have been introduced and studied in the literature.
Survival function (also known as reliability function) and hazard function (also known as failure rate function) of the FPS distribution are given, respectively, by:

$$S(x;\xi) = 1 - F(x;\xi) = \frac{(A[\theta(1 - F_\beta(x))] - a_0)}{A(\theta) - a_0} \tag{2.3}$$

and

$$h(x;\xi) = \frac{f(x;\xi)}{S(x;\xi)} = \frac{-\frac{\partial}{\partial x}(A[\theta(1-F_\beta(x))]-a_0)}{(A[\theta(1-F_\beta(x))]-a_0)} \tag{2.4}$$

3 Special Cases of the FPS distributions

3.1 Generalized Exponential-PS ditribution

The random variable Y has a generalized exponential(GE)distribution(Gupta and Kundu,1999) with parameters α and β if its cumulative distribution function(cdf) takes the form

$$F_Y(y) = (1-e^{-\beta y})^\alpha$$

where $\alpha > 0$, $\beta > 0$.The corresponding probability density function(pdf)is

$$f_Y(y) = \alpha\beta e^{-\beta y}(1-e^{-\beta y})^{\alpha-1}$$

Given Z, Let $X = min\{Y_1,\ldots,Y_Z\}$ and Z respectively be independent and identically distributed (iid) random variables from GE and PS distributions according to cdf(1) pdf(2),the cdf and pdf of X is given by

$$F_X(x;\xi) = 1 - \frac{A(\theta e^-(\beta x)^\alpha)}{A(\theta)-a_0}$$

,

$$f_X(x;\xi) = \alpha\beta e^{-\beta x}(1-e^{-\beta x})^{\alpha-1}$$

that it called Generalized exponential Power Series(GEPS) distribution. The associated hazard rate and survival function are given respectively by

$$h(x) = \frac{-\theta\alpha\beta e^{-\beta x}(1-e^{-\beta x})^{\alpha-1}A'(\theta e^{-(\beta x)^\alpha})}{A(\theta e^{-(\beta x)^\alpha})}$$

$$s(x) = \frac{A(\theta e^{-(\beta x)^\alpha})}{A(\theta)-a_0}$$

Remark. When $\alpha = 1$, we have a distribution which is known as exponential-power series distributions introduce by chahkandi and Ganjali that table1 shows the probability ,cumulative probability, survival and hazard functions.

3.2 Generalized Weibull-PS ditribution

Generalized Weibull(GW)distribution is special case of Life time distributons with parameters α and β if its cumulative distribution function(cdf) takes the form

$$F_Y(y) = (1-e^{-(\beta y)^\gamma})^\alpha$$

where $\alpha > 0$, $\beta > 0$ and $\gamma > 0$.The corresponding probability density function(pdf)is

$$f_Y(y) = \alpha\gamma\beta^\gamma y^{\gamma-1}e^{-(\beta y)^\gamma}(1-e^{-(\beta y)^\gamma})^{\alpha-1}$$

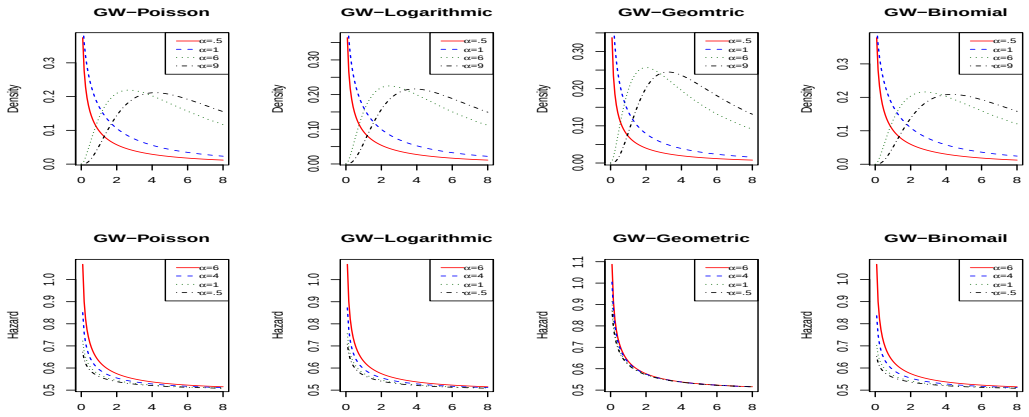


Figure 1: probability density and Hazard function of the GWPS distribution for different value of parameters

Let $X = min\{Y_1, \ldots, Y_Z\}$ and Z respectively be independent and identically distributed (iid) random variables from GW and PS distributions according to cdf(1) pdf(2),the cdf and pdf of X is given by

$$F_X(x) = 1 - \frac{A(\theta e^{-(\beta x)^\gamma})^\alpha}{A(\theta) - a_0}$$

$$f_X(x) = \theta\alpha\gamma\beta^\gamma x^{\gamma-1} e^{-(\beta x y)^\gamma} (1 - e^{-(\beta x)^\gamma})^{\alpha-1} \frac{-A'(\theta e^{-(\beta x)^\gamma})^\alpha}{A(\theta) - a_0}$$

that it called Generalized Weibull Power Series(GWPS) distribution that introduce by Mahmoudi and Shiran(2012). The associated hazard rate and survival function are given respectively by

$$h(x) = \frac{-\theta\alpha\gamma\beta^\gamma x^{\gamma-1} e^{-(\beta x)^\gamma} (1 - e^{-(\beta x)^\gamma})^{\alpha-1} A'(\theta e^{-(\beta x)^\gamma})^\alpha}{A(\theta e^{(\beta x)^\gamma})^\alpha}$$

$$s(x) = \frac{A(\theta e^{-(\beta x)^\gamma})^\alpha}{A(\theta) - a_0}$$

The plots of probability density and hazard rate function of GW-PS for different value of parameters are given in figure2.

Remark. When $\alpha = 1$, we have a distribution which is known as Weibull-power series distributions that table1 shows the probability ,cumulative probability, survival and hazard functions.

3.3 Inverse Weibull-PS ditribution

The generalized Weibull distribution is one of life time distributions with parameters α and λ if its cumulative distribution function(cdf) takes the form$F(y) = e^{-\lambda y^{-\alpha}}$,Where $\alpha > 0, \lambda > 0$.

The corresponding probability density function (pdf) is $f_Y(y) = \alpha\lambda y^{-(\alpha+1)}e^{-\lambda y^{-\alpha}}$ Let $X = min\{Y_1, \ldots, Y_Z\}$ and Z respectively be independent and identically distributed (iid) random variables from IW and PS distributions according to cdf(1) pdf(2),the cdf and pdf of X is given by

$$F_X(x) = 1 - \frac{A(\theta(1 - e^{-\lambda x^{-\alpha}}))}{A(\theta) - a_0}.$$

$$f_X(x) = \theta\alpha\lambda x^{-(\alpha-1)}e^{-\lambda x^{-\alpha}} \frac{-A'(\theta(1 - e^{-\lambda x^{-\alpha}}))}{A(\theta) - a_0}$$

that it called inverse Weibull Power Series(IWPS) distribution. The associated hazard rate and survival function are given respectively by

$$h(x) = \frac{-\theta\alpha\lambda x^{-(\alpha-1)}e^{-\lambda x^{-\alpha}} A'(\theta(1 - e^{-\lambda x^{-\alpha}}))}{A(\theta(1 - e^{-\lambda x^{-\alpha}}))}$$

$$s(x) = \frac{A(\theta(1 - e^{-\lambda x^{-\alpha}})}{A(\theta) - a_0}$$

3.4 Generalized Rayleigh-PS ditribution

The generalized Rayleigh distribution is one of life time distributions with parameters α β γ if its cumulative distribution function(cdf) takes the form$F_Y(y) = (1 - e^{-(\beta y)^2})^{\alpha}$,Where $\alpha > 0, \beta > 0$. The corresponding probability density function (pdf) is $f_Y(y) = 2\alpha\beta^2 e^{-(\beta y)^2}(1 - e^{-(\beta y)^2})^{\alpha-1}$. Let $X = min\{Y_1, \ldots, Y_Z\}$ and Z respectively be independent and identically distributed (iid) random variables from IW and PS distributions according to cdf(1) pdf(2),the cdf and pdf of X is given by

$$F_X(x) = 1 - \frac{A(\theta e^{-(\beta x)^2})^{\alpha}}{A(\theta) - a_0}$$

$$f_X(x) = 2\theta\alpha\beta^2 e^{-(\beta x)^2}(1 - e^{-(\beta x)^2})^{\alpha-1} \frac{-A'(\theta e^{-(\beta x)^2})}{A(\theta) - a_0}$$

that it called inverse Weibull Power Series(IWPS) distribution. The associated hazard rate and survival function are given respectively by

$$h(x) = \frac{-2\theta\alpha\beta^2 e^{-}(\beta x)^2(1 - e^{-(\beta x)^2})^{\alpha-1} A'(\theta e^{-(\beta x)^2})}{A(\theta e^{-(\beta x)^2})^{\alpha}}$$

$$s(x) = \frac{A(\theta e^{-}(\beta y)^2)}{A(\theta) - a_0}$$

The plots of probability density and hazard rate function of GR-PS for different value of parameters are given in figure2.

Table1 shows the the properties of some others lifetime distributions. Some of the other lifetime distributions are excluded from this table such as Gamma and lognormal distributions. Those distributions do not have nice forms although they still can be applied in this class numerically.

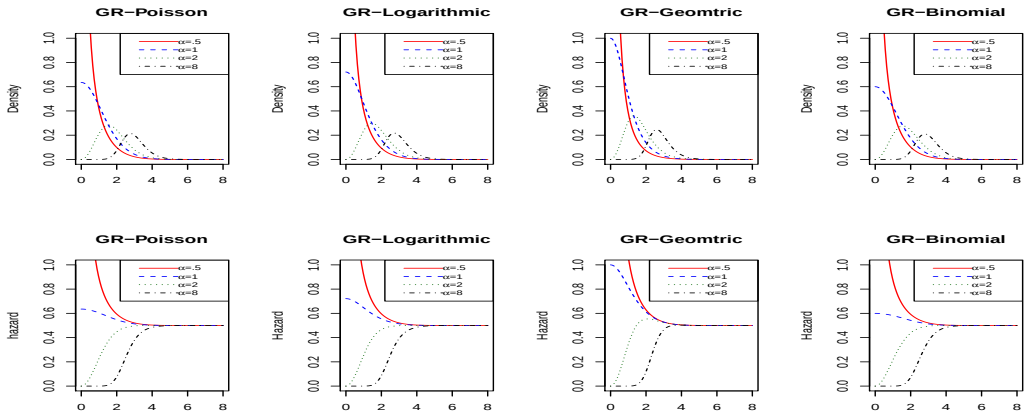


Figure 2: probability density and Hazard function of the GR-PS distribution for different value of parameters

Table 1: PSL distribution and its properties

Distribution	$f_x(x;\xi)$	$F_x(x;\xi)$	$S_x(x;\xi)$	$h_x(x;\xi)$
Exponential	$\frac{-\frac{\partial}{\partial x}A[\theta e^{-\beta x})]}{A(\theta)-a_0}$	$1-\frac{A[\theta e^{-\beta x}]-a_0}{A(\theta)-a_0}$	$\frac{A[\theta e^{-\beta x}]-a_0}{A(\theta)-a_0}$	$\frac{-\frac{\partial}{\partial x}A[\theta e^{-\beta x})]}{A[\theta e^{-\beta x}]-a_0}$
Weibull	$\frac{-\frac{\partial}{\partial x}A[\theta e^{-\beta x^{\alpha}})]}{A(\theta)-a_0}$	$1-\frac{A[\theta e^{-\beta x^{\alpha}})]-a_0}{A(\theta)-a_0}$	$\frac{A[\theta e^{-\beta x^{\alpha}})]-a_0}{A(\theta)-a_0}$	$\frac{-\frac{\partial}{\partial x}A[\theta e^{-\beta x^{\alpha}})]}{A[\theta e^{-\beta x^{\alpha}})]-a_0}$
Pareto	$\frac{-\frac{\partial}{\partial x}A[\theta\beta^{\alpha}x^{-\alpha})]}{A(\theta)-a_0}$	$1-\frac{A[\theta\beta^{\alpha}x^{-\alpha})]-a_0}{A(\theta)-a_0}$	$\frac{A[\theta\beta^{\alpha}x^{-\alpha})]-a_0}{A(\theta)-a_0}$	$\frac{-\frac{\partial}{\partial x}A[\theta\beta^{\alpha}x^{-\alpha})]}{A[\theta\beta^{\alpha}x^{-\alpha})]-a_0}$
Rayleigh	$\frac{-\frac{\partial}{\partial x}A[\theta e^{\frac{x^2}{2\beta^2}}]}{A(\theta)-a_0}$	$1-\frac{A[\theta e^{\frac{x^2}{2\beta^2}}]-a_0}{A(\theta)-a_0}$	$\frac{A[\theta e^{\frac{x^2}{2\beta^2}}]-a_0}{A(\theta)-a_0}$	$\frac{-\frac{\partial}{\partial x}A[\theta e^{\frac{x^2}{2\beta^2}}]}{A[\theta e^{\frac{x^2}{2\beta^2}}]-a_0}$

4 Estimation of the parameters by maximum likelihood and its related EM algorithm

4.1 Estimation by maximum likelihood

In this section we want to obtain the maximum likelihood estimates of parameters. Let $X_1,\ldots,X_n$ be a random sample with observed values $x_1,\ldots,x_n$ from a PSL distribution with parameters $\xi=(\theta,\underline{\beta})$. The log-likelihood function based on the observed sample size of n,$y_{obs}=(x_i,i=1,\ldots,n)$ is obtained by:

$$\ell(\xi;y_{obs}) \quad = \quad -n\ell[A(\theta)-a_0]-n\ell(\theta)+\sum\ell f(x_i;\beta)+\sum\ell A'[\theta(1-F(x_i;\beta))]$$

where $A'[\theta(1-F(x_i;\beta))]=A'(\theta)|_{\theta(1-F(x_i;\beta))}$ and $A'(\theta)$is the derivation of $A(\theta)$ with respect to θ. The partial derivations are found to be

$$\frac{\partial\ell}{\partial\theta}=-\frac{nA'(\theta)}{A(\theta)-a_0}-\frac{n}{\theta}+\sum\frac{[1-F(x_i;\beta)]A''[\theta(1-F(x_i;\beta)]}{A'[\theta(1-F(x_i;\beta))]} \tag{4.1}$$

where $A''[\theta(1-F(x_i;\beta)] = A''(\theta)|_{\theta(1-F(x_i;\beta)}$ and $A''(\theta)$ is the second derivation of $A(\theta)$ with respect to θ and

$$\frac{\partial \ell}{\partial \beta} = \sum \frac{\frac{\partial f(x_i;\beta)}{\partial \beta}}{f(x_i;\beta)} + \sum \frac{-\theta \frac{\partial F(x_i;\beta)}{\partial \beta} A''[\theta(1-F(x_i;\beta)]}{A'[\theta(1-F(x_i;\beta)]} \tag{4.2}$$

The Maximum Likelihood Estimates(MLE) of ξ, say $\hat{\xi}$,is obtained by solving the nonlinear system $U_n(\xi) = 0$. The solution of this nonlinear system of equations has not a closed form, but can be found numerically. For interval estimation and hypothesis tests on the model parameters, we require the information matrix.The $(k+1)\times(k+1)$ information matrix is given by $I_n(\xi)$ where, the elements of $I_n(\xi)$ are the second partial derivatives of (A.5) and (A.7). In the large sample, MLE of ξ can be treated as being approximately multivariate normal with mean ξ and variance-covariance matrix, which is the inverse of the expected information matrix $J_n(\xi) = E(I;\xi)$ and the expectation is to be taken with respect to the distribution of X.

4.2 EM algorithm

Based on the underlying distribution, the maximum likelihood estimation of the parameters can be found analytically using an EM algorithm. NewtonRaphson algorithm is one of the standard methods to determine the MLEs of the parameters. To employ the algorithm, second derivatives of the log-likelihood are required for all iteration. EM algorithm is a very powerful tool in handling the incomplete data problem. It is an iterative method by repeatedly replacing the missing data with estimated values and updating the parameter estimates. It is especially useful if the complete data set is easy to analyze. As pointed out by Little and Rubin (1983), the EM algorithm will converge reliably but rather slowly (as compared to the Newton-Raphson method) when the amount of information in the missing data is relatively large. Recently, EM algorithm has been used by several authors such as Adamidis and Loukas (1998); Adamidis (1999); Chahkandi and Ganjali (2009); Alkarni (2012).
To start the algorithm, hypothetical complete-data distribution defined with density function $f(x,z;\xi) =$. Thus, it is straight forward to verify that the E-step of an EM algorithm requires the computation of the conditional expectation of $(Z|X;\xi)$, where $\xi^{(h)} = (\theta^{(h)}, \underline{\beta^{(h)}})$ is the current estimate of ξ. Using $f(z|x;\xi) = za_z\theta f_\beta(x)[\theta(1-F_\beta(x))]^{z-1}[-\frac{\partial}{\partial x}A[\theta(1-F_\beta(x))]^{-1}$, we have

$$E(Z|X;\xi) = 1 + \frac{\theta(1-F_\beta(x))A''[\theta(1-F_\beta(x))]}{A'[\theta(1-F_\beta(x))]}$$

. The EM algorithm is completed with M-step, which is complete data maximum likelihood overξ with the missing $Z's$ replaced by their conditional expectations. Thus an EM iteration

$$\frac{\theta^{(h+1)}A'(\theta^{(h+1)})}{A(\theta^{(h+1)} - a_0)} = \frac{\sum z_i^{(h)}}{n}$$

$$\sum \frac{\partial \ln f(x_i;\beta^{(h+1)})}{\partial \beta^{(h+1)}} = \sum (1 - z_i^{(h)}) \frac{\partial}{\partial \beta^{(h+1)}} \ln[1 - F(x_i;\beta^{(h+1)})]$$

where $z_i^{(h)} = 1 + \frac{\theta^{(h)}(1-F(x;\beta^{(h)}))A''[\theta^{(h)}(1-F(x;\beta^{(h)})]}{A'[\theta^{(h)}(1-F(x;\beta^{(h)}))]}$.
It can be seen that only a one-dimensional search such as NewtonRaphson is required for M-step of an EM cycle.

4.3 Asymptotic variance and covariance of MLEs

Applying the usual large sample approximation, MLE of ξ can be treated as begin approximately bivariate normal with mean ξ and variance-covariance matrix, which is the inverse of the expected information matrix $J(\xi) = E(I;\xi)$, where $I = I(\xi; y_{obs})$ is the observed information matrix with elements $I_{ij} = \frac{-\partial^2 \ell}{\partial \xi_i \partial \xi_j}$ with $i, j = 1, 2, \ldots$ and the expectation is to be taken with respect to the distribution of X. Differentiating (5) and (6), the elements of the symmetric, second-order observed information matrix are found to be

$$\begin{aligned} I_{11} &= \frac{\partial^2 \ell}{\partial \theta^2} = -\frac{nA''(\theta)[A(\theta)-a_0] - n[A'(\theta)]^2}{[A(\theta)-a_0]^2} + \frac{n}{\theta^2} \\ &- \sum \frac{[1-F(x_i;\beta)]^2 A'''[\theta(1-F(x_i;\beta)] - [1-F(x_i;\beta)]^2[A''[\theta(1-F(x_i;\beta))]]^2}{A'[\theta(1-F(x_i;\beta))]} \end{aligned}$$

$$\begin{aligned} I_{12} = I_{21} &= -\sum \frac{\partial F(x_i;\beta)}{\partial \beta} \frac{[A''[\theta(1-F(x_i;\beta)] - \theta(1-F(x_i;\beta)A'''[\theta(1-F(x_i;\beta))]]}{A'[\theta(1-F(x_i;\beta))]} \\ &- \frac{\frac{\partial F(x_i;\beta)}{\partial \beta}(1-F(x_i;\beta))[A''[\theta(1-F(x_i;\beta))]]^2}{[A'[\theta(1-F(x_i;\beta))]]^2} \end{aligned}$$

$$\begin{aligned} I_{22} &= -\theta \sum \frac{\frac{\partial^2 F(x_i;\beta)}{\partial \beta^2} A''[\theta(1-F(x_i;\beta)] - \theta(\frac{\partial F(x_i;\beta)}{\partial \beta})^2 A'''[\theta(1-F(x_i;\beta))]}{A'[\theta(1-F(x_i;\beta))]} \\ &+ \theta[\frac{\frac{\partial F(x_i;\beta)}{\partial \beta} A''[\theta(1-F(x_i;\beta))]}{A'[\theta(1-F(x_i;\beta))]}]^2 \end{aligned}$$

The elements of the expected information matrix $J(\xi)$ are calculated by taking the expectations of I_{ij}, i,j=1,2 with respect to the distribution of X, i.e. this expectation is required: $E(lnA\prime(\theta e^{-}\beta x_i))$. we can calculate this expectation for each element of power-series distributions.when we obtain the expectations of I_{ij}, i,j=1,2 we would have the matrix $J(\xi)$, the inverse of $J(\xi)$, evaluated at $\hat{\xi}$ provides the asymptotic variance-covariance matrix of MLEs. Alternative estimates can be obtained from the inverse of the observed information matrix since it is a consistent estimator of $J^{-1}(\xi)$.

5 Application of the F-PS distribution

In this section, we fit FPS distributions to real set. The data set is given by Leiblein and Zelen (1956) on the endurance test of deep groove ball bearings discussed , which include the number of

Table 2: parameter estimates(with std.),K-S statistics, p-value, AIC and BIC for Leiblein and Zelen data.

Distribution	$MLE(std)$	$K-S$	$P-Value$	AIC	BIC
GE-P	$\hat{\alpha}=4.72439289, \hat{\beta}=0.03088544, \hat{\theta}=-0.12983222$	0.099518	0.972	214.2892	217.4228
GE-L	$\hat{\alpha}=4.68270010, \hat{\beta}=0.03165577, \hat{\theta}=-0.37921442$	0.10167	0.9663	214.2862	217.4198
GE-G	$\hat{\alpha}=4.71243089, \hat{\beta}=0.03108482, \hat{\theta}=-0.09453810$	0.10012	0.9705	214.2883	217.4219
GE-B	$\hat{\alpha}=0.77753509, \hat{\beta}=0.02980041, \hat{\theta}=-1.15228085$	0.094042	0.9836	214.2617	217.3952
GW-P	$\hat{\alpha}=1.219593882, \hat{\beta}=0.002151981, \hat{\theta}=0.919003139$ $\hat{\gamma}=0.98997699$	0.1437	0.9162	189.275	190.9698
GW-L	$\hat{\alpha}=1.191146389, \hat{\beta}=0.002246758, \hat{\theta}=0.490502858$ $\hat{\gamma}=1.3214322$	0.1428	0.9198	189.2264	190.9212
GW-G	$\hat{\alpha}=1.3888842, \hat{\beta}=0.0023155981, \hat{\theta}=0.534503249$ $\hat{\gamma}=0.98997699$	0.1443	0.9153	189.269	190.9598
GW-B	$\hat{\alpha}=1.191906470, \hat{\beta}=0.002212253, \hat{\theta}=0.201947884$ $\hat{\gamma}=1.58339$	0.1462	0.9063	189.3061	191.001
GR-P	$\hat{\alpha}=1.31833516, \hat{\beta}=0.01082137, \hat{\theta}=1.66560706$	0.11583	0.9104	389.1759	392.3095
GR-L	$\hat{\alpha}=1.83379903, \hat{\beta}=0.01059785, \hat{\theta}=0.95765106$	0.090961	0.9884	388.6836	391.8172
GR-G	$\hat{\alpha}=1.4824384, \hat{\beta}=0.0103635, \hat{\theta}=0.6954085$	0.10344	0.9611	388.9392	392.0728
GR-B	$\hat{\alpha}=1.27819385, \hat{\beta}=0.01103999, \hat{\theta}=0.42607813$	0.11948	0.891	389.2563	392.3898
IW-P	$\hat{\alpha}=0.3776692, \hat{\lambda}=27.7744343, \hat{\theta}=-204.2402199$	0.10613	0.972	214.2355	217.3691
IW-L	$\hat{\alpha}=2.4747970, \hat{\lambda}=4835.1386466, \hat{\theta}=0.9517133$	0.15134	0.7219	217.9702	221.1038
IW-G	$\hat{\alpha}=0.4284007, \hat{\lambda}=41.8460927, \hat{\theta}=-1146.7264757$	0.092237	0.9941	217.1235	220.257
IW-B	$\hat{\alpha}=0.8477569, \hat{\lambda}=77.4578973, \hat{\theta}=-1.2914360$	0.087796	0.9969	214.9114	218.045

revolutions (in millions) to failure for each of n=23 ball bearings under the life test. The MLEs of the parameters (with std), the maximized log-likelihood, the kolmogorov-smirnov statistic with its respective p-value, the AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) for the GEPS,GWPS, GRPS and IWPS models are given in Table1. The results for the first models,GE-PS, show that the GE-P distribution yields the best fit among the GE-L,GE-G and GE-B. For other model,we can choose the best fit by the K-S test, The K-S test statistic takes the smallesd value with the largest value of its respective p-value for GR-L and IW-B distribution. Also this conclusion is confirmed from the values of the AIC and BIC for the fitted models given in Table1.

References

[1] Adamidis, K. (1999),*An EM algorithm for estimating negetive binomial parameters*, Austral. New Zeland J. Stat., 41(2), 213-221.

[2] Adamidis, A., Loukas,S. (1998), *A lifetime distribution with decreasing failure rate*, Stat. Prob. Lett., 39, 35-42.

[3] Alkarni, S. and A. Oraby, (2012). *A compound class of poisson and lifetime distributions.* J. Stat. Applied Pro., 1: 45-51.

[4] Alkarni, S. (2012). *A new family of logarithmic lifetime distributions.* J. Mathe. stat., 8(4),435-440.

[5] Barlow, R.E., Marshall, A.W., Proschan, F. (1963),*Properties of probability distributions with monotone hazard rate*, Ann. Math. Statist.,34,375-389.

[6] Chahkandi, M., Ganjali, M., (2009). *On some lifetime distributions with decreasing failure rate.* Comput. Stat,. Data Anal., 53, 44334440.

[7] Kus, C., (2007), *A new lifetime distribution.* Comput. Stat., Data Anal., 51(9), 4497-4509.

[8] Mahmoudi, E., Jafari, A. A. (2012). *Generalized exponentialpower series distributions.* Computational Statistics and Data Analysis, 56(12), 4047-4066.

[9] Tahmasbi, R., Rezaei, S., (2008). *A two-parameter lifetime distribution with decreasing failure rate.* Comput. Stat., Data Anal., 52: 3889-3901.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Estimation of stress-strength reliability for the Pareto distribution in the presence of outliers

Jabbari Nooghabi, M. [1]

Department of Statistics, Ferdowsi University of Mashhad

Abstract

The present paper discussed the problem of estimating the stress-strength reliability parameter $R = P(X < Y)$, where X and Y follow the Pareto distribution, independently. Assuming X has the Pareto distribution in the presence of outliers and the random variable Y is not contaminated with outliers. Different estimators of R were derived when only the shape parameters are unknown. In particular, we obtain the maximum likelihood, moment method and least squares estimator of the unknown parameters and derived estimates of R. Also, in each situation the shrinkage estimation of the stress-strength reliability parameter is derived by using a prior guess R_0. Monte Carlo simulation study has been used to compare the different methods of estimation in any situation.

Keywords: Shrinkage estimation, Pareto distribution, Maximum likelihood estimator, Moment method, Least squares, Stress strength model, Outliers.

1 Introduction

The problem of estimating $R = P(X < Y)$ has been considered by several authors. Rezaei et al. [11] prepared a history of work on this problem in parametric and non-parametric methods when the variables are followed different types of distribution. A primary search shows that there are only little papers for estimating R when outliers affected on the variable(s). But, in any actual

[1] jabbarinm@um.ac.ir

example in the reliability and stress-strength area at least one of the variables may contaminated by outliers. Because there are noise in the process or in the life testing. In some application of R, we should obtain the treatment effect for a set of response variables which the statistical units are divided by two groups as experiment and control because of removing any other unsuitable effects. In this situation, may be some observations of the response variable (say k of n) are followed another distribution ie. data are contaminated by outliers. For more details refer to Jabbari Nooghabi [6]. Gunasekera [5] considered that X and Y follow independently two-parameter Pareto distribution and proposed several generalized variable method to estimate the reliability parameter. Based on the p-value as a basis for hypothesis testing, he investigated the generalized size, generalized adjusted and unadjusted powers of the test and generalized coverage probabilities by using the simulation study. Mahmoud et al. [8] estimated R, when the variables follow the Lomax distribution with common scale parameter. They presented the estimators of the reliability parameter in the situation where the stress measurements and the strength measurements are both in terms of records. Rezaei et al. [11] have been considered the estimation of R when X and Y are two independent generalized Pareto distributions with different parameters. They obtained the maximum likelihood estimator of R and its asymptotic distribution and also asymptotic confidence interval of it. Afzal Beg and Singh [2] have estimated $R = P(Y < X)$ where X and Y are distributed from the Pareto distribution using censored samples. They have used the minimum variance s-unbiased estimator, Bayesian and maximum likelihood methods. Odat [10] derived the maximum likelihood estimator of $P(X > Y)$ and its asymptotical distribution when X and Y are two independent Pareto distribution. Wong [12] used the interval estimation of $P(Y < X)$ when X and Y are independently followed generalized Pareto distribution with a common scale parameters. Masoom Ali and Woo [9] considered estimation of the reliability for two independent exponentiated Pareto distributions.

In addition, according to Jabbari Nooghabi and Khaleghpanah Nooghabi [7], the way of coping with outliers was ignoring the cases which were out of data range, so it leads losing information.

In this paper, we assume that the response observations for the experiment group have "good" and outlier data and the other observations for control group have only "good" data. Different methods of estimation of $R = P(X < Y)$ when X follows the Pareto distribution in the presence of outliers and Y independently has the homogenous case of the Pareto distribution are derived. We assume that all the parameters of the two Pareto distribution are different instead of the other researches which are mentioned above. The estimators are obtained when only the shape parameters are unknown. Therefore, we derive the maximum likelihood, moment method and least squares estimators the unknown parameters and obtain estimator of R. The different proposed methods have been compared using Monte Carlo simulations and their results have been discussed.

2 Main results

Assuming an sample of size n from X is taken such that k of them distributed from the Pareto distribution with pdf

$$f_2(x;\alpha,\beta,\theta) = \frac{\alpha(\beta\theta)^\alpha}{x^{\alpha+1}}, \qquad 0 < \beta\theta \le x, \quad \alpha > 0,\ \beta > 1,\ \theta > 0, \tag{2.1}$$

and the remaining $(n-k)$ random variables are distributed as:

$$f_1(x;\alpha,\theta)=\frac{\alpha\theta^{\alpha}}{x^{\alpha+1}},\qquad 0<\theta\leq x,\quad \alpha>0. \tag{2.2}$$

Then, according to the model of outliers the joint pdf of $(X_1,X_2,...,X_n)$ in the presence of k outliers is given by

$$\begin{aligned} f(x_1,x_2,..,x_n;\alpha,\beta,\theta) &= \frac{\alpha^n\theta^{n\alpha}\beta^{k\alpha}}{C(n,k)}(\prod_{i=1}^{n}x_i)^{-(\alpha+1)} \\ &\times \sum_{A_1=1}^{n-k+1}\sum_{A_2=A_1+1}^{n-k+2}\ldots\sum_{A_k=A_{k-1}+1}^{n}\prod_{j=1}^{k}\mathbf{I}(x_{A_j}-\beta\theta), \end{aligned} \tag{2.3}$$

where $\mathbf{I}$ is the indicator function. Also, the marginal pdf of X can be obtained as:

$$f(x;\alpha,\beta,\theta)=b\frac{\alpha(\beta\theta)^{\alpha}}{x^{\alpha+1}}\mathbf{I}(x-\beta\theta)+\bar{b}\frac{\alpha\theta^{\alpha}}{x^{\alpha+1}}\mathbf{I}(x-\theta),\quad \alpha>0,\ \beta>1,\ \theta>0, \tag{2.4}$$

where $(X_1,X_2,...,X_n)$ are not independent (refer to Dixit and Jabbari Nooghabi [3, 4] and Jabbari Nooghabi and Khaleghpanah Nooghabi [7]). In the next section, the stress-strength reliability parameter is derived.

2.1 The reliability parameter

Let X follows the Pareto distribution in the presence of outliers which is defined in the pervious section and Y be an rv which are independently from X distributed from the homogenous case of the Pareto distribution, ie.

$$f(y;\nu,\lambda)=\frac{\nu\lambda^{\nu}}{y^{\nu+1}}\mathbf{I}(y-\lambda),\quad \nu>0,\ \lambda>0. \tag{2.5}$$

So $R=P(X<Y)$ based on an sample of size n and m respectively from X and Y is as follows.

$$R=P(X<Y)=1-\frac{\nu}{\alpha+\nu}(b\beta^{\alpha}+\bar{b})\left(\frac{\theta}{\lambda}\right)^{\alpha}. \tag{2.6}$$

Now, we would like to estimate R when only the shape parameter of the two random variables are unknown.

2.2 Estimate R when only the shape parameters are unknown

Based on the application of the Pareto distribution, it is reasonable to assume that the shape parameter of the Pareto distribution is unknown and the threshold parameter is known. For example, in a motor insurance, a claim of at least θ as a compensation can be made and claims below it are not entertained. So, the parameter θ is known and we can fit the Pareto distribution

with parameters α and θ to the data of claims, where θ is known and α is unknown. For more details refer to Dixit and Jabbari Nooghabi [3]. Therefore in the subsection, we assume that the shape parameters α and ν are unknown and the parameters β, θ and λ are known. The ML estimator of the shape parameters of the two random variables are as follows.

$$\hat{\alpha}_{ml1} = \frac{n}{\sum_{i=1}^{n} \ln(X_i) - n\ln(\theta) - k\ln(\beta)}, \qquad \sum_{i=1}^{n} \ln(X_i) > \ln(\theta^n \beta^k), \tag{2.7}$$

and

$$\hat{\nu}_{ml1} = \frac{m}{\sum_{i=1}^{m} \ln(Y_i) - m\ln(\lambda)}, \qquad \sum_{i=1}^{m} \ln(Y_i) > \ln(\lambda^m). \tag{2.8}$$

So, according to the invariant property of the ML estimators, the MLE of R is

$$\hat{R}_{ml1} = 1 - \frac{\hat{\nu}_{ml1}}{\hat{\alpha}_{ml1} + \hat{\nu}_{ml1}} (b\beta^{\hat{\alpha}_{ml1}} + \bar{b}) \left(\frac{\theta}{\lambda}\right)^{\hat{\alpha}_{ml1}}. \tag{2.9}$$

Now, by using the guess R_0, the first shrinkage estimator of R can be obtained when the mean square error (MSE) of the estimator is became minimum. So, we use $\tilde{R}_{11} = \alpha_{11}\hat{R}_{ml1} + (1-\alpha_{11})R_0$ where R_0 is a prior estimate. Therefore, α_{11} is estimated such that minimized the MSE of $\tilde{R}_{11}$, ie.

$$MSE(\tilde{R}_{11}) = E(\tilde{R}_{11} - R)^2 = E[(\alpha_{11}\hat{R}_{ml1} + (1-\alpha_{11})R_0) - R]^2, \tag{2.10}$$

where

$$\alpha_{11} = \frac{(R - R_0)[E(\hat{R}_{ml1} - R_0)]}{E(\hat{R}^2_{ml1}) - 2R_0E(\hat{R}_{ml1}) + R_0^2}, \quad 0 \le \alpha_{11} \le 1. \tag{2.11}$$

One should note that this value depends on the unknown parameter R. So substituting $\hat{R}_{ml1}$ instead of R, implies that

$$\hat{\alpha}_{11} = \frac{(\hat{R}_{ml1} - R_0)[E(\hat{R}_{ml1} - R_0)]}{E(\hat{R}^2_{ml1}) - 2R_0E(\hat{R}_{ml1}) + R_0^2}. \tag{2.12}$$

Therefore, the first shrinkage estimator of R is found as

$$\tilde{R}_{11} = \hat{\alpha}_{11}\hat{R}_{ml1} + (1-\hat{\alpha}_{11})R_0. \tag{2.13}$$

Now, we have to find $E(\hat{R}_{ml1})$ and $E(\hat{R}^2_{ml1})$.

Theorem 2.1. *$E(\hat{R}_{ml1})$ and $E(\hat{R}^2_{ml1})$ are as follows.*

$$\begin{aligned} E(\hat{R}_{ml1}) &= 1 - \frac{2}{\Gamma(n)\Gamma(m)}\{b\sum_{j=0}^{\infty}(-1)^j\sum_{i=0}^{j}C(j,i)(n\alpha)^{\frac{n+i}{2}}[A_1(\beta,\theta,\lambda)]^{\frac{n-i}{2}} \\ &\times BesselK\left(-n+i, 2\sqrt{n\alpha A_1(\beta,\theta,\lambda)}\right)\sum_{l=0}^{j-i}C(j-i,l)(-1)^{j-i-l}(m\nu)^{l+1}\Gamma(m-1-l) \\ &+ \bar{b}\sum_{j=0}^{\infty}(-1)^j\sum_{i=0}^{j}C(j,i)(n\alpha)^{\frac{n+i}{2}}[A_2(\theta,\lambda)]^{\frac{n-i}{2}} \\ &\times BesselK\left(-n+i, 2\sqrt{n\alpha A_2(\theta,\lambda)}\right)\sum_{l=0}^{j-i}C(j-i,l)(-1)^{j-i-l}(m\nu)^{l+1}\Gamma(m-1-l)\}, \end{aligned} \tag{2.14}$$

and

$$\begin{aligned} E(\hat{R}^2_{ml1}) &= 1 - 2E(\hat{R}_{ml1}) + \frac{2}{\Gamma(n)\Gamma(m)}\{b^2\sum_{j=0}^{\infty}(-1)^j(j+1)\sum_{i=0}^{j}C(j,i)(n\alpha)^{\frac{n+i}{2}}[2A_1(\beta,\theta,\lambda)]^{\frac{n-i}{2}} \\ &\times BesselK\left(-n+i, 2\sqrt{2n\alpha A_1(\beta,\theta,\lambda)}\right)\sum_{l=0}^{j-i}C(j-i,l)(-1)^{j-i-l}(m\nu)^{l+2}\Gamma(m-2-l) \\ &+ 2b\bar{b}\sum_{j=0}^{\infty}(-1)^j(j+1)\sum_{i=0}^{j}C(j,i)(n\alpha)^{\frac{n+i}{2}}[A_3(\beta,\theta,\lambda)]^{\frac{n-i}{2}} \\ &\times BesselK\left(-n+i, 2\sqrt{n\alpha A_3(\beta,\theta,\lambda)}\right)\sum_{l=0}^{j-i}C(j-i,l)(-1)^{j-i-l}(m\nu)^{l+2}\Gamma(m-2-l) \\ &+ \bar{b}^2\sum_{j=0}^{\infty}(-1)^j(j+1)\sum_{i=0}^{j}C(j,i)(n\alpha)^{\frac{n+i}{2}}[2A_2(\theta,\lambda)]^{\frac{n-i}{2}} \\ &\times BesselK\left(-n+i, 2\sqrt{n\alpha 2A_2(\theta,\lambda)}\right)\sum_{l=0}^{j-i}C(j-i,l)(-1)^{j-i-l}(m\nu)^{l+2}\Gamma(m-2-l)\}, \end{aligned} \tag{2.15}$$

where $A_1(\beta,\theta,\lambda) = [\ln(\lambda) - \ln(\beta\theta)]$, $A_2(\theta,\lambda) = [\ln(\lambda) - \ln(\theta)]$, $A_3(\beta,\theta,\lambda) = 2\ln(\lambda) - 2\ln(\theta) - \ln(\beta)$ and $BesselK$ is the Bessel function of the second kind (see Abramowitz and Stegun [1]).

Proof. Proof is obvious by using some elementary algebra. □

To get the second and third shrinkage estimators of R, we shall use the generalized likelihood ratio test (GLRT) for testing $H_0 : R = R_0$ vs. $H_1 : R = R_1$. Then the p-value of the test and its

square root can be estimators of the weight to obtain the second and third estimators of R.
The GLRT for testing H_0 vs. H_1 is of the form: reject H_0 when $\Lambda(x,y) < c_1$ or $\Lambda(x,y) > c_2$, where

$$\Lambda(x,y) = \frac{\sup_{H_0} L(\alpha,\nu)}{\sup_H L(\alpha,\nu)}, \quad \text{and} \quad L(\alpha,\nu) \propto \frac{\alpha^n \theta^{n\alpha} \beta^{k\alpha}}{C(n,k)} (\prod_{i=1}^{n} x_i)^{-(\alpha+1)} \nu^m \lambda^{m\nu} (\prod_{i=1}^{m} y_i)^{-(\nu+1)}.$$

It is obvious that $H_0 : R = R_0$ is equivalent to $H_0 : \nu = \frac{\alpha(1-R_0)}{(b\beta^\alpha+\bar{b})\left(\frac{\theta}{\lambda}\right)^\alpha-(1-R_0)}$ and to find the maximum likelihood estimator of α and ν under H_0, one should replace the equation in the likelihood function and maximized it respect to α and ν, respectively. It is not possible to find a closed form of the ML estimator of α and ν under H_0, but we can solve the following equation numerically to obtain the MLE of α under H_0.

$$\begin{aligned} h(\alpha) &= \frac{n}{\alpha} + k\ln(\beta) + n\ln(\theta) - \sum_{i=1}^{n} \ln(x_i) + \frac{m}{\alpha} \\ &- \frac{m[b\beta^\alpha \ln(\beta)\left(\frac{\theta}{\lambda}\right)^\alpha + \left(\frac{\theta}{\lambda}\right)^\alpha \ln\left(\frac{\theta}{\lambda}\right)(b\beta^\alpha+\bar{b})] + (1-R_0)}{(b\beta^\alpha+\bar{b})\left(\frac{\theta}{\lambda}\right)^\alpha - (1-R_0)} \\ &- \frac{[b\beta^\alpha \ln(\beta)\left(\frac{\theta}{\lambda}\right)^\alpha + \left(\frac{\theta}{\lambda}\right)^\alpha \ln\left(\frac{\theta}{\lambda}\right)(b\beta^\alpha+\bar{b})]\alpha(1-R_0)[m\ln(\lambda) - \sum_{i=1}^{m}\ln(y_i)]}{[(b\beta^\alpha+\bar{b})\left(\frac{\theta}{\lambda}\right)^\alpha - (1-R_0)]^2} = 0. \quad (2.16) \end{aligned}$$

Then, to get the MLE of ν under H_0 we should substitute the solution of Equation (18) in $\nu = \frac{\alpha(1-R_0)}{(b\beta^\alpha+\bar{b})\left(\frac{\theta}{\lambda}\right)^\alpha-(1-R_0)}$. Therefore, the GLRT is derived based on a numerical method.
So, the second shrinkage estimation of R is taken by using the following formula

$$\tilde{R}_{21} = \alpha_{21}\hat{R}_{ml1} + (1-\alpha_{21})R_0, \quad (2.17)$$

where $(1-\alpha_{21})$ is the p-value of the GLRT. Also, to find the third shrinkage estimator of R the square root of the p-value is used.

$$\tilde{R}_{31} = \alpha_{31}\hat{R}_{ml1} + (1-\alpha_{31})R_0, \quad (2.18)$$

where $(1-\alpha_{31}) = \sqrt{\text{p}-\text{value}}$.
Now, the moment estimator (MM) of α and ν is obtained by using the first moment of X (or $\frac{1}{X}$) and Y (or $\frac{1}{Y}$), respectively. Therefore,

$$\hat{\alpha}_{mm1} = \begin{cases} \frac{\mu_1}{\mu_1 - \theta(b\beta+\bar{b})} & \alpha > 1, \\ \frac{\mu_1'}{\theta^{-1}(b\beta^{-1}+\bar{b})-\mu_1'} & \alpha \le 1, \end{cases} \quad (2.19)$$

where $\mu_1 = E(X) = \frac{\alpha}{\alpha-1}\theta(b\beta+\bar{b})$ and $\mu_1' = E(X^{-1}) = \frac{\alpha}{\alpha+1}\theta^{-1}(b\beta^{-1}+\bar{b})$. Also

$$\hat{\nu}_{mm1} = \begin{cases} \frac{\delta_1}{\delta_1-\lambda} & \nu > 1, \\ \frac{\delta_1'}{\lambda^{-1}-\delta_1'} & \nu \le 1, \end{cases} \quad (2.20)$$

where $\delta_1 = E(Y) = \frac{\nu}{\nu-1}\lambda$ and $\delta_1' = E(Y^{-1}) = \frac{\nu}{\nu+1}\lambda^{-1}$.
So, the moment estimator of R (we say) is

$$\hat{R}_{mm1} = 1 - \frac{\hat{\nu}_{mm1}}{\hat{\alpha}_{mm1} + \hat{\nu}_{mm1}}(b\beta^{\hat{\alpha}_{mm1}} + \bar{b})\left(\frac{\theta}{\lambda}\right)^{\hat{\alpha}_{mm1}}. \tag{2.21}$$

Also, same as the first shrinkage estimator of R respect to the MLE, the shrinkage estimator of R related to the moment method is

$$\tilde{R}_{41} = \hat{\alpha}_{41}\hat{R}_{mm1} + (1 - \hat{\alpha}_{41})R_0, \tag{2.22}$$

where

$$\hat{\alpha}_{41} = \frac{(\hat{R}_{mm1} - R_0)[E(\hat{R}_{mm1} - R_0)]}{E(\hat{R}^2_{mm1}) - 2R_0E(\hat{R}_{mm1}) + R_0^2}.$$

Further, similarly to the MM we can find the shrinkage estimator of R related to the least squares estimator (LSE). The LSE of α and ν can be derived by using the reliability function of X and Y, respectively. So,

$$\hat{\alpha}_{ls1} = \frac{\sum_{i=1}^n z_{x_i}\ln(x_i) - n\bar{z}_x\overline{\ln(x)}}{\sum_{i=1}^n[\ln(x_i) - \overline{\ln(x)}]^2}, \tag{2.23}$$

and

$$\hat{\nu}_{ls1} = \frac{\sum_{j=1}^m z_{y_j}\ln(y_j) - m\bar{z}_y\overline{\ln(y)}}{\sum_{j=1}^m[\ln(y_j) - \overline{\ln(y)}]^2}, \tag{2.24}$$

where $z_{x_i} = -\ln(1 - F_X(x_i)) = -\ln\left(1 - \frac{i}{n+1}\right)$, $i = 1, 2, ..., n$, $\bar{z}_x = \frac{1}{n}\sum_{i=1}^n z_{x_i}$, $\overline{\ln(x)} = \frac{1}{n}\sum_{i=1}^n \ln(x_i)$, $z_{y_j} = -\ln(1 - F_Y(y_j)) = -\ln\left(1 - \frac{j}{m+1}\right)$, $j = 1, 2, ..., m$, $\bar{z}_y = \frac{1}{m}\sum_{j=1}^m z_{y_j}$ and $\overline{\ln(y)} = \frac{1}{m}\sum_{j=1}^m \ln(y_j)$.
Therefore, the LS estimator of R (we say) is

$$\hat{R}_{ls1} = 1 - \frac{\hat{\nu}_{ls1}}{\hat{\alpha}_{ls1} + \hat{\nu}_{ls1}}(b\beta^{\hat{\alpha}_{ls1}} + \bar{b})\left(\frac{\theta}{\lambda}\right)^{\hat{\alpha}_{ls1}}. \tag{2.25}$$

Finally, the shrinkage estimator of R related to the LS is derived.

$$\tilde{R}_{51} = \hat{\alpha}_{51}\hat{R}_{ls1} + (1 - \hat{\alpha}_{51})R_0, \tag{2.26}$$

where

$$\hat{\alpha}_{51} = \frac{(\hat{R}_{ls1} - R_0)[E(\hat{R}_{ls1} - R_0)]}{E(\hat{R}^2_{ls1}) - 2R_0E(\hat{R}_{ls1}) + R_0^2}.$$

3 Numerical study and comparison the results

In this section, we perform a numerical study to illustrate the performance of the proposed estimators. The following values of the parameters are used in this numerical study. The sample size which is taken from random variable X is 6 and 10. The number of outliers in X random variable is taken to be 1 and 3. The sample size of random variable Y is 10 and 30. The true value of R is taken to be 0.5 and 0.8 and the initial estimate of R (ie R_0) is taken to be 0.35, 0.5 and 0.65 when R=0.5 and 0.65, 0.8 and 0.95 when R=0.8.
Based on one thousand replication of sampling form the distributions with different values of the parameters, the simulation are done. Here, by using the above values of R, we obtain the values of the shape parameter of random variable Y such as:

$$\nu = \frac{\alpha(1-R_0)}{(b\beta^{\alpha}+\bar{b})\left(\frac{\theta}{\lambda}\right)^{\alpha}-(1-R_0)}. \tag{3.1}$$

The estimators are calculated and the bias and mean square error (MSE) of the estimators and shrinkage estimators are found. The bias and MSE (in the parenthesis) of the estimators are shown in Tables 1.
In this research, we have addressed the problem of estimating the reliability parameter R for the Pareto distributions when the strength variable is contaminated with outliers in three cases. According to Table 1, it has been seen that the MLE of R has the least MSE among the MM and LS estimators. Also, the MSE of the estimators is decreasing respect to n, k and m. For the shrinkage estimators, it is obvious that the shrinkage estimators are more efficient than the classic estimators specially for small sample size. The efficiency of the first shrinkage estimator related to the MLE ($\tilde{R}_{11}$) is more than the other types of shrinkage estimators. In general, it has been seen that the shrinkage estimators can be arranged in terms of overall performance as follows (from best to worst); $\tilde{R}_{11} \to \tilde{R}_{21} \to \tilde{R}_{31} \to \tilde{R}_{51} \to \tilde{R}_{41}$.
Our result is same as the approach of Afzal Beg and Singh [2] which is the performance of ML estimator of the reliability parameter is better than the other classical estimator (in that paper minimum variance unbiased estimator). Also, our result is related with the result of Rezaeia et al. [11] which is discussed the estimator of R in the generalized Pareto distribution in the case that the maximum likelihood estimator works quite well.

Table 1. Biases and MSEs (in the parenthesis) of the estimators
when only the shape parameters are unknown for α=3, β=1.5, θ=1 and λ=1.

n	k	m	R	R_0	$\hat{R}_{ml1}$	$\tilde{R}_{11}$	$\tilde{R}_{21}$	$\tilde{R}_{31}$	$\hat{R}_{mm1}$	$\tilde{R}_{41}$	$\hat{R}_{ls1}$	$\tilde{R}_{51}$
6	1	10	0.5	0.35	-0.238841	-0.170210	-0.237053	-0.237948	-0.274981	-0.190513	-0.141050	-0.145051
					(0.128482)	(0.042882)	(0.064461)	(0.068450)	(0.171513)	(0.088256)	(0.143945)	(0.071175)
6	1	10	0.5	0.5	-0.072812	-0.021634	-0.025048	-0.042703	-0.151317	-0.362226	0.011885	-0.004681
					(0.084060)	(0.019401)	(0.028974)	(0.029545)	(0.296953)	(0.1776751)	(0.116354)	(0.053162)
6	1	10	0.5	0.65	0.078133	0.144240	0.121190	0.104512	0.002103	0.102458	0.132217	0.129830
					(0.033660)	(0.021042)	(0.027510)	(0.029150)	(0.076813)	(0.066139)	(0.094861)	(0.048752)
6	3	10	0.5	0.35	-0.420681	-0.284072	-0.219375	-0.286993	-0.599514	-0.237517	-0.055218	-0.109490
					(0.069761)	(0.014812)	(0.052750)	(0.066451)	(1.194876)	(0.185643)	(0.122375)	(0.087860)
6	3	10	0.5	0.5	-0.140562	-0.024380	-0.046860	-0.081131	-0.159015	-0.029861	0.035037	0.022001
					(0.051901)	(0.008391)	(0.012240)	(0.035961)	(0.147501)	(0.054682)	(0.115663)	(0.041981)
6	3	10	0.5	0.65	-0.081512	-0.070170	0.059871	0.005573	-0.107918	-0.054145	0.150060	0.151443
					(0.057660)	(0.025961)	(0.048022)	(0.055356)	(1.367465)	(0.4887338)	(0.097442)	(0.072951)
10	1	30	0.5	0.35	-0.176820	-0.150491	-0.176432	-0.176631	-0.180304	-0.176863	-0.152105	-0.152531
					(0.038103)	(0.022650)	(0.035771)	(0.037743)	(0.061614)	(0.053475)	(0.080038)	(0.045281)
10	1	30	0.5	0.5	-0.004421	-0.000314	-0.001423	-0.002501	-0.011424	-0.015016	-0.005248	-0.013917
					(0.004932)	(0.000035)	(0.000491)	(0.001562)	(0.005166)	(0.003730)	(0.010981)	(0.002115)
10	1	30	0.5	0.65	0.133140	0.149471	0.142904	0.139063	0.100651	0.072760	0.114331	0.118852
					(0.031812)	(0.022341)	(0.024143)	(0.029052)	(0.073791)	(0.060025)	(0.042123)	(0.032391)
10	3	30	0.5	0.35	-0.197230	-0.150581	-0.158602	-0.170151	-0.226247	-0.182218	-0.126589	-0.134173
					(0.039862)	(0.022681)	(0.025523)	(0.030948)	(0.064112)	(0.040656)	(0.042432)	(0.039351)
10	3	30	0.5	0.5	-0.041032	-0.000561	-0.011490	-0.021701	-0.098943	-0.055217	-0.006209	0.018501
					(0.008602)	(0.000024)	(0.000670)	(0.001401)	(0.019062)	(0.018583)	(0.015586)	(0.013763)
10	3	30	0.5	0.65	0.117001	0.149343	0.136457	0.128856	0.120389	0.044675	0.141123	0.138071
					(0.030297)	(0.022310)	(0.026761)	(0.029352)	(0.062961)	(0.053264)	(0.046101)	(0.031072)
6	1	10	0.8	0.65	-0.213196	-0.155521	-0.175723	-0.190304	-0.292068	-0.163757	-0.175802	-0.170832
					(0.043268)	(0.024413)	(0.033830)	(0.040461)	(0.097456)	(0.047901)	(0.048156)	(0.041873)
6	1	10	0.8	0.8	-0.03419	-0.00153	-0.01120	-0.01956	-0.01950	-0.00474	-0.03048	-0.01057
					(0.00643)	(0.00002)	(0.00070)	(0.00211)	(0.08164)	(0.05360)	(0.09750)	(0.03180)
6	1	10	0.8	0.95	0.139610	0.149765	0.149901	0.149024	0.145946	0.149093	0.141927	0.148558
					(0.040413)	(0.022439)	(0.024176)	(0.028212)	(0.062982)	(0.053253)	(0.050941)	(0.042110)
6	3	10	0.8	0.65	-0.269396	-0.168578	-0.197835	-0.225553	-0.279942	-0.162781	-0.132490	-0.149271
					(0.063568)	(0.036346)	(0.049347)	(0.056213)	(0.099364)	(0.078275)	(0.074286)	(0.062728)
6	3	10	0.8	0.8	-0.068961	-0.011720	-0.023674	-0.040396	-0.042523	-0.004734	-0.021826	0.002607
					(0.008002)	(0.001369)	(0.005660)	(0.006485)	(0.075406)	(0.054283)	(0.049514)	(0.035215)
6	3	10	0.8	0.95	0.133961	0.149732	0.149901	0.148754	0.148043	0.148390	0.142215	0.149527
					(0.038823)	(0.020421)	(0.022475)	(0.025138)	(0.054001)	(0.042183)	(0.050964)	(0.032756)
10	1	30	0.8	0.65	-0.253001	-0.160860	-0.190491	-0.214563	-0.247805	-0.168610	-0.162292	-0.148908
					(0.039041)	(0.017032)	(0.021596)	(0.029657)	(0.114891)	(0.077452)	(0.092753)	(0.062324)
10	1	30	0.8	0.8	-0.007212	-0.002083	-0.002654	-0.045593	-0.068918	-0.020103	0.001326	-0.004025
					(0.003489)	(0.002330)	(0.002651)	(0.002843)	(0.065554)	(0.040085)	(0.071306)	(0.035337)
10	1	30	0.8	0.95	0.137731	0.149708	0.149920	0.149052	0.148843	0.149251	0.145057	0.148513
					(0.030048)	(0.012415)	(0.022483)	(0.023224)	(0.044375)	(0.032280)	(0.041663)	(0.030134)
10	3	30	0.8	0.65	-0.178761	-0.150352	-0.161768	-0.168390	-0.175161	-0.161243	-0.157283	-0.163028
					(0.035623)	(0.012614)	(0.026780)	(0.029852)	(0.054253)	(0.046805)	(0.040367)	(0.036825)
10	3	30	0.8	0.8	-0.002971	-0.001433	-0.007124	-0.014605	-0.033492	-0.106683	-0.005136	0.002330
					(0.002323)	(0.000185)	(0.001252)	(0.001840)	(0.047871)	(0.014762)	(0.032417)	(0.006990)
10	3	30	0.8	0.95	0.147790	0.146981	0.152953	0.153240	0.146081	0.151250	0.147042	0.149238
					(0.031942)	(0.012501)	(0.022503)	(0.023674)	(0.048075)	(0.033126)	(0.041824)	(0.028293)

References

[1] Abramowitz, M. and Stegun, A. I. (1970), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York.

[2] Afzal Beg, M. and Singh, N. (1979), Estimation of $P_r(Y < X)$ for the Pareto Distribution. *IEEE Transactions on reliability*, **28**(5), 411-414.

[3] Dixit, U. J. and Jabbari Nooghabi, M. (2011), Efficient estimation in the Pareto distribution with the presence of outliers, *Statistical Methodology*, **8**(4), 340-355.

[4] Dixit, U. J. and Jabbari Nooghabi, M. (2011), Efficient Estimation of the parameters of the Pareto Distribution in the Presence of Outliers, *Communications of the Korean Statistical Society*, **18**(6), 817-835.

[5] Gunasekera, S. (2015). Generalized inferences of $R = Pr(X > Y)$ for Pareto distribution. *Stat Papers*, **56**(2), 333-351.

[6] Jabbari Nooghabi, M. (2016), Shrinkage estimation of $P(Y < X)$ in the exponential distribution mixing with exponential distribution. *Communications in Statistics-Theory and Methods*, **45**(5), 1477-1486.

[7] Jabbari Nooghabi, M. and Khaleghpanah Nooghabi, E. (2016), On Entropy of a Pareto distribution in the presence of outliers. *Communications in Statistics-Theory and Methods*, **45**(17), 5234-5250.

[8] Mahmoud, Mohamed A. W., El-Sagheer, Rashad M., Soliman, Ahmed A. and Abd Ellah, Ahmed H. (2016), Bayesian estimation of $P[Y < X]$ Based on Record Values from the Lomax Distribution and MCMC Technique. *Journal of Modern Applied Statistical Methods*, **15**(1), 488-510.

[9] Masoom Ali, M. and Woo, J. (2010), Estimation of tail-probsbility and reliability in exponentiated Pareto case. *Pak. J. Statist.*, **26**(1), 39-47.

[10] Odat, N. (2010), Estimation of Reliability Based on Pareto Distribution. *Applied Mathematical Sciences*, **4**(55), 2743-2748.

[11] Rezaei, S., Tahmasbi, R. and Mahmoodi, M. (2010). Estimation of $P[Y < X]$ for generalized Pareto distribution. *Journal of Statistical Planning and Inference*, **140**, 480-494.

[12] Wong, A. (2012), Interval estimation of $P(Y < X)$ for generalized Pareto distribution. *Journal of Statistical Planning and Inference*, **142**, 601607.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Coherent systems with heterogeneous components

Kelkinnama, M. [1]

Department of Mathematical Sciences, Isfahan University of Technology

Abstract

In this paper, we study the aging properties and stochastic ordering of coherent systems with heterogeneous components which components lifetime distributions follow the proportional hazard rate models. The results are based on the new concept of the survival signature.

Keywords: Coherent system, Survival signature, Heterogeneous components, Stochastic orders, Aging properties.

1 Introduction

The coherent systems are basic concepts in reliability theory and survival analysis. For a coherent system consisting of n independent and identically distributed (i.i.d.) components lifetime, the signature $\mathbf{s}$ is defined as the n-dimensional probability vector whose ith element is $s_i = Pr(T = X_{i:n})$, where T is the system lifetime and $X_{1:n}, ..., X_{n:n}$ are the ordered components lifetimes. The signature vector is a distribution-free function that depends on the systems design. Using signature, the reliability function of system can be represented as

$$\bar{F}_T(t) = \sum_{i=1}^{n} s_i \bar{F}_{i:n}(t) = \sum_{i=1}^{n} s_i \sum_{j=0}^{i-1} \binom{n}{j} F^j(t) \bar{F}^{n-j}(t). \tag{1.1}$$

Recently, Coolen and CoolenMaturi [1] defined a new metric for the coherent systems and called it *survival signature*. Consider an n-component system with components of r different types.

[1] m.kelkinnama@cc.iut.ac.ir

Suppose that the system has m_k components of type $k, k = 1, ..., r$. Assume that the lifetimes of components of the same type are exchangeable and that the lifetimes of components of different types are independent. Then the survival signature of the system is defined as a nonnegative function ϕ, where $\phi(i_1, ..., i_r)$ for $i_k = 0, ..., m_k, k = 1, ..., r$, represents the probability that the system works when precisely i_k components of type k are working. The function ϕ does not depend on the component distributions and depends only on the system design. Coolen and Coolen-Maturi [1] also showed that, under the stronger assumption that, for $k = 1, ..., r$, components of type k have i.i.d. lifetimes with common distribution function F_k, then the following representation for system's reliability function can be obtained

$$\bar{F}_T(t) = \sum_{i_1=0}^{m_1} ... \sum_{i_r=0}^{m_r} \phi(i_1, ..., i_r) \prod_{k=1}^{r} \binom{m_k}{i_k} F_k^{m_k - i_k}(t) \bar{F}_k^{i_k}(t). \tag{1.2}$$

specifically, when all the components are i.i.d., that is $r = 1, m_1 = n$ the (1.2) reduces to (1.1).

Comparisons of systems with heterogeneous components is an interesting topic in reliability studies. Many papers in this topic considered the series and parallel systems, (see. for example, [3], [4], [5], and [9]).

The expression (1.2) illustrates that systems with heterogeneous components can be compared, and that conditions under which one system provides better performance than another can be identified. Samaniego and Navarro [7] using the notion of survival signature, compared coherent systems.

Also, for studing the stochastic behaviors of coherent systems with different types of components, Erylimaz [2] investigated a mixture form for reliability function of system lifetime in the case of two types of dependent components.

In this paper, we investigate the aging properties and stochastic comparisons of coherent systems consisting of independent components with different distributions. We assume that the distributions of components follow the proportional hazard rate models.

Independent random variables $X_1, X_2, ..., X_n$ are said to follow the proportional hazard rates (PHR) model if for $i = 1, 2, ..., n$, the reliability function of X_i can be expressed as, $\bar{F}_i(x) = [\bar{F}(x)]^{\lambda_i}$, for $\lambda_i > 0$, where $\bar{F}(x)$ is the baseline reliability function. If $r(t)$ denotes the hazard rate corresponding to $\bar{F}$, then the hazard rate of X_i is $\lambda_i r(t), i = 1, 2, ..., n$. 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 stochastic orders that will be used in the sequel.

Definition 1. *(Shaked and Shanthikumar [8]) If the ratios below are well defined, the random variable X is said to be smaller than random variable Y in the*

(i) likelihood ratio order (denoted by $X \leq_{lr} Y$) if $g(x)/f(x)$ is increasing in x;

(ii) hazard rate order (denoted by $X \leq_{hr} Y$) if $\bar{G}(x)/\bar{F}(x)$ is increasing in x;

(iii) reversed hazard rate order (denoted by $X \leq_{rh} Y$) if $G(x)/F(x)$ is increasing in x;

(iv) stochastic order (denoted by $X \leq_{st} Y$) if $\bar{F}(x) \leq \bar{G}(x)$ for all x.

It is well known that $X \leq_{lr} Y \Longrightarrow X \leq_{hr(rh)} Y \Longrightarrow X \leq_{st} Y$

Definition 2. *Let* $\mathbf{x} = (x_1, ..., x_n)$ *and* $\mathbf{y} = (y_1, ..., y_n)$ *be two real vectors, denote* $x_{(1)} \leq ... \leq x_{(n)}$ *the increasing arrangement of* $x_1, ..., x_n$. $\mathbf{x}$ *is said to be majorized by* $\mathbf{y}$ *(denoted as* $\mathbf{x} \preceq^m \mathbf{y}$*) if* $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ *and* $\sum_{i=1}^j x_{(i)} \geq \sum_{i=1}^j y_{(i)}$ *for all* $j = 1, ..., n-1$.

Definition 3. *A real valued function* ϕ *defined on a set* $A \subseteq R^n$ *is said to be Schur-convex (Schur-concave) on* A *if* $\mathbf{x} \succeq^m \mathbf{y}$ *implies* $\phi(\mathbf{x}) \geq (\leq)\phi(\mathbf{y})$ *for any* $\mathbf{x}, \mathbf{y} \in A$.

Lemma 1.1. *[6]A permutation-symmetric continuously differentiable function* $\phi(\mathbf{X})$ *is Schur-concave (Schur-convex) if and only if*

$$(X_i - X_j)\Big(\frac{\partial \phi(\mathbf{X})}{\partial X_i} - \frac{\partial \phi(\mathbf{X})}{\partial X_j}\Big) \leq (\geq)0$$

for all $i \neq j$.

2 Main results

2.1 Aging properties

Consider a coherent system with two types of components say type A and B, respectively. Assume that the lifetimes of m_1 *components of type A are i.i.d. with reliability function* $\bar{F}_A$ *and hazard rate function* r_A*, the lifetimes of* m_2 *components of type B are i.i.d. with reliability function* $\bar{F}_B$ *and hazard rate function* r_B*, and the components of two types have independent lifetimes. Let*

$$P(\phi, \bar{F}_A(t), \bar{F}_B(t)) = \sum_{i_1=1}^{m_1} \sum_{i_2=0}^{m_2} i_1(\phi(i_1, i_2) - \phi(i_1 - 1, i_2)) \binom{m_1}{i_1} \binom{m_2}{i_2} \times$$
$$F_A^{m_1 - i_1}(t) \bar{F}_A^{i_1}(t) F_B^{m_2 - i_2}(t) \bar{F}_B^{i_2}(t),$$
$$Q(\phi, \bar{F}_A(t), \bar{F}_B(t)) = \sum_{i_1=0}^{m_1} \sum_{i_2=1}^{m_2} i_2(\phi(i_1, i_2) - \phi(i_1, i_2 - 1)) \binom{m_1}{i_1} \binom{m_2}{i_2} \times$$
$$F_A^{m_1 - i_1}(t) \bar{F}_A^{i_1}(t) F_B^{m_2 - i_2}(t) \bar{F}_B^{i_2}(t),$$
$$R(\phi, \bar{F}_A(t), \bar{F}_B(t)) = \sum_{i_1=0}^{m_1} \sum_{i_2=0}^{m_2} \phi(i_1, i_2) \binom{m_1}{i_1} \binom{m_2}{i_2} F_A^{m_1 - i_1}(t) \bar{F}_A^{i_1}(t) F_B^{m_2 - i_2}(t) \bar{F}_B^{i_2}(t).$$

Then the density function of the system lifetime, T, can be written as

$$f_T(t) = r_A(t) P(\phi, \bar{F}_A(t), \bar{F}_B(t)) + r_B(t) Q(\phi, \bar{F}_A(t), \bar{F}_B(t)),$$

therefore, the hazard rate and reversed hazard rate of the system are

$$r_T(t) = r_A(t) \frac{P(\phi, \bar{F}_A(t), \bar{F}_B(t))}{R(\phi, \bar{F}_A(t), \bar{F}_B(t))} + r_B(t) \frac{Q(\phi, \bar{F}_A(t), \bar{F}_B(t))}{R(\phi, \bar{F}_A(t), \bar{F}_B(t))}$$

and

$$\tilde{r}_T(t) = r_A(t) \frac{P(\phi, \bar{F}_A(t), \bar{F}_B(t))}{1 - R(\phi, \bar{F}_A(t), \bar{F}_B(t))} + r_B(t) \frac{Q(\phi, \bar{F}_A(t), \bar{F}_B(t))}{1 - R(\phi, \bar{F}_A(t), \bar{F}_B(t))}$$

At the following, we suppose that the distribution of components following the proportional hazard rate model.

Theorem 2.1. *Consider a coherent system with two subsystems connected to each other in series (Figure 2). The parallel subsystems have i.i.d. components which distributed by $\bar{F}_A(t)$ and $\bar{F}_B(t)$, respectively. The survival signature of system is given in Table 1 in parentheses. Suppose that the distributions of components follow a proportional hazard rate model as $\bar{F}_B(t) = [\bar{F}_A(t)]^\lambda$ for some $\lambda > 0$.*

(i) If $r_A(t)$ is increasing function of t (i.e. $\bar{F}_A$ is IFR) then $r_T(t)$ is increasing in t, for all $\lambda > 0$.

(ii) If $r_A(t)$ is decreasing function of t (i.e. $\bar{F}_A$ is DFR) then $\tilde{r}_T(t)$ is decreasing in t, for all $\lambda > 0$.

Proof. (i) For given system, the two ratios $\frac{P(\phi,x,x^\lambda)}{R(\phi,x,x^\lambda)}$ and $\frac{Q(\phi,x,x^\lambda)}{R(\phi,x,x^\lambda)}$ are decreasing function of $x, 0 \leq x \leq 1$ for all $\lambda > 0$. Letting $x = \bar{F}_A(t)$, under the assumptions of the theorem, the result is obtained.

(ii) For given system, the two ratios $\frac{P(\phi,x,x^\lambda)}{1-R(\phi,x,x^\lambda)}$ and $\frac{Q(\phi,x,x^\lambda)}{1-R(\phi,x,x^\lambda)}$ are increasing function of $x, 0 \leq x \leq 1$ for all $\lambda, \lambda > 0$, hence similar to part (i) the result holds. □

Theorem 2.2. *Consider the coherent system as in Theorem 2.1. Suppose that the distributions of components follow a proportional hazard rate model as $\bar{F}_A(t) = [\bar{F}_0(t)]^{\lambda_1}$ and $\bar{F}_B(t) = [\bar{F}_0(t)]^{\lambda_2}$ for some reliability function $\bar{F}_0(t)$ and some $\lambda_1, \lambda_2 > 0$.*

(i) If $r_0(t)$ is increasing function of t (i.e. $\bar{F}_0$ is IFR) then $r_T(t)$ is increasing in t, for all $\lambda_1, \lambda_2 > 0$.

(ii) If $r_0(t)$ is decreasing function of t (i.e. $\bar{F}_0$ is DFR) then $\tilde{r}_T(t)$ is decreasing in t, for all $\lambda_1, \lambda_2 > 0$.

Proof. The proof is similar to Theorem 2.1. □

2.2 Comparison study

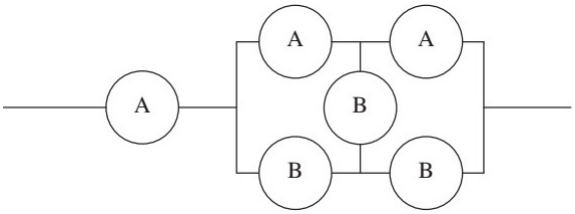


Figure 1: System 1

Samaniego and Navarro [7] compared coherent systems with different types of components using the survival signature. Their result is given below.

Theorem 2.3. *[7] Consider the two systems shown in Figures 1 and 2 with survival signatures given in Table 1. For $i = 1, 2$, let T_i be the lifetime of system i. Assume that the lifetimes of components of type A are i.i.d. with reliability $\bar{F}_A$, the lifetimes of components of type B are i.i.d. with reliability $\bar{F}_B$ and components of different types have independent lifetimes. If $\bar{F}_A(t) \leq \bar{F}_B(t)$ for all t, then $T_1 \leq_{st} T_2$.*

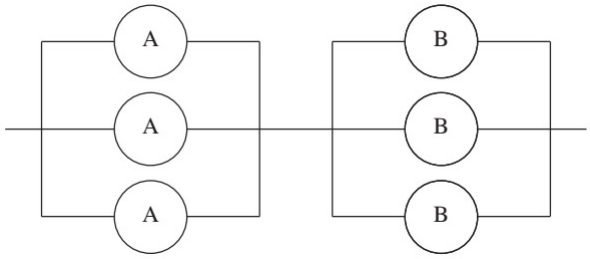


Figure 2: System 2

Table 1: The survival signature of the system in Figure 1 (Figure 2)

$\phi_1(i_1,i_2)(\phi_2(i_1,i_2))$	$i_2=0$	$i_2=1$	$i_2=2$	$i_2=3$
$i_1=0$	0(0)	0(0)	0(0)	0(0)
$i_1=1$	0(0)	0(1)	$\frac{1}{9}$(1)	$\frac{1}{3}$(1)
$i_1=2$	0(0)	0(1)	$\frac{4}{9}$(1)	$\frac{2}{3}$(1)
$i_1=3$	1(0)	1(1)	1(1)	1(1)

At the continue, we stochastically compare the coherent systems.

Theorem 2.4. *Consider two coherent systems with lifetimes $T_i, i=1,2$ composed of two types of components. Assume that the lifetimes of components of type A are i.i.d. with reliability function $\bar{F}_0$, and the lifetimes of components of type B are i.i.d. with reliability function $\bar{F}_0^\lambda$, $\lambda>0$ and components of different types are independent.*

(i) $T_1 \leq_{st} T_2$ for all $\bar{F}_0$ if $R(\phi_1,x,x^\lambda) \leq R(\phi_2,x,x^\lambda)$ for all $x, 0\leq x\leq 1$.

(ii) $T_1 \leq_{hr} T_2$ for all $\bar{F}_0$ if $\frac{R(\phi_2,x,x^\lambda)}{R(\phi_1,x,x^\lambda)}$ is decreasing in $x, 0\leq x\leq 1$.

(iii) $T_1 \leq_{lr} T_2$ for all $\bar{F}_0$ if $\frac{P(\phi_2,x,x^\lambda)+\lambda Q(\phi_2,x,x^\lambda)}{P(\phi_1,x,x^\lambda)+\lambda Q(\phi_1,x,x^\lambda)}$ is decreasing in $x, 0\leq x\leq 1$.

Proof. By using Definition 1 and letting $\bar{F}_0(t)=x$ we get the results. □

Now, we investigate the conditions at Theorem 2.4 for the two systems considered at Theorem 2.3 which their components obey the proportional hazard model. The following theorems are obtained by some tedious computations and therefore the proofs are omitted.

Theorem 2.5. *Consider the two coherent systems given at Theorem 2.3. Then*

(i) $T_1 \leq_{hr} T_2$ for all $\bar{F}_0$ and $0<\lambda\leq 2$.

(ii) $T_1 \leq_{hr} T_2$ for all $\bar{F}_0$ and $0<\lambda\leq 2$.

(iii) $T_1 \leq_{lr} T_2$ for all $\bar{F}_0$ and $0<\lambda\leq 2$.

Note that for $0<\lambda\leq 1$ we have $\bar{F}_0 \leq \bar{F}_0^\lambda$ and based on Theorem 2.3, $T_1 \leq_{st} T_2$. Hence, Theorem 2.5 shows that this result holds, for $1<\lambda\leq 2$.

Theorem 2.6. *Consider the two coherent systems as in Theorem 2.5.*

(i) Assume that at System 1, $\bar{F}_A = \bar{F}_0$, and $\bar{F}_B = [\bar{F}_0]^\lambda$ and at System 2, $\bar{F}_A = \bar{F}_*$, and $\bar{F}_B = [\bar{F}_*]^\lambda$ for some reliability functions $\bar{F}_0, \bar{F}_*$ and $\lambda > 0$. If $\bar{F}_0(t) \leq \bar{F}_*(t)$ for all $t > 0$ then $T_1 \leq_{st} T_2$ for all $\lambda > 0$.

(ii) Assume that at System 1, $\bar{F}_A = [\bar{F}_0]^{\lambda_1}$, and $\bar{F}_B = [\bar{F}_0]^{\lambda_2}$ and at System 2, $\bar{F}_A = [\bar{F}_*]^{\lambda_1}$, and $\bar{F}_B = [\bar{F}_*]^{\lambda_2}$, for some reliability functions $\bar{F}_0, \bar{F}_*$ and $\lambda_1, \lambda_2 > 0$.

If $\bar{F}_0(t) \leq \bar{F}_*(t)$ for all $t > 0$ then $T_1 \leq_{st} T_2$ for every λ_1, λ_2 such that $\lambda_1 \geq \lambda_2$.

At the following theorems, we consider coherent system in Figure 1 with survival signature given in Table 1, and compare the lifetime of this system under different conditions.

Theorem 2.7. *Consider the coherent system illustrated in Figure 1. Let T_1 be the lifetime of the system when $\bar{F}_A = \bar{F}_0$ and $\bar{F}_B = \bar{F}_0^{\lambda_1}$, and T_2 be the lifetime of the system when $\bar{F}_A = \bar{F}_0$ and $\bar{F}_B = \bar{F}_0^{\lambda_2}$. Then*

(i) $T_1 \geq_{st} T_2$ if $0 < \lambda_1 < \lambda_2$.

(ii) $T_1 \geq_{hr} T_2$ if $0 < \lambda_1 < \lambda_2 \leq 1$.

Proof. (i) For given system we have,

$R(\phi, x, x^\lambda) = 2x^{(3+3\lambda)} - 3x^{(3+2\lambda)} + 2x^{(2+2\lambda)} - 2x^{(2+3\lambda)} + x^3 + x^{(1+2\lambda)}$

It can be shown that it is a decreasing function of λ, for all $x, 0 \leq x \leq 1$. Therefore, for all $t > 0$,

$$\bar{F}_{T_1}(t) = R(\phi, \bar{F}_0(t), [\bar{F}_0(t)]^{\lambda_1}) \geq R(\phi, \bar{F}_0(t), [\bar{F}_0(t)]^{\lambda_2}) = \bar{F}_{T_2}(t).$$

□

Theorem 2.8. *Consider the coherent system illustrated in Figure 1. Let T_1 be the lifetime of the system when $\bar{F}_A = \bar{F}_0$ and $\bar{F}_B = \bar{F}_0^\lambda$, and T_2 be the lifetime of the system when $\bar{F}_A = \bar{F}_1$ and $\bar{F}_B = \bar{F}_1^\lambda$. Then if $\bar{F}_0(t) \leq \bar{F}_1(t)$ for all $t > 0$, then $T_1 \leq_{st} T_2$ for all $\lambda > 0$.*

Proof. For given system, similar to proof of Part (i) of Theorem 2.8, we have $R(\phi, x, x^\lambda)$ is an increasing function of $x, 0 \leq x \leq 1$, for all $\lambda > 0$. Therefore, for all $t > 0$,

$$\bar{F}_{T_1}(t) = R(\phi, \bar{F}_0(t), [\bar{F}_0(t)]^\lambda) \leq R(\phi, \bar{F}_1(t), [\bar{F}_1(t)]^\lambda) = \bar{F}_{T_2}(t).$$

□

Theorem 2.9. *Consider the coherent system illustrated in Figure 1. Let T_1 be the lifetime of the system when $\bar{F}_A = \bar{F}_0^{\lambda_1}$ and $\bar{F}_B = \bar{F}_0^{\lambda_2}$, and T_2 be the lifetime of the system when $\bar{F}_A = \bar{F}_0^{\lambda_1^*}$ and $\bar{F}_B = \bar{F}_0^{\lambda_2^*}$. Suppose that $0 < \lambda_1 \leq \lambda_2$ and $0 < \lambda_1^* \leq \lambda_2^*$. If $(\lambda_1, \lambda_2) \succeq^m (\lambda_1^*, \lambda_2^*)$ then $T_1 \leq_{st} T_2$.*

Proof. For the given system, we have

$$\frac{\partial R(\phi, x^{\lambda_1}, x^{\lambda_2})}{\partial \lambda_2} - \frac{\partial R(\phi, x^{\lambda_1}, x^{\lambda_2})}{\partial \lambda_1} =$$
$$- ln(x)(3x^{3\lambda_1} - x^{\lambda_1+2\lambda_2} - 3x^{3\lambda_1+2\lambda_2} + 2x^{2\lambda_1+3\lambda_2})$$

It can be shown that this function is nonnegative for $\lambda_1 \leq \lambda_2$.Hence, $(\lambda_2 - \lambda_1)\big(\frac{\partial \bar{F}_T(t)}{\partial \lambda_2} - \frac{\partial \bar{F}_T(t)}{\partial \lambda_1}\big)$ is nonnegative for $\lambda_1 \leq \lambda_2$ and hence using Lemma 1.1 we get that the reliability function of system is Schur-convex in (λ_1, λ_2) and using Definitions 2 and 3, the desired result holds. □

At the following, we compare a coherent system with heterogeneous components and the coherent system with homogeneous components.

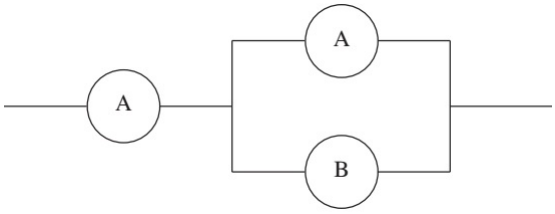


Figure 3: System 3

Table 2: The survival signature of the System 3

(i_1, i_2)	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$	$(2,0)$	$(2,1)$
$\phi(i_1, i_2)$	0	0	0	$\frac{1}{2}$	1	1

Theorem 2.10. *Consider a coherent system given in Figure 3 with two types of components with survival signature given in Table 2 and denote its lifetime by T_1. Also, consider a similar coherent system with the same structure which consists of i.i.d. components with common reliability function $\bar{\bar{F}}(t) = \frac{1}{2}(\bar{F}_A(t) + \bar{F}_B(t))$ and denote the system's lifetime by T_2. If $\bar{F}_A(t) \leq \bar{F}_B(t)$ for all $t > 0$ then $T_1 \leq_{st} T_2$.*

Proof. For the system with heterogeneous components, we have

$$\bar{F}_{T_1}(t) = \bar{F}_A(t)(1 - \bar{F}_A(t))\bar{F}_B(t) + \bar{F}_A^2(t).$$

The signature vector of this system is $\mathbf{s} = (\frac{1}{3}, \frac{2}{3}, 0)$, hence, the reliability function of the system with i.i.d. components is

$$\begin{aligned}\bar{F}_{T_2}(t) &= \bar{\bar{F}}^2(t)(2 - \bar{\bar{F}}(t)) \\ &= \frac{1}{2}(\bar{F}_A(t) + \bar{F}_B(t))^2\big(1 - \frac{1}{4}(\bar{F}_A(t) + \bar{F}_B(t))\big)\end{aligned}$$

Then one can show that $\bar{F}_{T_1}(t) \leq \bar{F}_{T_2}(t)$ holds if $\bar{F}_A(t) \leq \bar{F}_B(t)$ for all $t > 0$. □

Theorem 2.11. *Consider a coherent system given in Figure 3 with lifetime T_1 and survival signature given in Table 2. Suppose that $\bar{F}_A = \bar{F}_0^{\lambda_1}$ and $\bar{F}_B = \bar{F}_0^{\lambda_2}$. Also, consider a coherent system with the same structure but with i.i.d. components with common reliability function $[\bar{F}_0(t)]^{\bar{\lambda}}$ where, $\bar{\lambda} = \frac{1}{2}(\lambda_1 + \lambda_2)$ and denote the lifetime of such system by T_2. If $\lambda_1 \leq \lambda_2$ then $T_1 \geq_{st} T_2$.*

Proof. This can be proved similar to Theorem 2.10 and therefore is omitted. □

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, 115130.*

[2] *Erylimaz, S. (2015). Systems composed of two types of nonidentical and dependent components, Naval Research Logistics,* **62***(5), 388394.*

[3] *Fang, L., Balakrishnan, N. (2016). Ordering results for the smallest and largest order statistics from independent heterogeneous exponential-Weibull random variables, Statistics,* **50***, 1195-1205.*

[4] *Kochar, S.C., Torrado, N. (2015). On stochastic comparisons of largest order statistics in the scale model, Communications in Statistics -Theory and Methods,* **44***, 4132-4143.*

[5] *Kochar, S.C. (2012). Stochastic Comparisons of Order Statistics and Spacings: A Review. ISRN Probability and Statistics,* **2012***, Article ID 839473, 47 pages, 2012. doi:10.5402/2012/839473.*

[6] *Marshall, A. W., Olkin, I. and Arnold, B. C. (2011),* Inequalities: Theory of Majorization and Its Applications*, Springer, New York.*

[7] *Samaniego F. J. ,Navarro J. (2016) , On comparing coherent systems with heterogeneous components, Advances in Applied Probability,* **48***, 124.*

[8] *Shaked M, Shanthikumar JG. (2007).* Stochastic orders*, Springer, New York.*

[9] *Torrado, N. (2015). Comparisons of smallest order statistics from Weibull distributions with different scale and shape parameters, Journal of the Korean Statistical Society,* **44***, 6876.*

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Comparison of coherent systems using reverse mean residual order

Khaleghpanah Nooghabi, E. [1] Rezaei, M. [2] and Chahkandi, M. [3]

[1,2,3] Department of Statistics, University of Birjand

Abstract

The performance ordering of the coherent systems is well known topic in the reliability. Many works and researches in different criteria have been done. Now, we are interested in ordering the coherent system with dependent identical distributed components in terms of reversed mean residual lifetime. The results are based on the representation of the system distribution as a distorted distribution function of the common components' distribution.

Keywords: Distortion, Reverse mean residual lifetime, Stochastic order, Copula.

1 Introduction

The study on performance ordering of coherent system is a principle subject in the reliability which have been done based on different criteria. Most of the results are on the case that, the system components are i.i.d. see for example [2, 5, 6]. Navarro *et al.* [3] obtained some ordering results for coherent systems with identical distributed (ID) components by using stochastic, hazard rate and likelihood ratio orders. They also achieved a representation for reliability function based on distortion function to compare the system lifetimes in [3]. An increasing continuous function $q : [0,1] \to [0,1]$ is called distortion function introduced in [8]. For a distribution F, define $F_q = q(F)$, then F_q is also a distribution which is called distortion distribution.

[1]e.khaleghpanah@birjand.ac.ir
[2]mjrezaei@birjand.ac.ir
[3]mchahkandi@birjand.ac.ir

Recently, Navarro and Gomis *et al.* [4] obtained comparisons in the mean residual life (MRL) order for coherent systems with ID component lifetimes. In this paper, we extend the results of [4] for the case of reverse mean residual order (RMRL). Let X and Y be two non-negative random variables with distribution functions F and G. Reversed hazard rate order has been defined in [6]. X is smaller than Y in the reversed hazard rate order which is written as $X \leqslant_{RHR} Y$ if $\frac{G(t)}{F(t)}$ is increasing in $t \in (-\infty, +\infty)$. Equivalently, $X \leqslant_{RHR} Y$ iff $(t - Y|Y \leqslant t) \leqslant_{ST} (t - X|X \leqslant t)$.
As an another order, X is said smaller than Y in the reversed mean residual life order, which is shortly denoted by $X \leqslant_{RMRL} Y$, if $E(t - Y|Y \leqslant t) \leqslant E(t - X|X \leqslant t)$ for all t. Obviously, the RMRL order concluds $E(X) \leqslant E(Y)$, and also the RHR order implies the RMRL order. The mentioned relations are summarized in Table 1.

Table 1: Relationship between some relevant stochastic orders

$X \leqslant_{RHR} Y$	$\Longrightarrow$	$X \leqslant_{RMRL} Y$
$\Downarrow$		$\Downarrow$
$X \leqslant_{ST} Y$	$\Longrightarrow$	$E(X) \leqslant E(Y)$

In Section 2, the fundamental theorem is presented to obtain comparison results in the RMRL ordering.

2 Main Results

First of all, Belzunce *et al.* [1] presented some conditions which is provided a reverse connection between means order and RMRL order, that is available in the following theorem.

Theorem 2.1. *Let $X \sim F$ and $Y \sim G$ be non-negative random variables with finite means such that $E(X) \leqslant E(Y)$, then $X \leqslant_{RMRL} Y$ if there exist a t_0 in which $\frac{F(t)}{G(t)}$ is decreasing in $t \leqslant t_0$ and is increasing in $t \geqslant t_0$.*

In the next lemma sufficient conditions for equivalency of $\frac{F_X}{G_Y}$ based on the ratio of $\frac{q_X}{q_Y}$ is given.

Lemma 2.2. *Assume that X and Y be two non-negative random variables with distorted distributions q_X and q_Y based on the same continuous distribution F. Then, the following condonations are equivalent.*

(a) $\exists t_0 \in (0, \infty)$ *such that* $\dfrac{F_X}{G_Y}$ *is decreasing in* $(0, t_0)$ *and is increasing in* (t_0, ∞).

(b) $\exists u_0 \in (0, 1)$ *such that* $\dfrac{q_X}{q_Y}$ *is decreasing in* $(0, u_0)$ *and is increasing in* $(u_0, 1)$.

Proof. Assume that (a) holds and take $u_0 = F(t_0) \in (0, 1)$. If F is continuous and $0 < u_1 \leqslant u_2 \leqslant u_0$, there exists $0 \leqslant t_1 \leqslant t_2 \leqslant t_0$ such that $u_i = F(t_i)$ for $i = 1, 2$. As $\frac{F_X}{G_Y}$ is decreasing in $(0, t_0)$, then

$$\frac{q_X(u_2)}{q_Y(u_2)} = \frac{F_X(t_2)}{G_Y(t_2)} \leqslant \frac{F_X(t_1)}{G_Y(t_1)} = \frac{q_X(u_1)}{q_Y(u_1)},$$

so $\frac{q_X}{q_Y}$ is decreasing in $(0, u_0)$. Similarly, if $u_0 \leqslant u_1 \leqslant u_2 < 1$ and F is continuous, there exists $t_0 \leqslant t_1 \leqslant t_2$ such that $u_i = F(t_i)$ for $i = 1, 2$, the result is obtained. Then (a) implies (b) and the reverse implication can be proved obviously. □

As a consequence, we obtain conditions for comparing system based on RMRL order in the next theorem.

Theorem 2.3. *Suppose that X and Y be two non-negative random variables with distorted distributions q_X and q_Y based on the same continuous distribution F. If there exist a $u_0 \in (0, 1)$, such that $\frac{q_X}{q_Y}$ is decreasing in $(0, u_0)$ and is increasing in $(u_0, 1)$, then $X \leqslant_{RMRL} Y$ whenever $E(X) \leqslant E(Y)$.*

Proof. It is immediately proved based on Theorem 2.1 and Lemma 2.2. □

Now, consider $H = \frac{q_X}{q_Y}$, Theorem 2.3 is equivalent to say that, if H is decreasing in $(0, u_0)$ and increasing in $(u_0, 1)$ i.e. (H is bathtube), then the correspond random variables X and Y have RMRL order when their means are ordered. In special case $t_0 = 0$, without the condition about the means, we have $X \leqslant_{RMRL} Y$. The next examples are given to clarify the subject.
The connection between distortion distribution and distribution function of coherent system is the next discussed topic. In order to, consider a coherent system with n identical distributed components $X_1, ..., X_n \sim F$, structure function ϕ and lifetime T. If $P_1, ..., P_r$ are minimal path sets of the syetem, then

$$F_T(t) = \sum_{j=1}^{r} P(X_{P_j} \leqslant t) - \sum_{i<j} P(X_{P_i} \cap X_{P_j} \leqslant t) + ... + (-1)^{r+1} P(X_{P_1} \cap ... \cap X_{P_r} \leqslant t).$$

The dependence between identical components can be expressed by copula C, hence, distribution function of each minimal path set P_j can be written as follow.

$$\begin{aligned} F_{P_j}(t) &= P(X_{P_j} \leqslant t) = P(\min_{i \in P_j} X_i \leqslant t) \\ &= C(F(t_1), ..., F(t_n)). \end{aligned} \tag{2.1}$$

So far, the distribution of the system lifetime based on equation (2.1) can be represented as,

$$F_T(t) = q(F(t)) \tag{2.2}$$

where q is distortion function. For more perception consider the following example. Distortion function for all coherent systems with $1 - 3$ independent identical distributed components are exhibited in Table 2.

Table 2: Distortion function of all systems with $1-3$ i.i.d. components.

i	$T_i = \phi(X_1, ..., X_n)$	$q_i(u)$
1	$X_{1:1} = X_1)$	u
2	$X_{1:2} = \min(X_1, X_2)$	u^2
3	$X_{2:2} = \max(X_1, X_2)$	$2u - u^2$
4	$X_{1:3} = \min(X_1, X_2, X_3)$	u^3
5	$\min(X_1, \max(X_2, X_3)$	$2u^2 - u^3$
6	$X_{2:3}$ (2-out of-3)	$3u^2 - 2u^3$
7	$\max(X_1, \min(X_2, X_3)$	$u + u^2 - u^3$
8	$X_{3:3} = \max(X_1, X_2, X_3)$	$3u - 3u^2 + u^3$

Example 2.4. Assume a system with three components whose the lifetime is, $T_7 = \max(X_1, \min(X_2, X_3))$ with minimal path sets $P_1 = \{1\}$ and $P_2 = \{2, 3\}$. Then

$$\begin{aligned} F_{T_7}(t) &= P(T_7 \leqslant t) = P(\{X_{P_1} \leqslant t\} \cup \{X_{P_2} \leqslant t\}) \\ &= P(X_1 \leqslant t) + P(\{X_2 \cap X_3\} \leqslant t) - P(\{X_1 \cap X_2 \cap X_3\} \leqslant t) \\ &= C(F(t), 1, 1) + C(1, F(t), F(t)) - C(F(t), F(t), F(t)) = q_7(F(t)). \end{aligned}$$

Thus, the distortion function of T_7 is,

$$q_7(u) = C(u, 1, 1) + C(1, u, u) - C(u, u, u),$$

in which, C is copula function. If the components are independent i.e. C is the product copula, then

$$q_7(u) = u + u^2 - u^3.$$

A straight computation, give us

$$H_{17}(u) = \frac{q_1(u)}{q_7(u)} = \frac{u}{u + u^2 - u^3}.$$

This ratio, H_{17} is decreasing in $(0, 0.5)$ and increasing in $(0.5, 1)$, therefore $T_1 \leqslant_{RMRL} T_7$ when $E(T_1) \leqslant E(T_7)$.
Note that, $T_1 \leqslant_{RMRL} T_7$ is satisfied, if $E(T_1) \leqslant E(T_7)$, for example if the component lifetime is Weibull distribution $F(t) = e^{-t^2} \quad t \geqslant 0$, then

$$\begin{aligned} F_{T_7}(t) &= e^{-t^2} + e^{-2t^2} - e^{-3t^2} \\ F_{T_1}(t) &= e^{-t^2} \end{aligned}$$

so the means are,

$$E(T_1) \approx 0.88622 \leqslant 1.195443 \approx E(T_7).$$

For another comparison, computation give us

$$H_{21}(u) = \frac{q_2(u)}{q_1(u)} = \frac{u^2}{u},$$

this function is increasing in $(0, 1)$, so $T_1 \leqslant_{RHR} T_2$ and $T_1 \leqslant_{RMRL} T_2$ is concluded without any condition about their expectations.

Theorem 2.5. *Assume that $T_1, ..., T_8$ be the lifetimes of coherent systems with $1-3$ i.i.d. components which is given in Table 2.1, then*

$$T_3 \leqslant_{RMRL} T_8 \leqslant_{RMRL} T_1 \leqslant_{RMRL} T_7 \leqslant_{RMRL} T_6 \leqslant_{RMRL} T_5 \leqslant_{RMRL} T_2 \leqslant_{RMRL} T_4,$$

whenever their means are ordered.

Proof. The proof is obtained by studying all the ratio of distortion function pairs similar to Example 2.4. □

If the ratio, based on the Theorem 2.3, is bathtub in $(0, 1)$, they have RMRL ordering (when their means are order). If the ratio is increasing in $(0, 1)$, the RHR and RMRL order (without any condition) holds. So, with more precise in the conclusion, we can rewritten the result of Theorem 2.5 in such way,

$$T_3 \leqslant_{RHR} T_8 \leqslant_{RHR} T_1 \leqslant_{RMRL} T_7 \leqslant_{RHR} T_6 \leqslant_{RHR} T_5 \leqslant_{RHR} T_2 \leqslant_{RHR} T_4.$$

References

[1] Belzunce, F., Martinez-Riquelme, C. and Ruiz, J. M. (2013), On sufficient conditions for mean residual life and related orders, *Computational Statistics and Data Analysis*, **61**, 199-210.

[2] Kochar, S., Mukerjee, H. and Samaniego, F. J. (1999), The signature of a coherent system and its application to comparison among systems, *Naval Research Logistics*, **46**, 507-523.

[3] Navarro, J., Aguila, Y., Sordo, M. A. and Surez-Llorens, A. (2013), Stochastic ordering properties for systems with dependent identically distributed components, *Applied Stochastic Models in Business and Industry*, **29**, 264-278.

[4] Navarro, J. and Gomis, G. M. (2015), Comparisons in the mean residual life order of coherent systems with identically distributed components, *Applied Stochastic Models in Business and Industry*, **32**, 33-47.

[5] Samaniego, F. J. (1985), On closure of the IFR class under formation of coherent systems, *IEEE Transactions on Reliability*, **34**, 69-72.

[6] Samaniego, F. J. (2007), *System signatures and their applications in engineering reliability*, International Series in Operations Research and Management Science, Springer, New York.

[7] Shaked, M. and Shanthikumar, J. G. (2007), *Stochastic Orders*, Springer-Verlag, New York.

[8] Yaari, M. E. (1987), The dual theory of choice under risk, *Econometrica*, **55**, 95-115.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Some monotonicity properties of mean residual life of a k-out-of-n system with nonidentical components

Khanjari Sadegh, M. [1] and Mohammadzadeh, T. [2]

[1,2] Department of Statistics, Faculty of Mathematical Sciences and Statistics, University of Birjand

Abstract

This article considers some monotonicity properties of the residual life and the mean residual life (MRL) of a k-out-of-n System in two cases, when all components of the system are working, also when the failed components of the system are known. We assume that the lifetimes of the system components are independent random variables but not necessarily identically distributed (inid), extending some results in literatures.

Keywords: Mean residual life, k-out-of-n System, Inid components, IFR, DFR.

1 Introduction

The study of reliability properties of k-out-of-n systems has been considered by many researchers. An important class of coherent systems is k-out-of-n systems. A system with n components has k-out-of-n structure if it fails as long as at least k of its components fail.(For a details on coherent systems see Barlow and Proschan, 1975). Let nonnegative and continuous random variables $X_1, \ldots, X_n$ denote the lifetimes of the system components. It is known that the lifetime of the k-out-of-n system is the kth order statistic, that is $X_{k:n}$. The MRL and the failure rate functions are very important in reliability and survival analysis as both of them uniquely characterize the distribution

[1] mkhanjari@birjand.ac.ir
[2]

function. For a random lifetime $T(>0)$ with a reliability function $\bar{F}(t) = P(T > t) = 1 - F(t)$ and a density function $f(t)$, the failure rate and the MRL functions are defined as

$$h(t) = \frac{f(t)}{\bar{F}(t)} = -d/dt(ln\bar{F}(t))$$

and

$$m(t) = E(T - t|T > t) = \frac{\int_0^\infty \bar{F}(t + x)dx}{\bar{F}(t)},$$

respectively.
It is well known that if distribution F has increasing(decreasing) failure rate IFR(DFR) property, that is $h(t)$ is increasing(decreasing) function, then it has decreasing(increasing) MRL property DMRL(IMRL), that is $m(t)$ is decreasing(increasing) function.

$$\text{IFR(DFR)} \Rightarrow \text{DMRL(IMRL)}$$

The reverse above implications are not valid(see Bryson, 1969 for counterexamples). Also they do not hold true for a system in the following sense. Let $T = \phi(X_1, \ldots, X_n)$ represent the lifetime of the system and suppose X_i's have IFR(DFR) property. Then in general T does not have DMRL(IMRL) property. Actually T may not have IFR(DFR) property. See for example, Samaniego (1985) when X_i's are independent and identically distributed (iid) random variables. Note that for the system lifetime T we may define different types of MRL. Therefore even if T be IFR(DFR) closed (that is T has IFR(DFR) property when X_i's have IFR(DFR) property) it may not have DMRL(IMRL) property. In iid case Samaniego (1985) obtained a necessary and sufficient condition under which to T be IFR(DFR) closed and showed that k-out-of-n system is IFR closed. Most of literatures on k-out-of-n systems are based on the iid assumption. See for example Asadi and Goliforushani (2008) and Goliforushni et al.(2011). But in inid case not much works have been done. It seems Sadegh (2008-a) was the first to obtain some properties of MRL of a parallel system with inid components. Kochar and Zuo (2010) further investigated

$$X_{k:n} - t|X_{r:n} > t$$

and

$$X_{k:n} - t|X_{r:n} < t < X_{k:n}$$

for $1 \leq r < k \leq n$.
Ding et al.(2012) obtained some stochastic comparisons between residual lifetimes of parallel systems and inactivity times of series systems with inid components.
In inid case we note that a series system is both IFR and DFR closed but this is not case for a k-out-of-n system.

We consider here two MRL functions for a k-out-of-n system as follows:

$$H_n^k(t) = E(X_{k:n} - t|X_{1:n} > t)$$

and

$$H^r_{n,k}(t) = E(X_{k:n} - t|N(t) = r, C_r = \{i_1, \ldots, i_r\})$$

for $1 \leq r < k \leq n$ in which $N(t)$ is the number of failed components of the system up to time t and C_r is a subset of $\{1, \ldots, n\}$ represents the set of indices of failed components of the system up to this time. In other words $H^r_{n,k}(t)$ measures the MRL of the system at time t when the failed components of the system up to this time are known.

In iid case Asadi and Goliforushani (2008) showed that if F is IFR(DFR) then $H^k_n(t)$ is decreasing(increasing) in t. They also showed that if F is IFR then $M^{r,k}_n(t) = E(X_{k:n} - t|X_{r:n} > t$, $1 \leq r \leq k \leq n$ is decreasing in t. Sadegh (2008-b) showed in iid case that

$$P(X_{k:n} - t > x|N(t) = r) = P(X_{k-r:n-r} - t > x|X_{1:n-r} > t) \tag{1.1}$$

and therefore

$$E(X_{k:n} - t|N(t) = r) = E(X_{k-r:n-r} - t|X_{1:n-t} > t) = H^{k-r}_{n-r}(t)$$

which is increasing(decreasing) in t if F is DFR(IFR). We note that in iid case

$$H^r_{n,k}(t) = H^{k-r}_{n-r}(t).$$

In the following section we obtain some monotonicity properties of $H^k_n(t)$ and $H^r_{n,k}(t)$ in inid case.

2 Main results

Let $X_1, \ldots, X_n$ denote the lifetimes of n components in a k-out-of-n system and suppose $F_i(x) = 1 - \bar{F}_i(x)$ is the distribution function of X_i.

Lemma 2.1. *In inid case we have* $(X_{k:n} - t|X_{1:n} > t) \leq_{st} (X_{k:n-1} - t|X_{1:n-1} > t)$.

Proof. See Sadegh(2011).
We recall that $X \leq_{st} Y$ if $\bar{F}(x) \leq \bar{G}(x)$, for all x. Above lemma implies that $H^k_n(t) = E(X_{k:n} - t|X_{1:n} > t)$ is a decreasing function of n. We now consider the behavior of $H^k_n(t)$ in terms of t. Note that $X_{k:n} - t|X_{1:n} > t$ is distributed as kth order statistic from the sample $X^i_t = (X_i - t|X_i > t)$, $i = 1, \ldots, n$. We have $P(X^i_t \leq x) = 1 - \frac{\bar{F}_i(t+x)}{\bar{F}_i(t)}$ and therefore

$$P(X_{k:n} - t > x|X_{1:n} > t) = P(\sum_{i=1}^{n} Z^i_{t,x} \leq k - 1) \tag{2.1}$$

where $Z^i_{t,x}$ is distributed as Binomial$(1, p_{i,x}(t) = 1 - \frac{\bar{F}_i(t+x)}{\bar{F}_i(t)})$ and $Z^i_{t,x}$, $i = 1, \ldots, n$ are independent random variables. It is easy to see that

$$P(\sum_{1}^{n} Z^i_{t,x} \leq k - 1) - P(\sum_{1}^{n-1} Z^i_{t,x} \leq k - 2) = \frac{\bar{F}_n(t+x)}{\bar{F}_n(t)} P(\sum_{1}^{n-1} Z^i_{t,x} = k - 1) \geq 0$$

that is

$$(X_{k:n} - t | X_{1:n} > t) \geq_{st} (X_{k-1:n-1} - t | X_{1:n-1} > t)$$

and therefore $H_n^k(t) \geq H_{n-1}^{k-1}(t)$.
The following lemma is required in the sequel.

Lemma 2.2. *$K_n^r(p) = \sum_{x=0}^{r} \binom{n}{x} p^x (1-p)^{n-x}$, $0 \leq r \leq n$ is a decreasing function of $0 < p < 1$.*

Proof. $d/dp(K_n^r(p)) = \sum_1^r x\binom{n}{x} p^{x-1}(1-p)^{n-x} - \sum_0^r (n-x)\binom{n}{x} p^x (1-p)^{n-x-1} = n\sum_0^{r-1} \binom{n-1}{x} p^x (1-p)^{n-x-1} - n\sum_0^r \binom{n-1}{x} p^x (1-p)^{n-x-1}$ which is obviously negative and the lemma follows.

In above lemma we note that if p is replaced by p_t in which p_t is increasing(decreasing) in t then $K_n^r(p_t)$ is also decreasing(increasing) in t.

Lemma 2.3. *If F_i's are IFR(DFR) distributions then $P(X_{k:n} - t > x | X_{1:n} > t)$ is decreasing(increasing) in t.*

Proof. From Equation (2.1) we have

$$P(X_{k:n} - t > x | X_{1:n} > t) = \sum_{j=0}^{k-1} \sum_{C_j} \prod_{i \in C_j} p_{i,x}(t) \prod_{i \in C - C_j} (1 - p_{i,x}(t)) \tag{2.2}$$

where $C = \{1, \ldots, n\}$ and C_j is a subset of C with cardinality j. It is easy to see that F_i's are IFR(DFR) if and only if $p_{i,x}(t) = 1 - \frac{\bar{F}_i(t+x)}{\bar{F}_i(t)}$, $1 \leq i \leq n$ are increasing(decreasing) in t. Now in view of Lemma 2.2 and Equation (2.2) the proof of the lemma simply follows.
Lemma 2.3 implies that if F_i's are IFR(DFR) distributions then $H_n^k(t)$ is decreasing(increasing) in t.

We now consider the MRL function $H_{n,k}^r(t)$. The following equation gives in inid case an expression which is similar to that of given in iid case stated in Equation (1.1).

$$P(X_{k:n} - t > x | N(t) = r, C_r = \{i_1, \ldots, i_r\})$$

$$= P(X_{(k-r:n-r)} > t + x | X_{(1:n-r)} > t) \tag{2.3}$$

where $X_{(j:n-r)}$ is jth order statistic from random sample $\{X_i | i \in C - C_r, i = 1, \ldots, n-r\}$.
From above equation we have

$$H_{n,k}^r(t) = E(X_{k:n} - t | N(t) = r, C_r = \{i_1, \ldots, i_r\})$$

$$= E(X_{(k-r:n-r)} | X_{(1:n-r)} > t) \tag{2.4}$$

Lemma 2.4. *If F_i's are IFR(DFR) distributions then $P(X_{k:n} - t > x | N(t) = r, C_r = \{i_1, \ldots, i_r\})$ is decreasing(increasing) in t.*

Proof. In view of Equation (2.3) and Lemma 2.3 the proof of the lemma follows.
Based on Lemmas 2.3 and 2.4 and Equation (2.4) we note that if F_i's are IFR(DFR) distributions then $H^r_{n,k}(t)$ is decreasing(increasing) in t.

REFERENCES

[1] Asadi, M. and Goliforushani, S. (2008), On the mean residual life function of coherent systems. *IEEE Trans. on Rel.*, **57(4)**, 574-580.

[2] Barlow, R.E. and Proschan, F. (1975), *Statistical Theory of Reliability and Life Testing*, Holt, Rinehart and Winston.

[3] Ding, W., Li, X. and Balakrishnan, N. (2012), Monotonicity Properties of Residual Lifetimes of Parallel Systems and Inactivity Times of Series Systems with Heterogeneous Components. *Prob. in Engineer. and Inform. Sciences*, **26**, 61-75.

[4] Kuchar,S. and Xu, M. (2010), On residual lifetimes of k-out-of-n systems with nonidentical components. *Prob. in Engineer. and Inform. Sciences*, **24**, 109-127.

[5] Goliforushani, S., Asadi, M. and Balakrishnan, N. (2011) On the residual and inactivity times of the components of used coherent sytstems. *J. Appl. Prob.*

[6] Sadegh, M.K. (2008-a), Mean past and mean residual life functions of a parallel system with nonidentical components. *Commu. in Stat.; Theory & Methods*, **37**, 1134-1145.

[7] Sadegh, M.K. (2008-b), A note on the regression of order statistics. *Commu. in Stat.; Theory & Methods*, **37**, 475-480.

[8] Sadegh, M.K. (2011), A note on the mean residual life function of a coherent system with exchangeable or nonidentical components. *Journal of Statistical Planning and Inference*, **141**, 3267-3275.

[9] Samaniego, F.J. (1985), on the closure of the IFR class under formation of coherent systems. *IEEE Trans. on Rel.*, **34**, 69-72.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Predicting the lifetime of a k-out-of-n:F system in the presence of an outlier

Khatib Astane, B. [1]

Department of Statistics, University of Neyshabur

Abstract

Consider a k-out-of-n:F system which the lifetimes of the components are not identically distributed such that there is only one outlier. Based on an observed data set, the prediction problem of the lifetime of the system is investigated in this paper. The exponential distribution is used to obtain the results in details. Finally an illustrative example is presented.

Keywords: Order statistics, Single outlier model, Proportional hazard rate model, Exponential distribution.

1 Introduction and preliminaries

As known in the literature of reliability, a k-out-of-n:F system consists of n components which fails if and only if at least k of its components fail. Such systems have various applications in engineering. For more details, we refer to Lawless (2003). Recently, the residual lifetime and inactivity time of such systems have been studied by some authors. See, for example, Asadi and Bayramoglu (2006), Bairamov and Arnold (2008) and Tavangar and Bariamov (2015). Many researchers have studied the k-out-of-n:F system assuming the lifetimes of the components are independent and identically distributed random variables. But, there are some situations in life-testing and reliability experiments in which the observations are independent but not identically distributed. For example, when the distribution of one of the observations is different from the

[1] khatib_b@neyshabur.ac.ir

others, one face with a single outlier model. In this paper, we would like to predict the lifetime of a survive k-out-of-n:F system which contains an outlier.

Let us denote the lifetimes of the components of a k-out-of-n:F system by $X_1, \ldots, X_n$ which are independent random variables, such that $X_1, \ldots, X_{n-1}$ come from a population with cumulative distribution function (cdf) $F(x)$ and probability density function (pdf) $f(x)$; moreover, X_n is an outlier from a different population with cdf $G(x)$ and pdf $g(x)$. Denote the corresponding order statistics by $X_{1:n} < \cdots < X_{n:n}$. Then, it is obvious that the lifetime of a k-out-of-n:F system is $X_{k:n}$. If one observe some of the smallest lifetime whereas the system is still functioning, then the problem of predicting the lifetime of the system may be of interest. More precisely, assume that one only observe

$$\mathbf{X} = \{X_{1:n}, \ldots, X_{r:n}\},$$

such that $1 \leq r < k$. To predict $X_{k:n}$ for $r < k \leq n$, let us first recall the pdfs of order statistics in the presence of an outlier from Balakrishnan (2007). The pdf of $X_{r:n}$ $(1 \leq r \leq n)$ is given by

$$\begin{aligned} f_{X_{r:n}}(y) &= \frac{(n-1)!}{(r-2)!(n-r)!}(F(y))^{r-2}G(y)f(y)(1-F(y))^{n-r} \\ &\quad + \frac{(n-1)!}{(r-1)!(n-r)!}(F(y))^{r-1}g(y)(1-F(y))^{n-r} \\ &\quad + \frac{(n-1)!}{(r-1)!(n-r-1)!}(F(y))^{r-1}f(y)(1-F(y))^{n-r-1}(1-G(y)), \end{aligned}$$

where the first and last terms vanish when $r = 1$ and $r = n$, respectively. The joint pdf of $\mathbf{X} = (X_{1:n}, \ldots, X_{r:n})$, at the point $\mathbf{y} = (y_1, \ldots, y_r)$, is as follows

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{y}) &= \frac{(n-1)!}{(n-r)!}\sum_{i=1}^{r} g(y_i)\prod_{\substack{j=1 \\ j\neq i}}^{r} f(y_j)(1-F(y_r))^{n-r} \\ &\quad + \frac{(n-1)!}{(n-r-1)!}\prod_{i=1}^{r} f(y_i)(1-F(y_r))^{n-r-1}(1-G(y_r)), \end{aligned} \tag{1.1}$$

where the last term vanishes when $r = n$. Moreover, the joint pdf of $(\mathbf{X}, X_{k:n})$, for $1 \leq r < k \leq n$,

is given by

$$\begin{aligned} f_{(\mathbf{X},X_{k:n})}(\mathbf{y},y_k) &= (n-1)! \\ &\times\Bigg\{\sum_{i=1}^{r} g(y_i)\prod_{\substack{j=1\\ j\neq i}}^{r} f(y_j)\frac{(F(y_k)-F(y_r))^{k-r-1}f(y_k)(\bar{F}(y_k))^{n-s}}{(s-r-1)!(n-k)!} \\ &+\prod_{i=1}^{r} f(y_i)\frac{(F(y_k)-F(y_r))^{k-r-2}(G(y_k)-G(y_r))f(y_k)(\bar{F}(y_k))^{n-k}}{(k-r-2)!(n-k)!} \\ &+\prod_{i=1}^{r} f(y_i)\frac{(F(y_k)-F(y_r))^{k-r-1}g(y_k)(1-F(y_k))^{n-k}}{(s-r-1)!(n-k)!} \\ &+\prod_{i=1}^{r} f(y_i)\frac{(F(y_k)-F(y_r))^{k-r-1}f(y_k)(1-F(y_k))^{n-k-1}\bar{G}(y_k)}{(k-r-1)!(n-k-1)!}\Bigg\}. \end{aligned} \tag{1.2}$$

where the second and last terms vanish when $k=r+1$ and $k=n$, respectively.

The rest of paper is as follows. In Section 2, the best unbiased predictor is presented. In Section 3, the predictor is derived for the case of Exponential distribution. In this section, it is assumed that $(n-1)$ component of a k-out-of-n:F system have exponential lifetimes and the lifetime of the other component is different which come from a proportional hazard rate model. The results are illustrated via a real data set in Section 4. Finally, some conclusions are stated in Section 5.

2 Best unbiased predictor

Assume that the data set $\mathbf{X}=\{X_{1:n},\ldots,X_{r:n}\}$ is available. It is well-known that $\hat{X}_{k:n}=\mathrm{E}[X_{k:n}|\mathbf{X}]$ is the best unbiased predictor for $X_{k:n}$ $(r<k\leq n)$, in the sense that it is unbiased (i.e., $\mathrm{E}(\hat{X}_{k:n})=\mathrm{E}(X_{k:n})$) and minimizes the mean square prediction error $E\left((\hat{X}_{k:n}-X_{k:n})^2\right)$. For more details about this predictor, we refer the readers to MirMostafaee and Ahmadi (2011) and Khatib and Ahmadi (2015). Using (1.2) and (1.1), we find

$$\hat{X}_{s:n}=\sum_{i=1}^{4}\varphi_i(r), \tag{2.1}$$

such that the functions $\varphi_i(r)$ for $i=1,2,3,4$ are presented as follows. By assuming $U_{r,k}\sim Beta(k-r,n-k+1)$, we get

$$\varphi_1(r)=\frac{b_1(r)}{a(r)}\mathrm{E}\Big(F^{-1}\big(U_{r,k}(1-F(y_r))+F(y_r)\big)\Big),$$

$$\begin{aligned}\varphi_2(r) &= \frac{b_2(r)}{a(r)}\mathrm{E}\Big\{F^{-1}\big(U_{r+1,k}(1-F(y_r))+F(y_r)\big) \\ &\times\Big(G\big(F^{-1}\big(U_{r+1,k}(1-F(y_r))+F(y_r)\big)\big)-G(y_r)\Big)\Big\},\end{aligned}$$

$$\varphi_3(r) = \frac{b_3(r)}{a(r)}\mathrm{E}\Big\{F^{-1}\big(U_{r,k}(1-F(y_r))+F(y_r)\big) \times \frac{g\big(F^{-1}\big(U_{r,k}(1-F(y_r))+F(y_r)\big)\big)}{f\big(F^{-1}\big(U_{r,k}(1-F(y_r))+F(y_r)\big)\big)}\Big\}$$

and

$$\varphi_4(r) = \frac{b_2(r)}{a(r)}\mathrm{E}\Big\{F^{-1}\big(U_{r,k}(1-F(y_r))+F(y_r)\big) \times\Big(1-G\big(F^{-1}\big(U_{r,k}(1-F(y_r))+F(y_r)\big)\big)\Big)\Big\},$$

where

$$a(r)=\sum_{i=1}^{r} g(y_i)\prod_{j=1_{j\neq i}}^{r} f(y_j)+(n-r)\prod_{i=1}^{r} f(y_i)(1-F(y_r))^{-1}(1-G(y_r)),$$

$$b_1(r)=\sum_{i=1}^{r} g(y_i)\prod_{j=1_{j\neq i}}^{r} f(y_j),$$

$$b_2(r)=(n-r)\prod_{i=1}^{r} f(y_i)(1-F(y_r))^{-1}$$

and

$$b_3(r)=\prod_{i=1}^{r} f(y_i).$$

3 Exponential distribution

In this section, we assume that $X_1,\ldots,X_{n-1}$ have the exponential distribution with the cdf

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

The reader may refer to Balakrishnan and Basu (1995) for more details about exponential distribution. Moreover, we consider a proportional hazard rate model for the distribution of the outlier. That is, we assume that X_n comes from the cdf

$$\bar{G}(x)=(\bar{F}(x))^{\alpha}, \tag{3.1}$$

where $\bar{F}(x)=1-F(x)$ is the survival function of the first $(n-1)$ observations. In such a situation, from (3.2), the best unbiased predictor of $X_{k:n}$ is given by

$$\hat{X}_{k:n}=\sum_{i=1}^{4}\xi_i(r), \tag{3.2}$$

where

$$\xi_1(r)=\frac{\alpha\psi}{\alpha\psi+(n-r)}\Bigg\{X_{r:n}+\frac{1}{\theta}\frac{(n-r)!}{(k-r-1)!(n-k)!}\sum_{i=0}^{k-r-1}\frac{(-1)^i\binom{k-r-1}{i}}{(n-k+i+1)^2}\Bigg\},$$

$$\begin{aligned}\xi_2(r) &= \frac{1}{\frac{\alpha}{n-r}\psi+1}\Bigg\{X_{r:n}-(B(k-r-1,n-k+1))^{-1}\\ &\quad\times\Bigg[X_{r:n}B(k-r-1,n-k+\alpha+1)-\frac{1}{\theta}\sum_{i=0}^{k-r-2}\binom{k-r-2}{i}\\ &\quad\times(-1)^i\Big(\frac{1}{(n-k+i+1)^2}-\frac{1}{(n-k+i+\alpha+1)^2}\Big)\Bigg]\Bigg\},\end{aligned}$$

$$\begin{aligned}\xi_3(r) &= \frac{\alpha(B(k-r,n-k+1))^{-1}}{\alpha\psi+(n-r)}\Bigg\{X_{r:n}B(k-r,n-k+\alpha-2)\\ &\quad+\frac{1}{\theta}\sum_{i=0}^{k-r-1}\binom{k-r-1}{i}(-1)^i\frac{1}{(n-k+i+\alpha+1)^2}\Bigg\}\end{aligned}$$

and

$$\begin{aligned}\xi_4(r) &= \frac{(B(k-r,n-k))^{-1}}{\frac{\psi}{n-r}+1}\Bigg\{X_{r:n}\frac{(k-r-1)!(n-k+\alpha-1)!}{(n-r+\alpha-1)!}\\ &\quad+\frac{1}{\theta}\sum_{i=0}^{k-r-1}\binom{k-r-1}{i}\frac{(-1)^i}{(n-k+\alpha+i)^2}\Bigg\},\end{aligned}$$

such that

$$\psi=\sum_{i=1}^{r}e^{\theta(\alpha-1)(X_{r:n}-X_{i:n})}.$$

In the next section, we present a simulated example to illustrate the obtained results.

4 Illustrative example

Here, we generate the lifetimes of the components of a system with 10 components such that 9 components come from a exponential distribution with $\theta=0.5$ and the outlier comes from the model (3.1) with $\alpha=0.2$. The ordered data are

$$0.418, 0.723, 1.052, 2.923, 3.451, 3.483, 5.262, 5.469, 6.299, 16.970.$$

Now, consider a 5-out-of-10:F system such that the above data are the lifetime of its components. Obviously, the exact lifetime of this system is $X_{5:10}=3.451$. Consider the time point that we only observe the first third smallest lifetimes, i.e., $(X_{1:10},X_{2:10},X_{3:10})=(0.418,0.723,1.052)$. Based

on these observations and using (3.2), the predicted lifetime of the system is 1.752. So, it is important to know if we have more observations then can we get a better predictor. In this example, if $(X_{1:10}, \ldots, X_{4:10}) = (0.418, 0.723, 1.052, 2.923)$ are observed, then the predicted lifetime of the system is 3.304 which is closer to the exact value.

5 Concluding Remark

In this paper the prediction problem of the lifetime of a k-out-of-n:F system was investigated in the presence of an outlier. The best unbiased predictor was derived in general. The results were obtained in details for the case of exponential distribution. Eventually an illustrative example was presented. In this example it was observed that according to a larger observed data set a more accurate predictor may be obtained. However, this issue seems logical, but it may be discussed via different criteria such as Mean Squared Prediction Error or Pitman Closeness.

References

[1] Asadi, M. and Bayramoglu, I. (2006). The mean residual life function of a k-out-of-n structure at the system level, *IEEE Trans. Reliab.* **55**, 314-318.

[2] Bairamov, I. and Arnold, B.C. (2008). On the residual life lengths of the remaining components in an $(n-k+1)$-out-of-n system, *Stat. Probab. Lett.*, **78**, 945-952.

[3] Balakrishnan, N. (2007), Permanents, Order Statistics, Outliers, and Robustness, *Rev. Mat. Compulut.*, **20**, 7-107.

[4] Balakrishnan, N. and Basu, A. P. (1995), *Exponential Distribution: Theory, Methods and Applications*, *Taylor & Francis, New York.*

[5] Khatib, B. and Ahmadi, J. (2014), Best linear unbiased and invariant reconstructor for the past records, *Bullet. Malay. Math. Sci. Soc.*, **37**, 1017-1028.

[6] Lawless, J. F. (2003). *Statistical models and methods for lifetime data.* Wiley Series in Probability and Statistics, Second Edition.

[7] MirMostafaee, S. M. T. K. and Ahmadi, J. (2011), Point prediction of future order statistics from an exponential distribution, *Statist. Prob. Lett.*, **81**, 360-370.

[8] Tavangar, M. and Bairamov, I. (2015). On conditional residual lifetime and conditional inactivity time of k-out-of-n systems, *Reliab. Eng. Syst. Safe.*, **144**, 225-233.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Some reliability measures in general form of distributions

Khorashadizadeh, M. [1]

[1] Department of Statistics, University of Birjand

Abstract

In this paper, some properties of concepts in reliability such as log-odds rate, proportional hazard model and stress- strength model are studied in a general form of distributions.

Keywords: Hazard rate function, Reversed hazard rate function, Log-odds rate function, Stress-Strength model.

1 Introduction

In probability and statistics, it is very important to obtain statistical and mathematical properties for a general class of distributions. For example the exponential family of distribution and its properties is well known in statistics. Al-Hussaini (1999) has proposed a general class of distributions, which includes more distribution than exponential family. His class of distribution in special case is defined as follow:

Let X be a non-negative continuous random variable with distribution function of form,

$$F(x) = 1 - e^{-\alpha k_\theta(x)}; \quad x > 0, \tag{1.1}$$

where $\alpha > 0$ and θ is a parameter vector (which may contains α) and $k_\theta(x)$ is a strictly increasing function of x with $k_\theta(0) = 0$ and $k_\theta(\infty) = \infty$. The probability density function of x is then given by:

$$f(x) = \alpha k'_\theta(x) e^{-\alpha k_\theta(x)}; \quad x > 0, \tag{1.2}$$

[1] m.khorashadizadeh@birjand.ac.ir

where $k'_\theta(x)$ is the derivative of $k_\theta(x)$ with respect to x. Note that if $\overline{F}(x) = P(X > x)$ has closed form, then $k_\theta(x) = -\frac{\ln \overline{F}(x)}{\alpha}$. The form of (1.1) can be viewed as the *exponential form of distributions*. Table 1 shows the $K_\theta(x)$ for some distributions.

Rezaei Roknabadi et al. (2009) has introduced the discrete version of exponential form of distributions by using transformation $Y = [X]$ and named it Telescopic form of distribution.

Table 1: $k_\theta(x)$ function for some lifetime distribution

PDF	$k_\theta(x)$	Name
$\alpha e^{-\alpha x}$	x	Exponential
$\alpha\theta x^{\theta-1}e^{-\alpha x^\theta}$	x^θ	Weibull
$\alpha x e^{-\alpha\frac{x^2}{2}}$	$\frac{x^2}{2}$	Rayleigh
$\alpha\theta e^{\theta x}e^{-\alpha(e^{\theta x}-1)}$	$e^{\theta x}-1$	Gompertz
$2\alpha x^{2r-1}(\frac{\theta}{x^2}+r)\times\exp\{-\frac{\theta}{x^2}-\alpha x^{2r}e^{-\frac{\theta}{x^2}}\}$	$x^{2r}e^{-\frac{\theta}{x^2}}$	Brittle-Fracture
$(\gamma+\theta x)e^{-\gamma x+\frac{\theta x^2}{2}}$	$x+\frac{\theta}{2\gamma}x^2$	Exponential-linear
$2\alpha x e^{-x^2}(1-e^{-x^2})^{\alpha-1}$	$-\ln(1-e^{-x^2})$	Burr X
$\frac{\alpha\theta x^{\theta-1}}{(1+x^\theta)^{\alpha+1}}$	$\ln(1+x^\theta)$	Burr XII
$\theta x^{-1-\theta}e^{-x^{-\theta}}$	$\ln(-e^{x^{-\theta}})$	Frechet
$\frac{2e^x}{(1+e^x)^2}$	$\ln(\frac{1+e^x}{2})$	Half-logistic

Many researchers have obtained the different properties of this general class of distribution such as Soliman (2002), Al-Hussaini (2010), Al-Hossain (2016) and Hashempour and Doostparast (2016).

In this form of exponential distribution some reliability measures have the following form:

- Reliability function: $\overline{F}(t) = \exp\{-\alpha k_\theta(t)\}$.
- Hazard rate function: $h(t) = \frac{f(t)}{R(t)} = \alpha k'_\theta(t)$.
- Reversed hazard rate function: $r(t) = \frac{f(t)}{F(t)} = \frac{\alpha k'_\theta(t)e^{-\alpha k_\theta(t)}}{1-e^{-\alpha k_\theta(t)}} = \frac{h(t)}{e^{\alpha k_\theta(t)}-1}$.
- Mean residual life: $\mu(t) = \int_t^\infty \exp\{\alpha(k_\theta(t)-k_\theta(u))\}du$.

In the rest of the paper we obtain some results for this general class of distribution for other reliability concepts such as log-odds rate, proportional hazard model and stress-strength model.

2 Log-odds rate function in general form

Recently, there has been some interest in studying the potential of log-odds function and log-odds rate in reliability theory. The motivations for the construction of these functions are: (i) they are easy to compute and interpret, (ii) the estimation of these functions is relatively simple, (iii) the behaviour of the other reliability functions can be ascertained through them, and (iv) the odds ratio can also be used to elicit subjective probability of failure (survival) in the context of inference in a Bayesian framework. The aging properties and stochastic orders for comparing various life distributions based on various reliability functions can be translated in terms of the odds function.

In the reliability literature, the failure or hazard rate function is one of the most important measure in characterizing the lifetime distributions and in modeling statistical data, to derive the appropriate model. Also, the monotonic behaviour of the failure rate, i.e. increasing failure rate (IFR) and decreasing failure rate (DFR), are the fundamental classes in the reliability analysis.

In engineering reliability, there are many situations where the reliability of the components is high or their hazard rate is non-monotone. These two problems indicate that there is a need for other reliability indicators than hazard rate. The LO and LOR functions are such indicators. Martz and Waller (1982) and Zimmer et al. (1998) studied the important and the most frequent continuous distributions in modeling the failure times data, such as log normal, Burr XII and inverse Gaussian, which have non-monotone hazard rate.

Recently, in order to solve the above problem, Zimmer et al. (1998) and Wang et al. (2003, 2008) introduced a new model for continuous time to failure based on the log-odds rate (LOR) which is comparable to the model based on the failure rate.

LOR function is defined as follows,

$$\mathrm{LOR}(x) = \frac{\partial}{\partial x}\mathrm{LO}(x),$$

where $\mathrm{LO}(x) = \ln \frac{F(x)}{R(x)}$ is the log-odds function. Hence,

$$\mathrm{LOR}(x) = \frac{f(x)}{F(x)R(x)} = \frac{h(x)}{F(x)} = \frac{r(x)}{R(x)}. \tag{2.1}$$

So, for distributions of form (1.1) we have:

$$\mathrm{LO(t)} = \ln(e^{\alpha k_\theta(t)} - 1), \tag{2.2}$$

$$\mathrm{LOR(t)} = \frac{\alpha k'_\theta(t)}{1 - e^{-\alpha k_\theta(t)}}. \tag{2.3}$$

Theorem 2.1. *If X be lifetime random variable then increasing log odds rate (ILOR) implies the increasing failure rate (IFR) and also decreasing failure rate (DFR) implies decreasing log odds rate (DLOR). That is,*

$$\begin{aligned} \mathrm{ILOR} &\implies \mathrm{IFR} \\ \mathrm{DFR} &\implies \mathrm{DLOR} \end{aligned}$$

Proof. The theorem can be proved by the fact that multiplying two non negative increasing function (decreasing function) is also increasing (decreasing). □

Theorem 2.2. *For any lifetime random variable, the log-odds rate function is increasing if and only if for any t,*

$$-r'(t) \leq r(t)h(t) \leq h'(t),$$

and decreasing if and only if,

$$h'(t) \leq r(t)h(t) \leq -r'(t).$$

Proof. Differentiating from (2.1) we have,

$$\mathrm{LOR}'(\mathrm{t}) = \frac{h'(t) - r(t)h(t)}{F(t)} = \frac{r'(t) + r(t)h(t)}{\overline{F}(t)},$$

which follows the required results. □

Remark 2.3. If the probability density function of random variable X is of form (1.2), then X has increasing log odds rate function if and only if,

$$\frac{k'^2_\theta(t)}{k''_\theta(t)} \left(\frac{1}{e^{\alpha k_\theta(t)} - 1} \right) < 1. \tag{2.4}$$

Zimmer et al. (1998) have shown that the logistic distribution is the only distribution that have constant log-odds rate. Consider the standard logistic cdf, reliability, pdf, and hazard functions:

$$\begin{aligned} F(x) &= \frac{e^x}{1+e^x}, \quad \overline{F}(x) = \frac{1}{1+e^x}, \quad f(x) = \frac{e^x}{(1+e^x)^2}, \\ h(x) &= \frac{f(x)}{\overline{F}(x)} = F(x), \end{aligned}$$

The log-odds function for the standard logistic distribution is:

$$\mathrm{LO}(x) = \ln \frac{F(x)}{\overline{F}(x)} = x,$$

thus the logistic distribution has linear LO and therefore constant LOR in x. Using the result of Remark 2.6 it is easy to see the three parameters Burr XII distribution with reliability function $\overline{F}(x) = \frac{1}{(1+(\theta x)^\gamma)^\alpha}$ is ILOR with respect to $\ln x$ for $\alpha > 1$ and DLOR for $\alpha < 1$.

3 Proportional hazard model

Another reliability measure that has important role in modeling and dependance structure among two distributions is the proportional hazard model defined by Cox (1959). Two random variables X and Y satisfy the proportional hazard rate model (PHM) with proportionality constant $\beta(>0)$, if for all x,

$$\overline{G}(x) = [\overline{F}(x)]^{\beta}; \quad \beta > 0. \tag{3.1}$$

Or equivalently,

$$g(x) = \beta f(x)[\overline{F}(x)]^{\beta-1}.$$

Let X and Y have exponentiated distribution functions as follow:

$$F(t) = 1 - \exp\{-\alpha_1 k_{\theta_1}(t)\}, \tag{3.2}$$
$$G(t) = 1 - \exp\{-\alpha_2 k^*_{\theta_2}(t)\}. \tag{3.3}$$

Then, X and Y are satisfying the PHM if and only if,

$$\frac{k_{\theta_1}(t)}{k^*_{\theta_2}(t)} = \beta\frac{\alpha_1}{\alpha_2}, \tag{3.4}$$

and if $\alpha_1 = \alpha_2 = \alpha$, PHM is satisfy if and only if

$$k_{\theta_1}(t) = \beta k_{\theta_2}(t).$$

For example if $X \sim Weibull(\alpha, 2)$ and $Y \sim Rayleigh(\alpha)$ then $k_{\theta_1}(x) = x^2$ and $k_{\theta_2}(y) = \frac{y^2}{2}$, so X and Y are satisfying the PHM with proportional constant $\beta = 2$.

4 Stress-Strength model in general form

The stress-strength model involving two independent random variables and is defined as $R = P(X_1 < X_2)$, where X_1 represents the stress variable and X_2 represents the strength variable . The literature review on the stress-strength models covers over 58 years of research. Birnbaum (1956) was one of the first researchers who dealt with the model $P(X_1 < X_2)$ in stress-strength content. Nowadays the stress-strength model is of substantial interest and usefulness in which provides a general measure of the difference between two populations and has applications in many areas such as clinical trials, genetics, and reliability. For example, if Y is the response for a control group, and X refers to a treatment group, R is a measure of the effect of the treatment. Or, if Y is the water pressure on the dam wall, and X be the strength of the dam, then the parameter R is of very important in maintenance. Different examples of applications of $P(X_1 < X_2)$ in engineering and medicine is presented in Johnson (1988), and the monograph by Kotz et al. (2003). We should also mention the recent works of Sengupta (2008), Kundu and Raqab (2009), Rezaei et al. (2010) and Panahi and Asadi (2010) which have obtained results due to estimation of $P(X > Y)$ and characterizations related to it.

If X and Y are two independent continuous random variables with pdf of forms,

$$f_X(x) = \alpha_1 k'_\theta(x) e^{-\alpha_1 k_\theta(x)}; \quad x > 0,$$

and

$$f_Y(y) = \alpha_2 k'_\theta(y) e^{-\alpha_2 k_\theta(y)}; \quad y > 0,$$

respectively, then we have,

$$\begin{aligned} R = P(Y < X) &= \int_0^\infty P(Y < x) f_X(x) dx \\ &= \int_0^\infty (1 - e^{-\alpha_2 k_\theta(x)}) \alpha_1 k'_\theta(x) e^{-\alpha_1 k_\theta(x)} dx \\ &= \frac{\alpha_2}{\alpha_1 + \alpha_2}. \end{aligned} \tag{4.1}$$

Theorem 4.1. *Let $h_1(t)$ and $h_2(t)$ are the hazard rate of X_1 and X_2 respectively, then the Stress-Strength reliability function is of form,*

$$R = P(X_2 < X_1) = \frac{\alpha_2}{\alpha_1 + \alpha_2} = \frac{h_2(t)}{h_1(t) + h_2(t)}, \tag{4.2}$$

if and only if the distribution functions of X_1 and X_2 is of form (1.1) with parameters $\theta_1 = \theta_2 = \theta$ and α_1 and α_2.

Proof. Using (4.1) the proof of 'if' part is satisfied. Now, suppose that for two random variables X_1 and X_2, the Stress-Strength reliability function is of form (4.2). Let $g(t, \theta)$ be an non-negative integrable function, then for any $t > 0$ we have,

$$\frac{\alpha_2 g(t,\theta)}{\alpha_1 g(t,\theta) + \alpha_2 g(t,\theta)} = \frac{h_2(t)}{h_1(t) + h_2(t)},$$

which leads to,

$$\alpha_i g(t,\theta) = h_i(t) = \frac{f_i(t)}{\overline{F}_i(t)}, \quad i = 1, 2.$$

So, integrating both sides of the above equation over the range $(0, x)$, we get,

$$\ln \overline{F}_i(x) = -\alpha_i \int_0^x g(t,\theta) dt \Longrightarrow \overline{F}_i(x) = e^{-\alpha_i \int_0^x g(t,\theta) dt}, \quad i = 1, 2. \tag{4.3}$$

Hence, by setting $k_\theta(x) = \int_0^x g(t,\theta) dt$ the required result is obtained. □

To obtain the MLE for R, firstly we need to obtain the MLEs for α_1 and α_2, which are depend on MLE of θ. Suppose $x_1, \ldots, x_n$ is a random sample from $f_X(x)$ and $y_1, \ldots, y_m$ is a random sample

from $f_Y(y)$ and also let $k_\theta(.)$ does not depend on α, then the log-likelihood function is,

$$\begin{aligned} L(\theta,\alpha_1,\alpha_2) &= n\ln(\alpha_1)+m\ln(\alpha_2)+\sum_{i=1}^{n}\ln(k'_\theta(x_i))+\sum_{j=1}^{m}\ln(k'_\theta(y_j)) \\ &- \alpha_1\sum_{i=1}^{n}k_\theta(x_i)-\alpha_2\sum_{j=1}^{m}k_\theta(y_j). \end{aligned}$$

The MLEs of parameters, say $\widehat{\alpha}_1$, $\widehat{\alpha}_2$ and $\widehat{\theta}$, can be obtain as the solutions of

$$\frac{\partial L}{\partial \alpha_1} = \frac{n}{\alpha_1}-\sum_{i=1}^{n}k_\theta(x_i)=0, \tag{4.4}$$

$$\frac{\partial L}{\partial \alpha_2} = \frac{m}{\alpha_2}-\sum_{j=1}^{m}k_\theta(y_j)=0, \tag{4.5}$$

$$\frac{\partial L}{\partial \theta} = \sum_{i=1}^{n}\frac{k_\theta^{'(1)}(x_i)}{k'_\theta(x_i)}+\sum_{j=1}^{m}\frac{k_\theta^{'(1)}(y_j)}{k'_\theta(y_j)}-\alpha_1\sum_{i=1}^{n}k_\theta^{(1)}(x_i)-\alpha_2\sum_{j=1}^{m}k_\theta^{(1)}(y_j)=0, \tag{4.6}$$

where $k_\theta^{(1)}(.)$ is the derivative of $k_\theta(.)$ with respect to θ. Using (4.4) and (4.5) we have,

$$\widehat{\alpha}_1(\theta)=\frac{n}{\sum_{i=1}^{n}k_\theta(x_i)} \quad \text{and} \quad \widehat{\alpha}_2(\theta)=\frac{m}{\sum_{j=1}^{m}k_\theta(y_j)}. \tag{4.7}$$

Substituting $\widehat{\alpha}_1(\theta)$ and $\widehat{\alpha}_2(\theta)$ in to (4.6), we obtain,

$$\sum_{i=1}^{n}\frac{k_\theta^{'(1)}(x_i)}{k'_\theta(x_i)}+\sum_{j=1}^{m}\frac{k_\theta^{'(1)}(y_j)}{k'_\theta(y_j)}=n\frac{\sum_{i=1}^{n}k_\theta^{(1)}(x_i)}{\sum_{i=1}^{n}k_\theta(x_i)}+m\frac{\sum_{j=1}^{m}k_\theta^{(1)}(y_j)}{\sum_{j=1}^{m}k_\theta(y_j)}.$$

Therefor, when $k_\theta(.)$ is known, $\widehat{\theta}$ can be estimated by iterative process and the scale parameters $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$ can be obtained from (4.7). So the MLE of R is,

$$\widehat{R}=\frac{\frac{m}{\sum_{j=1}^{m}k_{\widehat{\theta}}(y_j)}}{\frac{n}{\sum_{i=1}^{n}k_{\widehat{\theta}}(x_i)}+\frac{m}{\sum_{j=1}^{m}k_{\widehat{\theta}}(y_j)}}. \tag{4.8}$$

5 Conclusions

In this paper, some reliability concepts such as log-odds rate, proportional hazard model and stress-strength model are studied in a big class of distribution. Considering more properties of general form of distributions in discrete and continuous lifetime distribution are the future of the research.

References

[1] Al-Hossain, A. Y. (2016). Predictive Inference from the Exponentiated Weibull Model Given Adaptive Progressive Censored Data. Appl. Math, 10(3), 1177-1184.

[2] Al-Hussaini, E. K. (1999). Predicting observables from a general class of distributions. Journal of Statistical Planning and Inference, 79(1), 79-91.

[3] Al-Hussaini, E. K. (2010). Inference based on censored samples from exponentiated populations. Test, 19(3), 487-513.

[4] Birnbaum, Z. W. (1956). On a use of the Mann-Whitney statistic. In Proceedings of the third Berkeley symposium on mathematical statistics and probability (Vol. 1, pp. 13-17). Berkeley and LOB Angeles: University of California Press.

[5] Cox, D. R. (1959). The analysis of exponentially distributed life-times with two types of failure. Journal of the Royal Statistical Society. Series B (Methodological), 411-421.

[6] Hashempour, M., and Doostparast, M. (2016). Evidences in sequential order statistics coming from a general class of distributions. Reliability Theory and its Applications, 84.

[7] Johnson, R. A. (1988). 3 Stress-strength models for reliability. Handbook of statistics, 7, 27-54.

[8] Kotz, S., Lumelskii, Y. and Pensky, M. (2003). The stress-strength model and its generalizations: theory and applications. World Scientific.

[9] Kundu, D. and Raqab, M.Z. (2009). Estimation of $R = P(Y < X)$ for three-parameter Weibull distribution. Statistics and Probability Letters, 79, 17, 1839-1846.

[10] Martz, H. F. and Waller, R. A. (1982). Bayesian Reliability Analysis. New York: John Wiley and Sons.

[11] Panahi, H. and Asadi, S. (2010). Estimation of $R = P(Y < X)$ for two-parameter Burr Type XII distribution. World Academy of Science, Engineering and Technology, 72, 465-470.

[12] Rezaei Roknabadi, A.H., Mohtashami Borzadaran, G.R. and Khorashadizadeh, M. (2009). Some aspects of discrete hazard rate function in Telescopic families. Economic Quality Control, 24, 1, 35-42.

[13] Rezaei, S., Tahmasbi, R. and Mahmoodi, M. (2010). Estimation of $P(Y < X)$ for generalized Pareto distribution. Statistical Planning and Inference 140: 480- 494.

[14] Sengupta, S. (2008). Unbiased variance estimation in estimating $P(X > Y)$ for one and two parameter exponential populations. Metron-International Journal of Statistics, 66, 3, 329-339.

[15] 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.

[16] Wang, Y., Hossain, A. M. and Zimmer, W. J. (2003). Monotone log-odds rate distributions in reliability analysis. Commun. Statist. Theor. Meth. 32(11): 2227-2244.

[17] Wang, Y., Hossain, A. M. and Zimmer, W. J. (2008). Useful properties of the three-parameter Burr XII distribution. In: Ahsanullah M., Applied Statistics Research Progress. pp. 11-20.

[18] Zimmer, W. J., Wang, Y. and Pathak, P. K. (1998). Log-odds rate and monotone log-odds rate distributions. J. Qual. Technol. 30(4): 376-385.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Assessing effect of ranking quality in parametric reliability estimation based on ranked set samples

Mahdizadeh, M. [1]

Department of Statistics, Hakim Sabzevari University

Abstract

This article deals with reliability estimation from exponential population when data are collected by ranked set sampling. Performances of some estimators in the presence of ranking errors are compared with that of the usual estimator in simple random sampling.

Keywords: Covariate information, Judgment ranking, Stress-strength model.

1 Introduction

Ranked set sampling (RSS), introduced by McIntyre [2], is applicable in settings where full measurement of a unit is expensive and time consuming while approximate ranking of a small set is inexpensive and easy. It often provides improved efficiency over simple random sampling (SRS), given a fixed sample size.

To construct a ranked set sample of size $p = mr$ (with set size m and cycle size r) the following procedure is repeated m times. One first selects m sets of units, each of size m, from an infinite population. Each of these sets is ranked from smallest to largest with some method that does not require any actual measurement. After completing the ranking in m sets, the ith ranked unit is fully measured in the ith set. The resulting sample is denoted by $\{X_{(i:m)j} : i = 1, \ldots, m; j = 1, \ldots, r\}$, where $X_{(i:m)j}$ is the ith judgment order statistic in the jth cycle.

As mentioned in the above, the ranking is done without any actual measurement. The ranking errors, which are inevitable, affect any RSS based procedure. This article aims to investigate this problem in the context of reliability estimation from exponential populations.

[1] mahdizadeh.m@live.com

2 The estimators

Let X and Y be independent exponential random variables with mean α and β, respectively. The corresponding probability density functions are given by $f_X(x) = \frac{1}{\alpha}\exp\{-x/\alpha\}\,(x>0;\alpha>0)$ and $f_Y(y) = \frac{1}{\beta}\exp\{-x/\beta\}\,(y>0;\beta>0)$. Then, it can be shown that $\theta = P(X>Y) = \alpha/(\alpha+\beta)$. In the sequel, estimators of θ in different designs are introduced.

2.1 SRS case

Suppose $X_1,\ldots,X_p \overset{iid}{\sim} f_X$ and $Y_1,\ldots,Y_q \overset{iid}{\sim} f_Y$ be two independent samples. Then the maximum likelihood estimator of θ is given by

$$\hat{\theta}_{\text{SRS}} = \frac{\bar{X}}{\bar{X}+\bar{Y}}, \tag{2.1}$$

where $\bar{X} = \sum_{i=1}^{p} X_i/p$ and $\bar{Y} = \sum_{i=1}^{q} Y_i/q$. The uniformly minimum variance unbiased estimator of θ (see Vasermanis et al. [5] for details) is not considered here as its performance is similar to (2.1) according to simulation results.

2.2 RSS case

Let $\{X_{(i:m)j} : i=1,\ldots,m; j=1,\ldots,r\}$ and $\{Y_{(i:n)j} : i=1,\ldots,n; j=1,\ldots,s\}$ be two independent ranked set samples of total sizes $p=mr$ and $q=ns$ drawn from f_X and f_Y, respectively. The following estimators were developed by Muttlak et al. [4] under perfect ranking assumption.

The simplest estimator of θ is given by

$$\hat{\theta}^1_{\text{RSS}} = \frac{\bar{X}_{\text{RSS}}}{\bar{X}_{\text{RSS}}+\bar{Y}_{\text{RSS}}}, \tag{2.2}$$

where $\bar{X}_{\text{RSS}} = \sum_{i=1}^{m}\sum_{j=1}^{r} X_{(i:m)j}/p$ and $\bar{Y}_{\text{RSS}} = \sum_{i=1}^{n}\sum_{j=1}^{s} Y_{(i:n)j}/q$.

The best linear unbiased estimators of α and β are given by

$$\hat{\alpha}_{\text{BLUE}} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{r} c_{i:m}X_{(i:m)j}/d_{i:m}}{\sum_{i=1}^{m}\sum_{j=1}^{r} c^2_{i:m}/d_{i:m}}$$

and

$$\hat{\beta}_{\text{BLUE}} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{s} c_{i:n}Y_{(i:n)j}/d_{i:n}}{\sum_{i=1}^{n}\sum_{j=1}^{s} c^2_{i:n}/d_{i:n}},$$

where $c_{i:m} = \sum_{t=1}^{i} 1/(m-t+1)$ and $d_{i:m} = \sum_{t=1}^{i} 1/(m-t+1)^2$. Based on $\hat{\alpha}_{\text{BLUE}}$ and $\hat{\beta}_{\text{BLUE}}$, an estimator of θ can be constructed as

$$\hat{\theta}^2_{\text{RSS}} = \frac{\hat{\alpha}_{\text{BLUE}}}{\hat{\alpha}_{\text{BLUE}}+\hat{\beta}_{\text{BLUE}}}. \tag{2.3}$$

By setting the first derivative of the log-likelihood function ℓ with respect to α and β equal to zero, we arrive at

$$\frac{\partial \ell}{\partial \alpha} = -mr\alpha + \sum_{i=1}^{m}\sum_{j=1}^{r}(m-i+1)X_{(i:m)j} - \sum_{i=1}^{m}\sum_{j=1}^{r}\frac{(i-1)X_{(i:m)j}}{e^{X_{(i:m)j}/\alpha}-1} = 0 \tag{2.4}$$

and

$$\frac{\partial \ell}{\partial \beta} = -ns\beta + \sum_{i=1}^{n}\sum_{j=1}^{s}(n-i+1)Y_{(i:n)j} - \sum_{i=1}^{n}\sum_{j=1}^{s}\frac{(i-1)Y_{(i:n)j}}{e^{Y_{(i:n)j}/\beta}-1} = 0. \tag{2.5}$$

If the solutions of the above equations (found numerically) are denoted by $\hat{\alpha}_{\text{MLE}}$ and $\hat{\beta}_{\text{MLE}}$, then

$$\hat{\theta}^3_{\text{RSS}} = \frac{\hat{\alpha}_{\text{MLE}}}{\hat{\alpha}_{\text{MLE}} + \hat{\beta}_{\text{MLE}}} \tag{2.6}$$

is another estimator for θ. If the terms $X_{(i:m)j}/\left(e^{X_{(i:m)j}/\alpha}-1\right)$ and $Y_{(i:n)j}/\left(e^{Y_{(i:n)j}/\beta}-1\right)$ in (A.1) and (A.2) are replaced by their expectations, then the solutions have explicit forms as

$$\hat{\alpha}_{\text{MMLE}} = \frac{4}{rm(m+3)}\sum_{i=1}^{m}\sum_{j=1}^{r}(m-i+1)X_{(i:m)j}$$

and

$$\hat{\beta}_{\text{MMLE}} = \frac{4}{sn(n+3)}\sum_{i=1}^{n}\sum_{j=1}^{s}(n-i+1)Y_{(i:n)j}.$$

The final estimator is defined as

$$\hat{\theta}^4_{\text{RSS}} = \frac{\hat{\alpha}_{\text{MMLE}}}{\hat{\alpha}_{\text{MMLE}} + \hat{\beta}_{\text{MMLE}}}. \tag{2.7}$$

3 Simulation study

In this section, we use Monte Carlo simulation to compare different estimators. In doing so, $\rho = \alpha/\beta \in \{0.15, 1, 6\}$ with $\beta = 1$ was used. Also, $(p, q) \in \{(10, 10), (10, 20)\}$ and $m = n = 2, 5$ were chosen.

The relative efficiencies (REs) reported in Muttlak et al. [4] are computed under perfect ranking setup, which does not happen in practice. The imperfect rankings model that we utilize is the fraction-of-random-rankings model described by Frey et al. [1]. Under this model, each set in the RSS procedure is independently ranked either perfectly, with probability λ, or randomly, with probability $1-\lambda$, where $\lambda \in [0, 1]$ is a model parameter. Perfect rankings are obtained by setting $\lambda = 1$, and random rankings are obtained by setting $\lambda = 0$. In particular, we considered $\lambda \in \{1, 0.75, 0.5\}$.

The efficiency of $\hat{\theta}^i_{\text{RSS}}$ relative to $\hat{\theta}_{\text{SRS}}$ was estimated as follows. For each setup, 5,000 pairs of samples were generated in SRS and RSS. The two estimators were computed from each pair of samples in the corresponding designs, and their mean squared errors (MSEs) were determined. The RE is defined as

$$REi = \frac{\widehat{MSE}(\hat{\theta}_{\text{SRS}})}{\widehat{MSE}(\hat{\theta}^i_{\text{RSS}})}, \quad i \in \{1, 2, 3, 4\}.$$

The estimated REs are reported in Table 1 from which the following conclusions can be made:

- As ranking quality deteriorates, all of the REs decline. There are (a few cases) that the RSS estimator is outperformed by its SRS analog. Overall, $\hat{\theta}^1_{\text{RSS}}$ has better performance in the presence of ranking errors.

- The larger set size, the higher impact of ranking errors. For example compare values of RE1 for $(p, q) = (10, 10)$ and $m = 2, 5$ when λ changes from 1 to 0.75.

- As with most inference procedures based on RSS, the RE is expected to be increasing in set size, given a fixed total sample size. We note, however, that this property does not hold here. For example, examine values of RE2-RE4 for $(p, q) = (10, 10)$ and $m = 2, 5$ when $\lambda = 0.5$.

4 Application

The stress-strength model defines the reliability of a component as the probability that the strength of the unit (X) is greater than the stress (Y) imposed on it. Although the use of stress-strength models was originally motivated by problems in physics and engineering, it is not limited to these contexts. It is worth mentioning that $\theta = P(X > Y)$ provides a general measure of the difference between two populations, and has found applications in different fields such as economics, quality control, psychology, medicine and clinical trials. For instance, if Y is the response of a control group, and X is that of a treatment group, then θ is a measure of the treatment effect. This situation is exemplified in the following.

Murray et al. [3] conducted an experiment in which apple trees are sprayed with chemical containing fluorescent tracer, Tinopal CBS-X, at 2% concentration level in water. Two nine-tree plots were chosen for spraying. One plot was sprayed at high volume, using coarse nozzles on the sprayer to give a large average droplet size. The other plot was sprayed at low volume, using fine nozzles to give a small average droplet size. Fifty sets of five leaves were identified from the central five trees of each plot, and used to draw a ranked set sample with set size 5 and cycle size 10, from each plot. The variable of interest is the percentage of area covered by the spray on the surface of the leaves. The formal measurement entails chemical analysis of the solution collected from the surface of the leaves, and thereby is a time-consuming and expensive process. The judgment ranking within each set is based on the visual appearance of the spray deposits on the leaf surfaces when viewed under ultraviolet light. Clearly, the latter method is cheap, and fairly accurate if implemented by an expert observer.

The data are given in Table 2, where measurements obtained from the plot sprayed at high (low) volume constitute the control (treatment) group. Figure 1 shows histogram of data in each group along with a fitted exponential density. The parameter of the exponential distribution is simply estimated using the sample mean. Clearly, the exponentiality assumption for the data in both groups is supported. So approximations of θ can be obtained using the estimators presented in sub-section 2.2. The estimated values are as follows: $\hat{\theta}^1_{\text{RSS}} = 0.601$, $\hat{\theta}^2_{\text{RSS}} = 0.590$, $\hat{\theta}^3_{\text{RSS}} = 0.592$ and $\hat{\theta}^4_{\text{RSS}} = 0.588$.

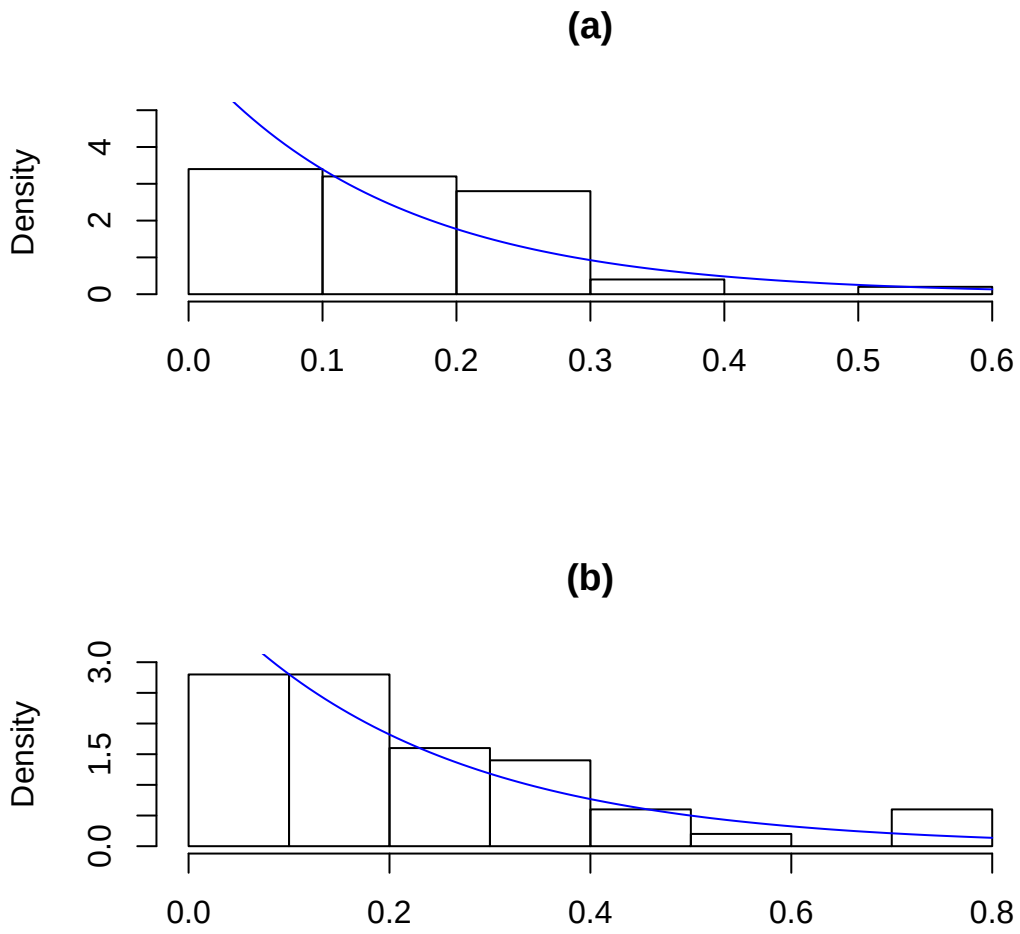


Figure 1: Histogram of apple trees data along with a fitted exponential density for control (a) and treatment (b) groups

5 Conclusion

The RSS is often used when a ranking of the sampling units can be obtained cheaply without having to actually measure the characteristic of interest, which may be time consuming or costly. Many statistical procedures based on RSS have been developed in the literature. This article is directed at the problem of reliability estimation from exponential populations when data are collected by RSS. In particular, effect of the judgment ranking on the precisions of different reliability estimators is investigated.

Given the simulation results, parametric reliability estimation in RSS is quite sensitive to the accuracy of ranking. In a subsequent work, we plan to evaluate robustness properties of nonparametric reliability estimation based on RSS.

References

[1] Frey, J., Ozturk, O., and Deshpande, J.V. (2007). Nonparametric tests for perfect judgment rankings. Journal of the American Statistical Association **102**, 708-717.

[2] McIntyre, G.A. (1952). A method for unbiased selective sampling, using ranked sets. Australian Journal of Agricultural Research **3**, 385-390.

[3] Murray, R.A., Ridout, M.S., and Cross, J.V. (2000). The use of ranked set sampling in spray deposit assessment. Aspects of Applied Biology **57**, 141-146.

[4] Muttlak, H.A., Abu-Dayyeh, W.A., Al-Saleh, M.F., and Al-Sawi, E. (2010). Estimating $P(Y < X)$ using ranked set sampling in case of the exponential distribution. Communications in Statistics: Theory and Methods **39**, 1855-1868.

[5] Vasermanis, E.K., Nechval, N.A., Rozevskis, U., and Rosite, K. (2003). Minimum variance unbiased estimation of availability. Transp. Telecommun **4**, 108-115.

Table 1: Estimated REs for different configurations

			$(p,q)=(10,10)$				$(p,q)=(10,20)$			
ρ	λ	m	RE1	RE2	RE3	RE4	RE1	RE2	RE3	RE4
0.15	1	2	1.37	1.45	1.45	1.44	1.34	1.40	1.40	1.39
		5	2.58	3.03	3.06	2.96	2.24	2.60	2.61	2.54
	0.75	2	1.26	1.25	1.24	1.21	1.18	1.15	1.15	1.11
		5	1.65	1.54	1.56	1.35	1.43	1.29	1.30	1.12
	0.5	2	1.18	1.12	1.13	1.07	1.06	1.01	1.01	0.97
		5	1.20	1.07	1.08	0.94	1.09	0.95	0.97	0.82
1	1	2	1.34	1.40	1.40	1.39	1.32	1.38	1.38	1.37
		5	2.26	2.61	2.63	2.55	2.17	2.47	2.49	2.41
	0.75	2	1.19	1.18	1.18	1.14	1.19	1.16	1.17	1.13
		5	1.49	1.42	1.43	1.27	1.40	1.30	1.31	1.16
	0.5	2	1.13	1.09	1.09	1.05	1.05	1.01	1.01	0.98
		5	1.15	1.06	1.07	0.95	1.12	1.00	1.02	0.89
6	1	2	1.49	1.55	1.56	1.53	1.40	1.46	1.47	1.45
		5	2.56	2.99	3.02	2.91	2.49	2.84	2.87	2.75
	0.75	2	1.20	1.18	1.18	1.14	1.25	1.22	1.23	1.18
		5	1.56	1.48	1.50	1.30	1.49	1.39	1.41	1.23
	0.5	2	1.14	1.09	1.09	1.04	1.07	1.03	1.03	0.99
		5	1.17	1.06	1.07	0.94	1.19	1.05	1.07	0.91

Table 2: Ranked set sample data for the percentage area covered on the surface of the leaves of apple trees

Group	Cycle	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
Control	1	0.003	0.028	0.244	0.057	0.143
	2	0.039	0.119	0.126	0.105	0.565
	3	0.034	0.118	0.130	0.218	0.296
	4	0.051	0.104	0.193	0.210	0.150
	5	0.032	0.141	0.130	0.250	0.229
	6	0.069	0.070	0.260	0.225	0.285
	7	0.100	0.091	0.244	0.130	0.347
	8	0.012	0.096	0.069	0.373	0.133
	9	0.046	0.117	0.126	0.223	0.273
	10	0.028	0.083	0.108	0.212	0.261
Treatment	1	0.036	0.137	0.183	0.270	0.487
	2	0.250	0.181	0.290	0.328	0.715
	3	0.089	0.032	0.269	0.419	0.315
	4	0.180	0.111	0.130	0.194	0.742
	5	0.100	0.009	0.184	0.277	0.122
	6	0.042	0.089	0.199	0.269	0.395
	7	0.044	0.083	0.227	0.177	0.742
	8	0.044	0.171	0.067	0.192	0.336
	9	0.009	0.017	0.217	0.438	0.544
	10	0.071	0.132	0.310	0.343	0.379

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Semiparametric regression for mean past lifetime

Mansourvar, Z. [1] and Asadi, M. [2]

[1,2] 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 the presence of covariates, regression models are needed to study the association between the mean past lifetime function and potential regression covariates. In this paper, we propose the proportional mean past life model for fitting survival data under left censoring. To estimate the model parameters, martingale estimating equations are developed, and the asymptotic properties of the resulting estimators are established.

Keywords: Counting process, Martingale estimating equation, Mean past lifetime, Proportional model, Reversed hazard rate.

1 Introduction

In reliability studies and survival analysis, several approaches have been considered in the literature to study the lifetime and aging properties of a system or any other living organism. The basic tools for studying the characteristics of lifetime distribution specially under right censored data are the hazard rate and the mean residual lifetime. There are many occasions in survival studies, where the lifetime data are left censored. For example, when we are following persons until they become HIV positive, we may record a failure when a subject first tests positive for the virus. However, we may not know exactly the time of first exposure to the virus, and therefore do not

[1] z.mansourvar@sci.ui.ac.ir
[2] m.asadi@sci.ui.ac.ir

know exactly when the failure occurred. Thus, the survival time is censored on the left side. The reversed hazard rate (RHR) and the mean past lifetime (MPL) facilitate the analysis of such left censored data, which are defined, respectively, by

$$r(t) = \lim_{\Delta t \to 0} \frac{P(t - \Delta t < T < \Delta t | T \leq t)}{\Delta t},$$

$$k(t) = E(t - T | T \leq t).$$

The RHR specifies the instantaneous rate of failure of a subject at time t given that it failed before time 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. In this example, the random variable of interest is $T_t = t - T | T < t$, and the MPL function is its expected value.

The MPL has recently been considered by some authors. [1] gave the necessary conditions for a function to be MPL, where the MPL is called the mean waiting time in his paper. The latter author has also shown that the reversed hazard ordering implies the MPL ordering. [2] defined some new classes of distributions based on the MPL, and obtained some results on the MPL ordering, where they called the MPL the "reversed mean residual". [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. [9] studied some implications of stochastic orders and aging notions for the proportional MPL model and for its extended mixture model.

In this study, to assess the association between the MPL and potential regression covariates, we propose the proportional MPL model that takes the form

$$k(t|Z) = k_0(t) \exp(Z^T \beta), \tag{1.1}$$

where $k(t|Z) = E(t - T | T \leq t, Z)$ and $k_0(t)$ is an unspecified baseline MPL, and β is a $p \times 1$ vector of regression parameters. In the presence of left censoring, we apply martingale estimating equations to estimate $k_0(t)$ and β. The rest of the paper is organized as follows. Section 2 is devoted to semiparametric inference procedures for estimating the non-parametric component $k_0(t)$ and parametric component β in model (1.1). Section 3 gives the asymptotic properties of the proposed estimators. Regularity conditions and theoretical proofs are collected in the Appendix.

2 Inference procedures

Suppose that the lifetime random variable T is left censored by the random variable C. We assume that T is independent of C given the covariate Z. Under left censoring we observe random vector (X, Δ, Z) where $X = \max(T, C)$ and $\Delta = I(T \geq C)$ with $I(\cdot)$ as the usual indicator function. Let (X_i, Δ_i, Z_i) be independent and identically distributed copies of $(X, \Delta, Z), i = 1, 2, \ldots, n$.

We now employ the counting process approach for the estimation of the parameters of the proposed model. Let $N_i(t) = I(X_i \geq t, \Delta_i = 1)$, $Y_i(t) = I(X_i \leq t)$. Define the sigma field $\mathcal{F}_t = \sigma\{N_i(u), Y_i(u), Z_i : \quad 0 \leq t \leq u < \tau;\ i = 1, \ldots, n\}$. 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

$$E\{dN_i(t)|\mathcal{F}_{t^+}; \beta_*, k_*(.)\} = Y_i(t)dR_i(t; \beta_*, m_*),$$

where the true values of β and $m_0(t)$ are defined by β_* and $m_*(t)$, respectively, and $R(t) = \int_t^\tau r(u)du$ and $\tau = \sup(t; F(t) < 1)$. Let

$$dM_i(t; \beta_*, m_*) = dN_i(t) - Y_i(t)dR_i(t|Z_i; \beta_*, k_*), \quad (i = 1, \ldots, n),$$

where $\{M_i(t; \beta_*, m_*), t \geq 0\}$, is a martingale with respect to $\mathcal{F}_t$. Therefore without assuming any particular form for $k_0(t)$, it is natural to estimate $k_*(t)$ and β_* from estimating equations parallel to the partial score equations

$$\begin{aligned} \sum_{i=1}^{n} \{dN_i(t) - Y_i(t)dR_i(t; \beta, k_0)\} = 0, \quad (0 \leq t < \tau), \\ \sum_{i=1}^{n} \int_0^\tau Z_i \{dN_i(t) - Y_i(t)dR_i(t; \beta, k_0)\} = 0. \end{aligned} \tag{2.1}$$

It is well known that the RHR of T given Z is [3, sect. 2]

$$r(t|Z) = \frac{1 - k'(t|Z)}{k(t|Z)}. \tag{2.2}$$

Then under model (1.1), one obtains that

$$k_0(t)dR_i(t|Z_i; \beta, k_0) = \exp(-Z_i^T\beta)dt - dk_0(t), \quad (i = 1, \ldots, n).$$

Therefore, by analogy with (2.1), the following estimating equations can be used for estimating $k_*(t)$ and β_* in model (1.1), respectively

$$\sum_{i=1}^{n} \left[k_0(t)dN_i(t) - Y_i(t)\left\{\exp(-Z_i^T\beta)dt - dk_0(t)\right\}\right] = 0, \tag{2.3}$$

$$\sum_{i=1}^{n} \int_0^\tau Z_i \left[k_0(t)dN_i(t) - Y_i(t)\left\{\exp(-Z_i^T\beta)dt - dk_0(t)\right\}\right] = 0. \tag{2.4}$$

In view of equation (2.3), it can be obtained that

$$dk_0(t) = \frac{\sum_{i=1}^n J(t)\left\{Y_i(t)\exp(-Z_i^T\beta)dt - k_0(t)dN_i(t)\right\}}{Y_\cdot(t)}, \tag{2.5}$$

where $Y_\cdot(t) = \sum_{i=1}^n Y_i(t)$ and $J(t) = I(Y_\cdot(t) > 0)$. Since $k_0(0) = 0$ in model (1.1), we suggest the following recursive estimating equation for estimation of $k_*(t)$

$$\hat{k}_0(t;\beta) = \int_0^t \frac{J(s)\sum_{i=1}^n \left\{Y_i(s)\exp(-Z_i^T\beta)ds - \hat{k}_0(s^-)dN_i(s)\right\}}{Y_\cdot(s)}, \tag{2.6}$$

where $\hat{k}_0(s^-)$ is the left hand limit of $\hat{k}_0(t)$ at s. The equation displayed in (2.6) has a recursive structure, but admits an explicit solution. The Volterra equation [4, pp. 90-91] can easily be adapted to obtain the solution for equation (2.6) as

$$\hat{k}_0(t;\beta) = Q(t)^{-1}\int_0^t Q(s^-)\frac{J(s)\sum_{i=1}^n Y_i(s)\exp(-Z_i^T\beta)}{Y_\cdot(s)}ds, \tag{2.7}$$

where

$$Q(t) = \prod_{[t,\tau]}\left\{1 - \frac{J(s)dN_\cdot(s)}{Y_\cdot(s)}\right\},$$

with $\prod_{[t,\tau]}$ denoting a product-integral, see [4, Sect. II.6] and $dN_\cdot(t) = \sum_{i=1}^n dN_i(t)$.

To estimate β_*, we replace $k_0(t)$ with $\hat{k}_0(t;\beta)$ in (2.4). Then in view of (2.5), the score equation for β_* can be simplified as

$$U\left\{\tau,\beta;\hat{k}_0(t;\beta)\right\} = \sum_{i=1}^n \int_0^\tau \left\{Z_i - \bar{Z}(t)\right\}\left\{\hat{k}_0(t;\beta)dN_i(t) - Y_i(t)\exp(-Z_i^T\beta)dt\right\}, \tag{2.8}$$

where $\bar{Z}(t) = J(t)\sum_{i=1}^n Y_i(t)Z_i/\sum_{i=1}^n Y_i(t)$. Let $\hat{\beta}$ denote the solution to $U\left\{\tau,\beta;\hat{k}_0(t;\beta)\right\} = 0$. The corresponding estimator of $k_*(t)$ is given by $\hat{k}_0(t) = \hat{k}_0(t;\hat{\beta})$.

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 $\hat{k}_0(t)$. The results are summarized in the following theorem with the proof given in the Appendix.

Theorem 3.1. *Under the regularity conditions stated in the Appendix, $\hat{\beta}$ exists and is unique. Moreover, $\hat{\beta}$ is strongly consistent for β_*, and $\hat{k}_0(t) \to k_*(t)$ almost surely uniformly in $t \in [0,\tau]$ as $n \to \infty$.*

Theorem 3.2. *Under the regularity conditions stated in the Appendix, we have*

(i) $n^{-1/2}U\{\tau, \beta_*; \hat{k}_0(t;\beta_*)\}$ *converges to a normal distribution with mean zero and a variance-covariance matrix that can be consistently estimated by* $\hat{\Sigma}$*, where* $\hat{\Sigma} = n^{-1}\sum_{i=1}^{n}\hat{\xi}_i^{\otimes 2}$*,*

$$\hat{\xi}_i = \int_0^\tau \{Z_i - \bar{Z}(t) - \tilde{Z}(t)\}\,\hat{k}_0(t)d\hat{M}_i(t),$$

with

$$\tilde{Z}(t) = \frac{J(t)Q(t^-)}{Y_\cdot(t)}\int_t^\tau Q(s)^{-1}\sum_{i=1}^{n}\{Z_i - \bar{Z}(s)\}dN_i(s),$$

$d\hat{M}_i(t)$ *is the corresponding estimator* $dM_i(t)$ *with all unknown parameters replace by their estimates and* $\upsilon^{\otimes 2} = \upsilon\upsilon^T$ *for a vector* υ*;*

(ii) $n^{1/2}(\hat{\beta} - \beta_*)$ *is asymptotically normal with mean zero and a variance-covariance matrix that can be consistently estimated by* $\hat{D}(\tau,\hat{\beta})^{-1}\hat{\Sigma}\,\hat{D}(\tau,\hat{\beta})^{-1}$*, where*

$$\hat{D}(\tau,\hat{\beta}) = -n^{-1}\sum_{i=1}^{n}\int_0^\tau \left\{Z_i - \bar{Z}(t) - \tilde{Z}(t)\right\}Y_i(t)Z_i^T\exp(-Z_i^T\hat{\beta})dt.$$

Theorem 3.3. *Under the regularity conditions stated in the Appendix,* $n^{1/2}\{\hat{k}_0(t) - k_*(t)\}$ $(0 \le t \le \tau)$ *converges weakly to a zero-mean Gaussian process whose covariance function at* (s,t) *can be consistently estimated by* $\hat{\Gamma}(s,t) = n^{-1}\sum_{i=1}^{n}\hat{\phi}_i(s)\hat{\phi}_i(t)$*, where*

$$\begin{aligned}\hat{\phi}_i(t) = &- Q(t)^{-1}\int_0^t Q(s^-)\frac{J(s)\hat{k}_0(s)}{n^{-1}Y_\cdot(s)}d\hat{M}_i(s)\\ &- \hat{B}(t,\hat{\beta})\hat{D}(\tau,\hat{\beta})^{-1}\int_0^\tau\{Z_i - \bar{Z}(t) - \tilde{Z}(t)\}\hat{k}_0(t)d\hat{M}_i(t),\end{aligned}$$

and

$$\hat{B}(t,\hat{\beta}) = Q(t)^{-1}\int_0^t Q(s^-)\frac{J(s)\sum_{i=1}^{n}Y_i(s)Z_i^T\exp(-Z_i^T\hat{\beta})}{Y_\cdot(s)}ds.$$

References

[1] Finkelstein, M. S. (2002), On the reversed hazard rate, *Reliability Engineering and System Safety*, **78**(1), 71-75.

[2] 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.

[3] Asadi, M. and Berred, A. (2012), Properties and estimation of the mean past lifetime, *Statistics*, **46**(3), 405-417.

[4] Andersen, P. K., Borgan, O., Gill, R. D. and Keiding, N. (2012), *Statistical models based on counting processes*, Springer Science and Business Media.

[5] Lin, D. Y. (2000), On fitting Cox's proportional hazards models to survey data, *Biometrika*, **87**(1), 37-47.

[6] Pollard, D. (1990), *Empirical processes: theory and applications*, Institute of Mathematical Statistics and the American Statistical Association.

[7] Bilias, 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.

[8] Rudin, W. (1964), *Principles of mathematical analysis*, Vol. 3, New York: McGraw-Hill.

[9] Rezaei, M. (2016), On proportional mean past lifetimes model, *Communications in Statistics-Theory and Methods*, **45**(14), 4035-4047.

Appendix

Let $\bar{Z}(t)$ and $\tilde{Z}(t)$ denote the quantities defined in the text. Also let $\mu(t)$ and $\tilde{\mu}(t)$ denote their limits, respectively. In order to study the asymptotic properties of the proposed estimators, we need the following regularity conditions:

(C1) $P(C \geq \tau) > 0$, and $N(\tau)$ in bounded almost surely.

(C2) The covariate Z is bounded.

(C3) $k_*(t)$ is right continuous with left-hand limits, and has bounded total variation on $[0, \tau]$.

(C4) The matrix $D(\tau, \beta) = -E\left[\int_0^\tau \{Z_i - \mu(t) - \tilde{\mu}(t)\} Y_i(t) Z_i^T \exp(-Z_i^T \beta) dt\right]$ is non-singular.

Proof of Theorem 3.1

It follows from (2.6) that

$$\hat{k}_0(t;\beta) - k_0(t) = -\int_0^t \frac{J(s)k_0(s)dM_\cdot(s)}{Y_\cdot(s)} - \int_0^t \frac{J(s)\left\{\hat{k}_0(s^-;\beta) - k_0(s)\right\} dN_\cdot(s)}{Y_\cdot(s)} - \int_0^t \{1 - J(s)\}\, dk_0(s), \tag{A.1}$$

where $dM_{\cdot}(t) = \sum_{i=1}^{n} dM_i(t)$. Since the latter term in (A.1) converges to zero uniformly in probability, using the Volterra equation [4, pp. 90-91], the foregoing equation has the solution given by

$$\hat{k}_0(t;\beta) - k_0(t) = -Q(t)^{-1} \sum_{i=1}^{n} \int_0^t Q(s^-) \frac{J(s)k_0(s)}{Y_{\cdot}(s)} dM_i(s), \tag{A.2}$$

where $Q(t) = \prod_{[t,\tau]} [1 - \{J(s)dN_{\cdot}(s)/Y_{\cdot}(s)\}]$. Define $A(t) = \int_t^{\tau} \{J(s)/Y_{\cdot}(s)\} dN_{\cdot}(s)$, we then have

$$A(t;\beta) = \int_t^{\tau} \frac{J(s)}{Y_{\cdot}(s)} dM_{\cdot}(s) + \int_t^{\tau} \frac{J(s)}{k_0(s)} \left\{ \frac{\sum_{i=1}^{n} Y_i(s) \exp(-Z_i^T\beta) ds}{Y_{\cdot}(s)} + dk_0(s) \right\}. \tag{A.3}$$

By Lenglart's inequality [4, p. 86], the first term in (A.3) converges uniformly in probability to zero. The second term, by the weak law of large numbers converges uniformly in probability to $a(t;\beta) = \int_t^{\tau} \{\lambda(s;\beta)ds + dk_0(s)\}/k_0(s)$ where $\lambda(t;\beta) = E\{Y_i(t)\exp(-Z_i^T\beta)\}/y(t)$ in which a function $y(t)$ is such that $\sup_{s\in[t,\tau]} |n^{-1}Y_{\cdot}(s) - y(s)| \xrightarrow{P} 0$.
Therefore, we have $Q(t;\beta) = \prod_{[t,\tau]}\{1 - dA(s;\beta)\}$ with $A(t;\beta) = a(t;\beta) + o_p(1)$. Using arguments along the lines of [4, pp. 263-264] for the Kaplan-Meier estimate, it can be demonstrated that $Q(t;\beta)$ converges uniformly in probability to $q(t;\beta) = \exp\{a(t;\beta)\}$. Now, again from the uniform weak law of large numbers, $J(t)k_0(t)/\{n^{-1}Y_{\cdot}(t)\}$ converges uniformly in probability to $k_0(t)/y(t)$. Here and subsequently, the $o_p(1)$ is uniform in t.
It remains to show that

$$\sum_{i=1}^{n} \int_0^t \{Q(s^-;\beta) - q(s;\beta)\} d\tilde{M}_i(s) \to 0,$$

uniformly in probability, where $d\tilde{M}_i(t) = J(t)k_0(t)dM_i(t)/n^{-1}Y_{\cdot}(t)$. Because $Q(t;\beta) = \exp\{A(t;\beta)\}$ is monotone, this follows from the Lemma in [5, p. 45]. Thus it can be obtained that

$$\hat{k}_0(t;\beta_*) - k_*(t) = -n^{-1}q(t;\beta_*)^{-1} \int_0^t q(s;\beta_*) \frac{k_*(s)}{y(s)} dM_{\cdot}(s) + o_p(1). \tag{A.4}$$

From Lenglart's inequality [4, p. 86], the above term converges uniformly in probability to zero. Hence (A.4) implies that $\hat{k}_0(t;\beta_*)$ converges almost surely to $k_*(t)$ in $t \in [0,\tau]$.

Denote the minus of the first derivative of $U\{\tau,\beta;\hat{k}_0(t;\beta)\}$ with respect to β by $I(\tau,\beta)$, where

$$I(\tau,\beta) = -\sum_{i=1}^{n} \int_0^{\tau} \{Z_i - \bar{Z}(t)\} \left\{ \frac{\partial \hat{k}_0(t;\beta)}{\partial\beta} dN_i(t) + Y_i(t)Z_i^T \exp(-Z_i^T\beta)dt \right\}.$$

Taking the derivative of representation (2.7) with respect to β and substituting that into the foregoing term and then interchanging the order of integrals, we obtain

$$I(\tau,\beta) = -\sum_{i=1}^{n} \int_0^{\tau} \left\{ Z_i - \bar{Z}(t) - \tilde{Z}(t) \right\} Y_i(t)Z_i^T \exp(-Z_i^T\beta)dt.$$

It can be proved that $\bar{Z}(t)$ and $\tilde{Z}(t)$ are of bounded variation and thus can be expressed as the difference of two increasing functions. Since the process $I(\tau,\beta)$ can be written as sums and products of monotone functions in t and all components of β, therefore it is manageable ([6, p. 38]; [7, Lemma A.1]). Using

the uniform strong law of large numbers [6, p. 41], it follows that $n^{-1}I(\tau,\beta)$ converges almost surely to a nonrandom function $D(\tau,\beta)$ uniformly in β, where

$$D(\tau,\beta) = -E\left[\int_0^\tau \{Z_i - \mu(t) - \tilde{\mu}(t)\}Y_i(t)Z_i^T \exp(-Z_i^T\beta)dt\right],$$

with $\mu(t)$ and $\tilde{\mu}(t)$ are the limit in probability of $\bar{Z}(t)$ and $\tilde{Z}(t)$, respectively.

It can also be checked that $n^{-1}U(\tau,\beta_*) \to 0$ almost surely, and $D(\tau,\beta)$ is non-singular by condition (C4). Therefore, the uniform convergence of $n^{-1}I(\tau,\beta)$ and the continuity of $D(\tau,\beta)$ imply that for all large n, there exists a small neighborhood of β_* in which $n^{-1}I(\tau,\beta)$ is non-singular. Hence it follows from the inverse function theorem [8, p. 221] that within a small neighborhood of β_*, there exists a unique solution $\hat{\beta}$ to $U(\tau,\beta) = 0$ for all large n. Since this neighborhood of β_* can be arbitrarily small, the preceding proof also implies that $\hat{\beta}$ is strongly consistent. It then follows from the uniform convergence of $\hat{k}_0(t;\beta_*)$ to $k_0(t;\beta_*)$ that $\hat{k}_0(t) \equiv \hat{k}_0(t;\hat{\beta}) \to k_0(t;\beta_*) \equiv k_*(t)$ almost surely uniformly in $t \in [0,\tau]$.

Proof of Theorem 3.2

(i) Consider a decomposition of $n^{-1/2}U\{\tau,\beta_*;\hat{k}_0(t;\beta_*)\}$ of the form

$$\begin{aligned} n^{-1/2}U\{\tau,\beta_*;\hat{k}_0(t;\beta_*)\} &= n^{-1/2}\sum_{i=1}^n \int_0^\tau \left\{Z_i - \bar{Z}(t)\right\}\left\{\hat{k}_0(t;\beta_*) - k_*(t)\right\} dN_i(t) \\ &\quad + n^{-1/2}\sum_{i=1}^n \int_0^\tau \left\{Z_i - \bar{Z}(t)\right\} k_*(t) dM_i(t). \end{aligned}$$

By the representation of $\hat{k}_0(t;\beta_*) - k_*(t)$ in (A.2) and interchanging the order of integration, the foregoing expression can be written as

$$n^{-1/2}U\{\tau,\beta_*;\hat{k}_0(t;\beta_*)\} = n^{-1/2}\sum_{i=1}^n \int_0^\tau \{Z_i - \bar{Z}(t) - \tilde{Z}(t)\}k_*(t)dM_i(t).$$

Since $\bar{Z}(t)$ and $\tilde{Z}(t)$ are of bounded variation, thus can be written as the difference of two increasing functions. Therefore, by the uniform strong law of large numbers and using the Lemma [5, p. 45], $n^{-1/2}U\{\tau,\beta_*;\hat{k}_0(t;\beta_*)\}$ can be decomposed as a sum of independent and identically distributed terms

$$n^{-1/2}U\{\tau,\beta_*;\hat{k}_0(t;\beta_*)\} = n^{-1/2}\sum_{i=1}^n \xi_i + o_p(1), \tag{A.5}$$

where

$$\xi_i = \int_0^\tau \{Z_i - \mu(t) - \tilde{\mu}(t)\}k_*(t)dM_i(t).$$

Utilizing the multivariate central limit theorem, $n^{-1/2}U\{\tau,\beta_*;\hat{k}_0(t;\beta_*)\}$ converges in distribution to zero-mean normal distribution whose variance-covariance matrix $\Sigma = E\{\xi_i^{\otimes 2}\}$ can be consistently estimated by $\hat{\Sigma} = n^{-1}\sum_{i=1}^n \hat{\xi}_i^{\otimes 2}$ as defined in Theorem 2(i).

(ii) A Taylor series expansion of the score function (2.8) around $\hat{\beta}$ gives

$$n^{1/2}(\hat{\beta} - \beta_*) = \{n^{-1}I(\tau,\beta^{**})\}^{-1}n^{-1/2}U\{\tau,\beta_*;\hat{k}_0(t;\beta_*)\},$$

where β^{**} is on the line segment between β_* and $\hat{\beta}$. To prove asymptotic normality of $n^{1/2}(\hat{\beta}-\beta_*)$ it suffices to prove weak convergence of $n^{-1/2}U\{\tau,\beta_*;\hat{k}_0(t;\beta_*)\}$ to a Gaussian process and to prove convergence in probability of $n^{-1}I(\tau,\beta^{**})$ to a non-singular matrix. From the uniform convergence of $n^{-1}I(\tau,\beta)$ to $D(\tau,\beta)$ along with the consistency of $\hat{\beta}$ and representation (A.5), asymptotic approximation for $n^{1/2}(\hat{\beta}-\beta_*)$ can be displayed by

$$n^{1/2}(\hat{\beta}-\beta_*)=n^{-1/2}D(\tau,\beta_*)^{-1}\sum_{i=1}^{n}\xi_i+o_p(1). \tag{A.6}$$

Thus, it follows that $n^{1/2}(\hat{\beta}-\beta_*)$ is asymptotically normal with mean zero and covariance matrix $D(\tau,\beta_*)^{-1}\Sigma\, D(\tau,\beta_*)^{-1}$, which can be consistently estimated by $\hat{D}(\tau,\hat{\beta})^{-1}\hat{\Sigma}\,\hat{D}(\tau,\hat{\beta})^{-1}$ as defined in Theorem 2(ii).

Proof of Theorem 3.3

To show the weak convergence of $n^{1/2}\{\hat{k}_0(t)-k_*(t)\}$, we first note that

$$n^{1/2}\{\hat{k}_0(t)-k_*(t)\}=n^{1/2}\{\hat{k}_0(t;\beta_*)-k_*(t)\}+n^{1/2}\{\hat{k}_0(t;\hat{\beta})-\hat{k}_0(t;\beta_*)\}.$$

It follows from (A.4) that

$$n^{1/2}\left\{\hat{k}_0(t;\beta_*)-k_*(t)\right\}=n^{-1/2}\sum_{i=1}^{n}\zeta_i(t)+o_p(1),$$

where

$$\zeta_i(t)=-q(t;\beta_*)^{-1}\int_0^t q(s;\beta_*)\frac{k_*(s)}{y(s)}dM_i(s),\quad (i=1,\ldots,n).$$

Taking the Taylor expansion of $\hat{k}_0(t,\hat{\beta})$, together with the consistency of $\hat{\beta}$ and the uniform strong law of large numbers, we have

$$n^{1/2}\{\hat{k}_0(t,\hat{\beta})-\hat{k}_0(t,\beta_*)\}=-B(t;\beta_*)\,n^{1/2}(\hat{\beta}-\beta_*)+o_p(1),$$

where

$$B(t;\beta)=q(t;\beta)^{-1}\int_0^t q(s;\beta)\frac{E\left\{Y_i(s)Z_i^T\exp(-Z_i^T\beta)\right\}}{y(s)}ds,$$

denotes the limit in probability of $-\partial\hat{k}_0(t;\beta)/\partial\beta$. Therefore, it follows from (A.6) that uniformly in $t\in[0,\tau]$,

$$n^{1/2}\{\hat{k}_0(t)-k_*(t)\}=n^{-1/2}\sum_{i=1}^{n}\phi_i(t)+o_p(1), \tag{A.7}$$

where

$$\phi_i(t)=\zeta_i(t)-B(t;\beta_*)D(\tau,\beta_*)^{-1}\xi_i.$$

Since $\phi_i(t)$, $i=1,\ldots,n$, are independent zero-mean random variables for each t, the multivariate central limit theorem implies that $n^{1/2}\{\hat{k}_0(t)-k_*(t)\}$ converges in finite-dimensional distribution to a zero-mean Gaussian process for $0\le t\le\tau$. Because any function of bounded variation can be expressed as the difference

of two increasing functions, the processes $\{\zeta_i(t);\ i = 1, \ldots, n\}$ can be written as sums or products of monotone functions of t and are thus manageable. It then follows from the functional central limit theorem [6, p. 53] that the first term on the right-hand side of (A.7) is tight. The second term is tight because $n^{-1/2} \sum_{i=1}^{n} \xi_i$ converges in distribution and $B(t, \beta_*)$ is a deterministic function. Thus $n^{1/2} \left\{ \hat{k}_0(t) - k_*(t) \right\}$ is tight and converges weakly to a zero-mean Gaussian process whose covariance function $\Gamma(s,t) = E\{\phi_i(s)\phi_i(t)\}$ can be consistently estimated by $\hat{\Gamma}(s,t)$ defined in Theorem 3.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

On a dynamic L_2 distance with applications

Mehrali, Y. [1] and Asadi, M. [2]

[1,2] Department of Statistics, University of Isfahan

Abstract

In this paper, we propose a dynamic form of L_2 distance between the probability distributions of two nonnegative continuous random variables. We investigate some properties of the proposed measure and explore its relation with some well known stochastic and aging concepts. Empirical results are also presented leading to an estimator for the proposed criterion. The results are then employed to compare two probability distributions based on real data sets.

Keywords: Aging, Information measures, Residual life, Stochastic ordering.

1 Introduction

The energy distance is a measure of discrimination between two distributions that introduced by Székely [9, 10]. If X and Y be independent random vectors on $\mathbb{R}^d$ with cumulative distribution functions ($c.d.f.$'s) $F(x)$ and $G(x)$, respectively, then the squared energy distance (ED^2) was defined as

$$ED^2(F,G) = 2E\|X-Y\| - E\|X-X'\| - E\|Y-Y'\|,$$

where X' is $i.i.d.$ with X, Y' is $i.i.d.$ with Y and $\|\cdot\|$ denotes the Euclidean norm. The energy distance (ED) between distributions F and G is defined as the square root of $ED^2(F,G)$. Székely

[1]yasermehrali@gmail.com
[2]m.asadi@sci.ui.ac.ir

[10] proved that if $d = 1$, then $ED^2(F,G)$ is twice Cramér's distance (CD) that introduced by Cramér' [1] (also known as L_2 distance), i.e.,

$$ED^2(F,G) = 2CD(F,G) = 2\int_{-\infty}^{\infty} (F(x) - G(x))^2 dx.$$

For more informations about background and applications of energy distance, see [11], [7], and references cited therein. Rizzo and Székely [8] also provided the *energy* package, for the statistical software *R*.

In the present paper, we consider a dynamic version of Cramér's distance, and call it dynamic Cramér's distance (DCD), as well as dynamic L_2 distance. The rest of this paper is organized as follows: In Section 2, we present a dynamic form of Cramér's distance for nonnegative continuous random variables. In Section 3, we obtain some properties of the introduced measure. Finally in Section 4, we provide some empirical results.

2 A dynamic form of L_2 distance

Hereafter, let X and Y be independent nonnegative continuous random variables with *c.d.f.*'s $F(x)$ and $G(x)$, survival functions (*s.f.*'s) $\bar{F}(x)$ and $\bar{G}(x)$, and mean residual life (MRL) functions $m_F(t)$ and $m_G(t)$, respectively. Then $CD(F,G)$ can be defined as

$$CD(F,G) = \int_0^{\infty} \left(\bar{F}(x) - \bar{G}(x)\right)^2 dx.$$

Now, we consider the dynamic version of CD and call it dynamic Cramér's distance (DCD) which is defined as

$$\begin{aligned} DCD(F,G;t) &= \int_t^{\infty} \left(\bar{F}_t(x) - \bar{G}_t(x)\right)^2 dx \\ &= m_F^{(2)}(t) + m_G^{(2)}(t) - 2m_{FG}(t), \end{aligned} \tag{A.1}$$

where $\bar{F}_t(x) = \frac{\bar{F}(x)}{\bar{F}(t)}$, $m_F^{(2)}(t)$ is MRL function of minimum of two components from F and $m_{FG}(t)$ is MRL function of minimum of two components from F and G. Now, we study monotonicity of DCD.

Theorem 2.1. *$DCD(F,G;t)$ is non decreasing (non increasing) function of t, iff*

$$DCD(F,G;t) \geq (\leq)\,(1-2w_t)\left(m_F^{(2)}(t) - m_G^{(2)}(t)\right), \tag{A.2}$$

where $w_t = \lambda_F(t) / (\lambda_F(t) + \lambda_G(t))$.

Theorem 2.2. *If $X \geq_{hr} Y$ $(X \leq_{hr} Y)$, X is IFR and Y is DFR, then $DCD(F,G;t)$ is non increasing (non decreasing) function of t.*

3 Some properties of DCD

In this section, we obtain some properties of DCD. The following theorem gives a lower bound for $DCD(F,G;t)$.

Theorem 3.1.

$$DCD(F,G;t) \geq m_F^{(2)}(t) + m_G^{(2)}(t) - 2\max\left(m_F(t), m_G(t)\right).$$

We have the following upper bound for $DCD(F,G;t)$ in general case.

Theorem 3.2.

$$DCD(F,G;t) \leq m_F(t) + m_G(t). \tag{A.1}$$

The following theorem gives an upper bound for $DCD(F,G;t)$ which is sharper than that in (A.1).

Theorem 3.3.

$$DCD(F,G;t) \leq \int_t^\infty \left|\bar{F}_t(x) - \bar{G}_t(x)\right| dx.$$

In the following theorem, we compare $DCD(F,G;t)$ and $CD(F,G)$.

Theorem 3.4. *If $X \geq_{hr} Y$ ($X \leq_{hr} Y$), X is NWU and Y is NBU, then*

$$DCD(F,G;t) \geq (\leq) CD(F,G).$$

The next theorem compares distances of three random variables.

Theorem 3.5. *Let X, Y and Z be three nonnegative random variables with p.d.f.'s F, G and H, respectively.*

(a) If $X \geq_{hr} Z \geq_{hr} Y$ or $X \leq_{hr} Z \leq_{hr} Y$, then

$$DCD(F,G;t) \geq \max\left\{DCD(F,H;t), DCD(H,G;t)\right\}.$$

(b) If $X \geq_{hr} Y$ or $X \leq_{hr} Y$ and Z be mixture of X and Y, then (a) holds.

Finally, we consider the behavior of $DCD(F,G;t)$ under transformations.

Theorem 3.6. *Let ϕ be an increasing function on $(0,\infty)$ with $a \leq \phi' \leq b$, $a, b > 0$, where ϕ' is derivation of ϕ, and Let $DCD(X,Y;t)$ be equivalent with $DCD(F,G;t)$. Then*

$$\begin{aligned} a\ DCD\left(X,Y;\phi^{-1}(t)\right) &\leq DCD\left(\phi(X),\phi(Y);t\right) \\ &\leq b\ DCD\left(X,Y;\phi^{-1}(t)\right). \end{aligned}$$

and

$$DCD(bX,bY;t) = b\ DCD\left(X,Y;\frac{t}{b}\right),\ b>0.$$

4 Empirical results

Let $X_1, \cdots, X_n$ and $Y_1, \cdots, Y_m$ be two independent random samples from F and G, respectively. Based on empirical $s.f.$, we can define the empirical estimation of $DCD(F, G; t)$ as distance of F_n and G_m.

Definition 1. *Under above considerations, the empirical dynamic Cramér's distance of F_n and G_m is defined as*

$$\begin{aligned} DCD(F_n, G_m; t) &= \int_t^\infty \left(\bar{F}_{nt}(x) - \bar{G}_{mt}(x)\right)^2 dx \\ &= \int_t^\infty \left(\frac{\bar{F}_n(x)}{\bar{F}_n(t)} - \frac{\bar{G}_m(x)}{\bar{G}_m(t)}\right)^2 dx. \end{aligned} \tag{A.1}$$

and the empirical Cramér's distance of F_n and G_m is defined as

$$CD(F_n, G_m) = \int_0^\infty \left(\bar{F}_n(x) - \bar{G}_m(x)\right)^2 dx, \tag{A.2}$$

which is related to the energy statistic that defined by Rizzo [6].

In following, we will find simple forms of (A.1) and (A.2).

Theorem 4.1. *Let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ and $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(m)}$ be ordered sample of F and G, respectively, and $U_{(1)} \le U_{(2)} \le \cdots \le U_{(n+m)}$ be ordered pooled sample. Let $M_i = \sum_{j=1}^m I\left(Y_j \le X_{(i)}\right)$, $i = 1, 2, \cdots, n$ and $N_j = \sum_{i=1}^n I\left(X_i \le Y_{(j)}\right)$, $j = 1, 2, \cdots, m$. If $X_{(k)} \le t < X_{(k+1)}$ and $Y_{(s)} \le t < Y_{(s+1)}$, then $DCD(F_n, G_m; t)$ can be rewritten as*

$$\begin{aligned} DCD(F_n, G_m; t) = &\frac{1}{(n-k)^2}\left\{\sum_{i=k+1}^n X_{(i)} - 2\sum_{i=k+1}^n iX_{(i)}\right\} \\ &+ \frac{1}{(m-s)^2}\left\{\sum_{j=s+1}^m Y_{(j)} - 2\sum_{j=s+1}^m jY_{(j)}\right\} \\ &+ 2\left(\frac{n}{n-k} - \frac{m}{m-s}\right) \\ &\times \left\{\frac{1}{n-k}\sum_{i=k+1}^n X_{(i)} - \frac{1}{m-s}\sum_{j=s+1}^m Y_{(j)}\right\} \\ &+ \frac{2}{(n-k)(m-s)}\left\{\sum_{i=k+1}^n M_i X_{(i)} + \sum_{j=s+1}^m N_j Y_{(j)}\right\}. \end{aligned}$$

Notice that

$$\sum_{i=k+1}^n M_i X_{(i)} + \sum_{j=s+1}^m N_j Y_{(j)} = \sum_{r=k+s+1}^{n+m} rU_{(r)} - \sum_{i=k+1}^n iX_{(i)} - \sum_{j=s+1}^m jY_{(j)}.$$

Table 1: Air-conditioning data

Plane	Failure times								
1 (X)	90	100	160	346	407	456	470	494	550
	570	649	733	777	836	965	983	1008	1164
	1474	1550	1576	1620	1643	1705	1835	2043	2113
	2214	2422							
2 (Y)	23	284	371	378	498	512	574	621	846
	917	1163	1184	1226	1246	1251	1263	1383	1394
	1397	1411	1482	1493	1507	1518	1534	1624	1625
	1641	1693	1788						
3 (Z)	97	148	159	163	304	322	464	532	609
	689	690	706	812	1018	1100	1154	1185	1401
	1447	1558	1597	1660	1678	1869	1887	2050	2074

Example 4.2. Table 1 gives failure times of air-conditioning systems of three different planes given by Proschan [5]. Times between failures are available in package *MixtureInf* of *R* Li:et:al:2016. Let X, Y and Z denote the failure times of the air-conditioning of planes 1 to 3, respectively. Zardasht and Asadi [12] used pairs (X, Y) of these data to estimate $R(t) = P(X_t > Y_t)$, where $X_t = X - t | X > t$. We use these data to evaluate $DCD(F_n, G_m; t)$. Figure 1 represents $DCD(F_n, G_m; t)$ for all pairs of these 3 planes. It shows that approximately pairs (X, Y) have most distance and pairs (X, Z) have least distance. Also, if compare results with figure 5 of Zardasht and Asadi [12], we see that both models $R(t)$ and $DCD(F_n, G_m; t)$ have the same behavior.

Using theorem 4.1 we have the following corollary.

Corollary 4.3. *In Theorem 4.1, if* $t \to 0^+$, *then* $CD(F_n, G_m)$ *can be rewritten as*

$$CD(F_n, G_m) = \frac{\overline{X}}{n} - \frac{2}{n^2}\sum_{i=1}^{n} iX_{(i)} + \frac{\overline{Y}}{m} - \frac{2}{m^2}\sum_{j=1}^{m} jY_{(j)} + \frac{2}{nm}\left\{\sum_{i=1}^{n} M_i X_{(i)} + \sum_{j=1}^{m} N_j Y_{(j)}\right\}.$$

Example 4.4. The top left frame of Figure 2 shows a sample picture of two parrots. The picture has 768×512 cells. The gray level of each cell has a value between 0 (black) and 1 (white). The top right, bottom left and bottom right frames are increased brightness, adjusted contrast and gamma corrected versions of the original frame, respectively. This sample picture and functions for adjusting it are available in package *EBImage* of *R* Pau:et:al:2010,Pau:et:al:2016. Let name the frames from top to bottom and left to right by X, $Y(= X + 0.3)$, $Z(= 2X)$ and $W(= \sqrt{X})$, respectively. Using corollary 4.3, we want see how much is difference of each pairs of these frames.

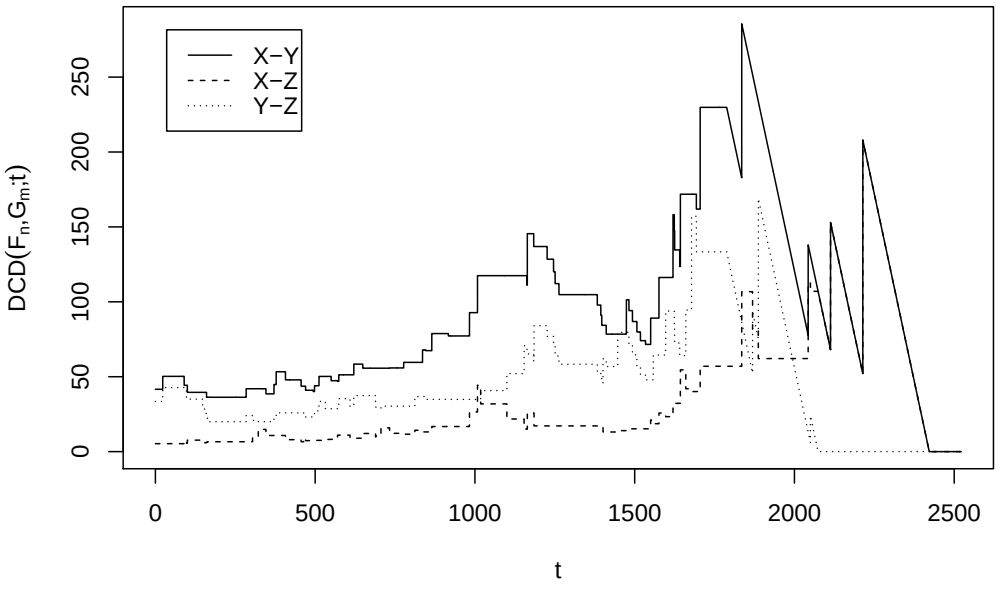


Figure 1: $DCD\,(F_n, G_m; t)$ for all pairs of air-conditioning data

Figure 2: Sample picture of two parrots with its adjustments

Table 2: $CD\left(F_n, G_m\right)$ for each pairs of Figure 2

	X	Y	Z	W
X	0	0.1336	0.1673	0.0882
Y	0.1336	0	0.0222	0.0116
Z	0.1673	0.0222	0	0.0519
W	0.0882	0.0116	0.0519	0

Table 2 represents $CD\left(F_n, G_m\right)$ for each pairs of these frames. $CD\left(F_n, G_m\right)$ states that the frame Z has most and the frame W has least distance with the original frame X. Also pairs (X, Z) have most and pairs (Y, W) have least distance between all pairs.

Let $X_1, \cdots, X_n$ be a random sample of a parametric family $F\left(x; \boldsymbol{\theta}\right)$, $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^k$. Based on empirical $s.f.$, we can define $DCD\left(F_n, F; t\right)$ as dynamic Cramér's distance of F_n and F.

Definition 2. *Under above considerations, the dynamic Cramér's distance of F_n and F is defined as*

$$\begin{aligned} DCD\left(F_n, F; t\right) &= \int_t^{\infty} \left(\bar{F}_{nt}\left(x\right) - \bar{F}_t\left(x; \boldsymbol{\theta}\right)\right)^2 dx \\ &= \int_t^{\infty} \left(\frac{\bar{F}_n\left(x\right)}{\bar{F}_n\left(t\right)} - \frac{\bar{F}\left(x; \boldsymbol{\theta}\right)}{\bar{F}\left(t; \boldsymbol{\theta}\right)}\right)^2 dx. \end{aligned} \tag{A.3}$$

In following, we will find simple form of (A.3).

Theorem 4.5. *Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ be ordered sample of $F\left(x; \boldsymbol{\theta}\right)$. If $X_{(k)} \leq t < X_{(k+1)}$, then $DCD\left(F_n, F; t\right)$ can be rewritten as*

$$\begin{aligned} DCD\left(F_n, F; t\right) = \frac{2n+1}{\left(n-k\right)^2} \sum_{i=k+1}^{n} X_{(i)} - \frac{2}{\left(n-k\right)^2} \sum_{i=k+1}^{n} i X_{(i)} - t \\ + m_F^{(2)}\left(t\right) - \frac{1}{n-k} \sum_{i=k+1}^{n} h\left(X_{(i)}\right), \end{aligned}$$

where

$$h\left(x\right) = 2\int_t^x \frac{\bar{F}\left(y; \boldsymbol{\theta}\right)}{\bar{F}\left(t; \boldsymbol{\theta}\right)} dy, \;\; x > t.$$

$DCD\left(F_n, F; t\right)$ enables us to do some inferences on parameter $\boldsymbol{\theta}$, for example in parameter estimation and goodness of fit testing. A dynamic estimator of $\boldsymbol{\theta}$ can be found for example by

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} DCD\left(F_n, F; t\right).$$

References

[1] Cramér, H. (1928). On the composition of elementary errors: Second paper: Statistical applications. *Scandinavian Actuarial Journal*, **11(1)**, 141-180.

[2] Li, S., Chen, J., & Li, P. (2016). *MixtureInf: Inference for Finite Mixture Models. R package version 1.1.* https://CRAN.R-project.org/package=MixtureInf.

[3] Pau, G., Fuchs, F., Sklyar, O., Boutros, M., & Huber, W. (2010): EBImage - an R package for image processing with applications to cellular phenotypes. *Bioinformatics*, **26(7)**, 979-981.

[4] Pau, G., Oles, A., Smith, M., Sklyar, O., & Huber, W. (2016). *EBImage: Image processing toolbox for R. R package version 4.14.2.* http://bioconductor.org/packages/release/bioc/html/EBImage.html.

[5] Proschan, F. (1963). Theoretical Explanation of Observed Decreasing Failure Rate. *Technometrics*, **5(3)**, 375-383.

[6] Rizzo, M. L. (2002). A test of homogeneity for two multivariate populations. *Proceedings of the American Statistical Association, Physical and Engineering Sciences Section.*

[7] Rizzo, M. L., & Székely, G. J. (2016a). Energy distance. *Wiley Interdisciplinary Reviews: Computational Statistics*, **8(1)**, 27-38.

[8] Rizzo, M. L., & Székely, G. J. (2016b). *energy: E-Statistics: Multivariate Inference via the Energy of Data. R package version 1.7-0.* https://CRAN.R-project.org/package=energy

[9] Székely, G. J. (1989). Potential and kinetic energy in statistics. *Lecture Notes, Budapest Institute of Technology (Technical University).*

[10] Székely, G. J. (2003). E-Statistics: The energy of statistical samples. *Bowling Green State University, Department of Mathematics and Statistics Technical Report* **(03-05)**, 1-18.

[11] Székely, G. J., & Rizzo, M. L. (2013). Energy statistics: A class of statistics based on distances. *Journal of Statistical Planning and Inference*, **143(8)**, 1249-1272.

[12] Zardasht, V., & Asadi, M. (2010). Evaluation of $P\left(X_t > Y_t\right)$ when both X_t and Y_t are residual lifetimes of two systems. *Statistica Neerlandica*, **64(4)**, 460-481.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Planning of step-stress accelerated degradation test with dependent competing risks

Mireh, S. [1] Khodadadi, A. [2] and Haghighi, F. [3]

[1,2] Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University
[3] Department of Mathematics, Statistics and Computer Science, University of Tehran

Abstract

Today, many products are designed to function for a long period of time before they fail. For such highly-reliable products, collecting step-stress accelerated degradation test (SSADT) data can provide useful reliability information. SSADT is more useful for developing products when there is inadequate knowledge for test conditions. Some efficient SSADT plans have been proposed when the underlying failure modes are independent. However, how to design an efficient SSADT plan for the dependent failure modes with linear degradation data is still a problem to be solved. The aim of this paper is to provide an SSADT plan for these cases. Copulas are used to describe the dependence between failure modes. Finally, we use the proposed method to deal with the SSADT design with dependent failure modes for a real data set and a simulation study.

Keywords: Competing risk, Copula function, Linear degradation, Reliability function, Step-stress accelerated degradation test.

[1] s_mireh@sbu.ac.ir
[2] a_khodadadi@sbu.ac.ir
[3] haghighi@khayam.ut.ac.ir

1 Introduction

Accelerated life testing (ALT) is the method of testing a product by subjecting it to situations (stress, temperatures and so on.) in excess of its normal service parameters with a view to find faults and ability modes of failure in a short quantity of time. It is often used to speed up tests. In ALT, data concerning the lifetime distribution at higher levels of stress is extrapolated to attain an estimate of the lifetime distribution at a normal level of stress via a link among accelerated and normal environments. Because most of today's manufacturing products grow to be more reliable, ALT has evolved in latest years to in reliability tests for which product lifetimes are predicted to closing far beyond the allotted testing time.

A few current works on ALT with competing risks are addressed as follows. [8] studied the design of ALT with competing failure risks while the time-to-failure due to a particular risk is modeled by using a Weibull or a lognormal distribution. A Bayesian technique to designing an ALT for repairable systems with multiple independent failure modes became mentioned by [4]. They brought a power law procedure for modeling a failure manner of the repairable systems under the ALT. [3] proposed a way for planning multiple step-stress ALT with competing risks.

Alongside ALT, degradation testing and analysis has been significant in reliability assessment for industries, because it is difficult to observe failure times, or even degradation measurements, under normal operating conditions. Even upon the use of methods incorporating censoring and accelerating techniques, the information obtained about the lifetime distribution remains very restrained.

In this case, an alternative approach is to measure the degradation of components at higher levels of stress for predicting a products lifetime at a use-stress level. Such a test is known as an accelerated degradation test (ADT). ADT combines degradation testing with ALT by means of trying out the products in harsh environments, and measuring the degradation of product during the accelerated test. ALT and ADT share similar concepts and practices for underlying lifetime-stress models.

Accelerated degradation models have been studied by many authors. [7], extended the accelerated degradation models to a very general class of models that included Gaussian process models, geometric Brownian motion models and gamma process models, and described the stochastic degradation models with several accelerating variables. They also used stochastic degradation models with several accelerated variables ([7]). Later, [6], proposed direct methods to predict the lifetime distribution of organic light-emitting diodes (OLEDs) based on the ADT of the products. See [9] for more thorough reviews of the (accelerated) degradation models, and remaining useful life estimation from the models.

Although ADT is an efficient life-test method, it is generally very expensive to carry out. Obviously, for a newly developed product or expensive product, it is very difficult to have enough test units available for testing purposes. Under the circumstance, a constant-stress ADT is not applicable. To handle this problem, the SSADT has been proposed. It is a useful tool for assessing the lifetime distribution of highly reliable products (under a typical-use condition) when the available test items are very few. Recently, [10] discussed the SSADT problem based on gamma process. [5] discussed multiple-step step-stress accelerated degradation models based on Wiener

and gamma processes in a Bayesian framework, and applied a Markov chain Monte Carlo (MCMC) method to estimate the parameters in analytically intractable models.

In this paper, we the study a statistical inference for SSADT with competing failures. These competing failures can be either independent or dependent. But in recent studies, the dependency assumption is not considered or dependency of failure modes is considered without the presence of degradation (in SSALT) and there is almost no literature on dependency analysis in SSADT. Recently, [1], studied reliability estimation from linear degradation and multiple failure time data with competing risks under a SSADT, but they assumed multiple failure times are independent.

In this paper, we propose a modeling approach for jointly analyzing linear degradation data and dependent failure-times which are simultaneously recorded during the SSADT. We apply the copula model to combine the marginal distribution of the competing failure models with the time to failure distribution of the products, and derived a general statistical model for SSADT with dependent competing failures. It is widely believed that the copula method is the most convenient, important tool for solving dependence problems. Finally, We derive a likelihood function for failure-times from degradation data to provide maximum likelihood estimates (MLEs) of the parameters of the proposed model and the reliability function.

This paper is organized as follows. In section two, we describe the model and design for a SSADT and express the likelihood. In section three, we provide two examples from real and simulation data in an attempt to compare the efficiency of the SSADT plan proposed in this paper considering dependency between failure modes using the copula function and that given by [1]. Finally, conclusion and possible directions for future study are discussed in section four.

2 The model and design for a SSADT

In this section, we use a cumulative exposure model for step-stress ADT data. Let n units be placed on the test. The testing time starts in $\tau_0 = 0$. All of the units are first subjected to a normal stress, S_0. Afterwards, at changing time stress τ_1, all surviving units are moved to the stress level S_1 until time τ_2. The stress on a unit is thus increased step by step until it fails. So, we have a sequence of changing time points of stress $\tau_i, i = 0, \ldots, m+1$.

As previously mentioned, we consider the linear degradation process: $\{Z(t); t > 0\}$ as $Z(t) = t/A, A \in R$, where A is a random variable with a distribution function of $F_A(a; \boldsymbol{\eta})$ for the q-dimensional vector of parameters $\boldsymbol{\eta}$, which is influenced by stress levels. The degradation process at time t under the SSADT can be written as:

$$Z(t|S) = \begin{cases} Z_0(t|A(S_0)) = \frac{t-\tau_0}{a_0}, & \tau_0 < t \leq \tau_1, \\ Z_1(t|A(S_1)) = \frac{\tau_1-\tau_0}{a_0} + \frac{t-\tau_1}{a_1}, & \tau_1 \leq t < \tau_2, \\ \quad \vdots & \quad \vdots \\ Z_m(t|A(S_m)) = \sum_{j=0}^{m} \frac{\tau_j-\tau_{j-1}}{a_{j-1}} + \frac{t-\tau_m}{a_m}, & \tau_m \leq t < \tau_{m+1}. \end{cases}$$

Let T_{ij} be the failure time of i-th level of stress for the j-th unit, where $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$ and Z_{ij} are degradation values at time T_{ij}. Suppose we have s failure modes at each level of stress. Therefore failure time is defined as: $T_{ij} = \min\{T_{ij}^1, T_{ij}^2, , T_{ij}^s\}$ where the failure-times

caused by each failure modes are dependent. The failure mode at the i-th level of stress for the j-th unit is denoted by V_{ij}. In this case, we make use of the data set $\{T_{ij}, V_{ij}, Z_{ij}\}$. Suppose that the rates of failures caused by different failure mechanisms are increasing functions of degradation values (e.g., [2]). We denote the failure rate at time t, corresponding to the k-th failure mode at the i-th level of stress, by $\lambda_i^k(z(t))$.

According to [1], the conditional reliability function corresponding to the k-th failure mode is given by:

$$R_i^k(t|\boldsymbol{a}_i) = \begin{cases} R_0^k(t|\boldsymbol{a}_0), & \tau_0 < t \le \tau_1, \\ R_1^k(t|\boldsymbol{a}_1) = R_0^k(\tau_1 + \frac{a_0}{a_1}(t-\tau_1)|\boldsymbol{a}_1), & \tau_1 \le t < \tau_2, \\ \quad \vdots & \quad \vdots \\ R_m^k(t|\boldsymbol{a}_m) = R_0^k\left(a_0\left[\sum_{j=1}^m \frac{\tau_j-\tau_{j-1}}{a_{j-1}} + \frac{t-\tau_m}{a_m}\right]|\boldsymbol{a}_m\right), & \tau_m \le t < \tau_{m+1}. \end{cases}$$

Our goal is to model the reliability function of the cause-specific survival times random vector $\{T^1, .., T^s\}$ in the presence of degradation which is defined as follows:

$$R_T(t) = P(T > t) = P(\min(T^1, \ldots, T^s) > t) = R(t, \ldots, t). \tag{A.1}$$

Using copula, equation (A.4) is as follows:

$$R(t^1, \ldots, t^s) = C(R^1(t^1), \ldots, R^s(t^s)) \tag{A.2}$$

where C is the survival copula. We use Franks family of copulas because they exhibit symmetric dependence in both tails. The Frank copula is defined as:

$$C_\theta(u, v) = -\frac{1}{\theta}\log\left[1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{(\exp(-\theta) - 1)}\right], \quad \theta \neq 0.$$

From the observed data at the end of each step-stress under the SSADT and equations (A.2), the likelihood has the following form:

$$\begin{aligned} L = \prod_{i=0}^{m}\prod_{j=1}^{n_i}\sum_{k=1}^{2} \boldsymbol{I}(V_{ij} = k)(a_{0j}/a_{ij})\lambda_0^k(z_{ij};\boldsymbol{\gamma_k})R_0^k(t_{ij}^\star) \\ \frac{\exp(-\theta R_0^k(t_{ij}^\star))[\exp(-\theta R_0^{3-k}(t_{ij}^\star) - 1)]}{(\exp(-\theta) - 1)\exp(-\theta C(R_0^1(t_{ij}^\star), \ldots, R_0^k(t_{ij}^\star)|\boldsymbol{a_{ij}};\boldsymbol{\gamma_k}))} f_{A_i}(a_{ij};\boldsymbol{\eta}) \end{aligned} \tag{A.3}$$

where $\boldsymbol{\gamma_k}$ is q-dimensional vector of parameters that correspond to the k-th competing risk, $\boldsymbol{I(.)}$ is an indicator function, and a_{ij} is an observed value of A for the j-th unit at the i-th level of stress, $\boldsymbol{a_{ij}} \equiv (a_{0j}, a_{1j}, \ldots, a_{ij})^T$, $f_{A_i}(\boldsymbol{a_{ij}};\boldsymbol{\eta}) = \prod_{l=0}^{i} f_A(a_{lj};\boldsymbol{\eta_l})$, for $i = 0, \ldots, m$ and $t^\star$ is the equivalent failure-time at normal stress S_0:

$$t^\star = a_0\left[\sum_{j=1}^{i}\frac{\tau_j - \tau_{j-1}}{a_{j-1}} + \frac{t - \tau_i}{a_i}\right]. \tag{A.4}$$

The MLEs of $\underline{\boldsymbol{\gamma}} \equiv (\boldsymbol{\gamma_1}, \ldots, \boldsymbol{\gamma_s})^T$ and $\underline{\boldsymbol{\eta}} \equiv (\boldsymbol{\eta_0}, \ldots, \boldsymbol{\eta_m})^T$, $\underline{\hat{\boldsymbol{\gamma}}}$ and $\underline{\hat{\boldsymbol{\eta}}}$, cannot be obtained in closed form, thus we use numerical methods.

The unconditional reliability (survival) function can be defined as:

$$R(t) = \int_0^\infty C(R^1(t), \ldots, R^s(t)|A = a) dF_A(a, \boldsymbol{\eta_0})$$

Thus the estimate of reliability function is:

$$\hat{R}(t) = \int_0^\infty \hat{C}(R^1(t), \ldots, R^s(t)|A = a) dF_A(a, \hat{\boldsymbol{\eta}}_{\mathbf{0}}) \tag{A.5}$$

3 Examples

3.1 Simulation study

We generated n data sets under SSADT by considering a simple step-stress test with a changing time point of $\tau_1 = 40$. As mentioned earlier, we have assumed that the failure rates of two failure modes depends on the amount of degradation, $\lambda^k(z(t)) = (\theta_k z(t))^{\nu_k}$ for $k = 1, 2$. Suppose that $\theta_1 = 0.06, \nu_1 = 5, \theta_2 = 0.06$ and $\nu_2 = 5$. We consider the dependency between failure modes using Frank copula with parameter $\theta = 20$. In addition, we assume that the distribution of A in $Z = t/A$ is Weibull with parameters $\alpha = 5$ and $\beta = 4$.

Let a_{0j}, a_{1j} denote the values of A for the j-th unit at the first (normal) and second level of stress. We observe a_{0j} according to linear degradation as t_{0j}/z_{0j}, where t_{0j}, and z_{0j} are failure time and degradation of j-th unit at normal stress, but a_{1j} has been generated from a Weibull distribution. From (A.3) and simulated a_{1j}, the failure times in second level of stress can be computed. The considered sample sizes are $n = 50, n = 100$ and $n = 200$, each with one thousand replications.

The estimation of parameters of failure rates $(\theta_1, \nu_1, \theta_2, \nu_2)$, the parameters of A_i, (α_0, β_0), and the copula parameter, θ, have been obtained using a numerical solution of the likelihood function (A.3). Table 1 shows estimations in two cases: independent and dependent failure modes. The estimated median of the lifetime $(\hat{M})$, which is derived from (A.3) $(R(\hat{t}_{0.5}) = 0.5)$, is shown in the Table 1. According to Table 1, it can be seen that the estimated parameters differ in cases with and without dependency. Thus it is important to consider the dependency between failure modes in analyzing the survival time. The results provide insight into the sampling behavior of the estimators. They indicate that the MLEs approximate the true values of the parameters as the sample size n increases. Similarly, the standard errors (SE) and mean relative errors (MRE) decrease with increasing the sample size.

The estimate of reliability function $(\hat{R}(t))$ in two dependent and independent cases was also derived. In the dependent case, $\hat{R}(t)$ has been derived from (A.3). Figure 1 shows the reliability functions estimated using the proposed method in both dependent and independent cases (solid and dashed lines) with different sample sizes. They indicate that increasing the sample size decreases the difference between two curves.

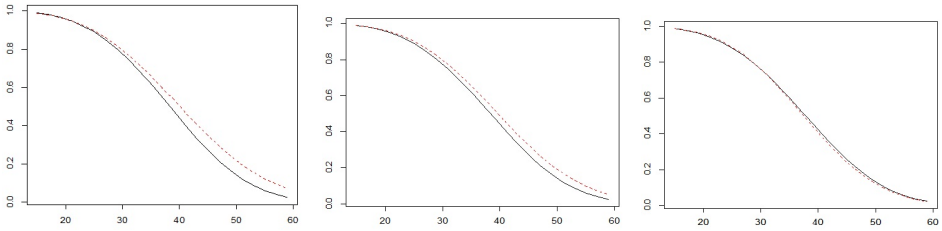
Table 1: The MLEs of the parameters, and the associated SE and MRE for different sample sizes ($\alpha_1 = 0.06$, $\nu_1 = 5$, $\alpha_2 = 0.06$, $\nu_2 = 5$, $\theta = 20$)

		Dependent			Independent		
n	Param.	Estimate	SE	MRE	Estimate	SE	MRE
n=50	α_0	5.139	0.572	0.092	5.139	0.572	0.092
	β_0	3.996	0.114	0.023	3.996	0.114	0.023
	α_1	0.059	0.005	0.067	0.037	0.024	0.399
	ν_1	5.228	0.893	0.136	3.628	2.193	0.389
	α_2	0.059	0.005	0.068	0.039	0.023	0.345
	ν_2	5.208	0.873	0.136	3.953	2.151	0.348
	θ	16.33	12.45	0.547	-	-	-
	τ_θ	0.779	-		0	-	-
	M	38.35	-	-	40.08	-	-
n=100	α_0	5.065	0.403	0.064	5.065	0.468	0.064
	β_0	3.996	0.082	0.016	3.996	0.096	0.016
	α_1	0.059	0.004	0.049	0.041	0.022	0.313
	ν_1	5.153	0.581	0.092	3.594	1.898	0.285
	α_2	0.059	0.004	0.046	0.044	0.021	0.261
	ν_2	5.144	0.550	0.088	3.937	1.799	0.346
	θ	16.55	11.54	0.515	-	-	-
	τ_θ	0.775	-	-	0	-	-
	M	38.03	-	-	39.24	-	-
n=200	α_0	5.048	0.278	0.044	5.048	0.278	0.044
	β_0	4.000	0.061	0.012	4.000	0.061	0.012
	α_1	0.059	0.003	0.036	0.053	0.003	0.110
	ν_1	5.080	0.408	0.065	4.987	0.511	0.082
	α_2	0.059	0.003	0.035	0.053	0.003	0.111
	ν_2	5.084	0.402	0.064	4.975	0.508	0.080
	θ	18.50	10.35	0.499	-	-	-
	τ_θ	0.803	-	-	0	-	-
	M	37.85	-	-	37.48	-	-

3.2 Real data

We use real data collected from 53 bus tires in a normal use environment as presented in a study by Bagdonavicius who have reported tire wear with two different failure modes: protector zone and side zone. These two failure modes are dependent since we assume the failure rate of both of them depends on degradation.

The method is similar to the previous example with respect to a changing time point of $\tau_1 = 60$. According to τ_1, 33 out of 53 failure times were at the use stress (S_0). As in the previous example, a_{0j}, a_{1j} are the values of A for the j-th unit at the first (normal) and second level of stress, where a_{1j} has been generated from a Weibull distribution with shape parameter, $\alpha = 10.60$, and scale parameter, $\beta = 4.50$. According to the data at normal stress, the estimation of shape and scale parameters are: $\alpha = 11.1190, \beta = 4.4826$.



(a) $n = 50$ (b) $n = 100$ (c) $n = 200$

Figure 1: The estimated reliability function in a dependent (solid line) and an independent (dashed line) case under the simple SSADT for simulation data with different sample sizes

All model parameters,$(\alpha_0, \beta_0, \theta_1, \nu_1, \theta_2, \nu_2, \theta)$, have been estimated using a numerical solution of the likelihood function (A.3). The last column of Table 2 is the estimated median of the lifetime $(\hat{M})$. According to Table 2, considering the dependency between failure modes is important. According to the relation between kendall's tau and parameter of copula, $\tau_\theta = 0.81$. Like the previous example, the estimate of reliability functions was also derived. Figure 2 shows the reliability functions estimated using a Nelson-Aalen nonparametric method (dashed line) and the proposed method in both dependent and independent cases (dotted and solid lines). The difference between the two curves (solid and dotted lines) is obvious in Figure 2. As we see, ignoring dependency between failure modes led to overestimation of the reliability function.

Table 2: The estimation of parameters with and without dependency for the bus tire data

	α_0	β_0	θ_1	ν_1	θ_2	ν_2	θ	M
Dependent case	11.119	4.4826	0.0655	12.168	0.0663	13.046	20.22	55.88
Independent case	11.119	4.4826	0.0616	11.611	0.0647	17.269	–	56.79

4 Coclusion

SSADT is one of the most commonly used methods for reducing the required sample size. When we have multiple failure modes in SSADT, it is important to take into account the dependency between them. This article has proposed a modeling approach for simultaneously analyzing dependent failure modes and degradation data under the SSADT. According to this method, we can estimate the reliability function in a dependent case and see the differences between the dependent and independent cases.

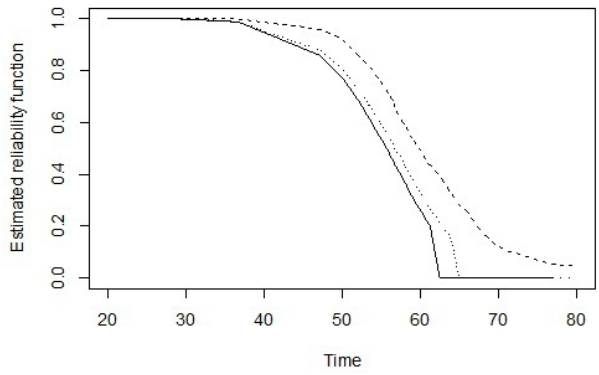


Figure 2: Nonparametric (dashed line) and parametric estimated reliability function in a dependent (solid line) and an independent (dotted line) case under the simple SSADT for bus tire data

References

[1] Haghighi, F. and Bae, S.J. (2015), Reliability Estimation from Linear Degradation and Failure Time Data With Competing Risks Under a Step-Stress Accelerated Degradation Test, *IEEE Transactions on Reliability*, **64**, 960-971.

[2] Liu, X., Li, J., Al-Khalifa, K., Hamouda, A., Coit, D. and Elsayed, E. (2013), Condition-based maintenance for continuously monitored degrading systems with multiple failure modes. *IIE Transactions*, **45**, 422-435.

[3] Liu, X. and Qiu, W. (2011), Modeling and planning of step-stress accelerated life tests with independent competing risks, *IEEE Transactions on Reliability*, **60**, 712-720.

[4] Liu, X. and Tang, L. (2010), Accelerated life test plans for repairable systems with multiple independent risks, *IEEE Transactions on Reliability*, **59**, 115-127.

[5] Pan, Z. and Balakrishnan, N. (2010), Multiple-steps step-stress accelerated degradation modeling based on Wiener and gamma processes, *Communications in Statistics - Theory and Methods*, **39**, 1384-1402.

[6] Park, J.I. and Bae, S.J. (2010), Direct prediction methods on lifetime distribution of organic light-emitting diodes from accelerated degradation tests, *IEEE Transactions on Reliability*, **59**, 74-90.

[7] Park, C. and Padgett, W.J. (2006), Stochastic degradation models with several accelerating variables, *IEEE Transactions on Reliability*, **55**, 379-390.

[8] Pascual, F. (2010), Accelerated life test planning with independent log-normal competing risks, *Journal of Statistical Planning and Inference*, **140**, 1089-1100.

[9] Si, X.S. , Wang, W., Hu, C.H. and Zhou, D.H. (2011), Remaining useful life estimation-a review on the statistical data driven approaches, *European Journal of Operational Research*, **213**, 1-14.

[10] 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**, 611-618.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Bayesian estimation for the STH distribution based on type II censored data

MirMostafaee, S.M.T.K. [1] and Mahdizadeh, M. [2]

[1] Department of Statistics, University of Mazandaran
[2] Department of Statistics, Hakim Sabzevari University

Abstract

In this paper, we consider type II right censored order statistics from the STH distribution which is a lifetime distribution. The maximum likelihood estimation is discussed. Then, we focus on the problem of Bayesian estimation of the unknown parameter. Since the integral related to the Bayes estimate can not be obtained explicitly, we propose the Markov chain Monte Carlo technique to derive the result. A real data example ends the paper.

Keywords: Bayesian estimation, MCMC technique, Censoring.

1 Introduction

Recently, Sarhan et al. [6] introduced a one parameter lifetime distribution with the following probability density function (pdf)

$$f(x) = \frac{\beta}{\beta+1}\left[\beta + (1+2\beta x)\mathrm{e}^{-\beta x}\right]\mathrm{e}^{-\beta x}, \quad x > 0,\ \beta > 0. \tag{A.1}$$

[1] m.mirmostafaee@umz.ac.ir
[2] mahdizadeh.m@live.com

Let us call this model the STH distribution. Sarhan et al. [6] referred to this distribution as $N(\beta)$ and we do the same in this paper. The cumulative distribution function of $X \sim N(\beta)$ is given by

$$F(x) = 1 - \frac{1}{\beta+1}\left[\beta + (1+\beta x)\mathrm{e}^{-\beta x}\right]\mathrm{e}^{-\beta x}, \quad x > 0. \tag{A.2}$$

In the sequel, we discuss the problem of estimation of the unknown parameter based on type II right censored order statistics (TOSs).

2 Main results

Let $\tilde{x} = (x_r, x_2, \cdots, x_r)$ be the observed TOSs extracted from a random sample of size n from $N(\beta)$ where $1 \le r \le n$, then the likelihood function is given by (see [1])

$$\begin{aligned} L(\beta, \tilde{x}) &= \frac{n!}{(n-r)!}[1-F(x_r)]^{n-r}\prod_{i=1}^{r} f(x_i) \\ &= \frac{n!\beta^r \mathrm{e}^{-\beta t}\left[\beta + (1+\beta x_r)\mathrm{e}^{-\beta x_r}\right]^{n-r}}{(n-r)!\,(1+\beta)^n}\prod_{i=1}^{r}\left[\beta + (1+2\beta x_i)\mathrm{e}^{-\beta x_i}\right], \end{aligned}$$

where $t = \sum_{i=1}^{r} x_i + (n-r)x_r$. The log likelihood function is

$$\begin{aligned} \ell(\beta, \tilde{x}) &= \log\frac{n!}{(n-r)!} - \beta t + r\log\beta + (n-r)\log\left[\beta + (1+\beta x_r)\mathrm{e}^{-\beta x_r}\right] \\ &\quad -n\log(1+\beta) + \sum_{i=1}^{r}\log\left[\beta + (1+2\beta x_i)\mathrm{e}^{-\beta x_i}\right]. \end{aligned}$$

The maximum likelihood estimate (MLE) of β, denoted by $\widehat{\beta}_M$, is obtained by maximizing $\ell(\beta, \tilde{x})$ with respect to (w.r.t.) β. Upon differentiating $\ell(\beta, \tilde{x})$ w.r.t. β and equating the result with zero, we get

$$\begin{aligned} \frac{\partial \ell(\beta, \tilde{x})}{\partial \beta} &= -t + \frac{r}{\beta} + (n-r)\frac{1-\beta x_r^2 e^{-\beta x_r}}{\beta + (1+\beta x_r)e^{-\beta x_r}} \\ &\quad -\frac{n}{1+\beta} + \sum_{i=1}^{r}\frac{1 + x_i e^{-\beta x_i}(1-2\beta x_i)}{\beta + (1+2\beta x_i)e^{-\beta x_i}} = 0. \end{aligned} \tag{A.1}$$

Next, we derive the Bayes estimate (BE) of β. suppose that β has the following prior distribution

$$\pi(\beta) = \frac{b^a}{\Gamma(a)}\beta^{a-1}e^{-b\beta}, \quad \beta > 0, \quad a, b > 0.$$

Then the posterior distribution of β given $\tilde{x}$ is

$$\pi(\beta|\tilde{x}) = \frac{1}{C_0}\beta^{r+a-1}\mathrm{e}^{-(b+t)\beta}h(\beta, \tilde{x}),$$

where

$$h(\beta,\tilde{x})=\left[\beta+(1+\beta x_r)\mathrm{e}^{-\beta x_r}\right]^{n-r}\prod_{i=1}^{r}\left[\beta+(1+2\beta x_i)\mathrm{e}^{-\beta x_i}\right],$$

and

$$C_0=\int_0^\infty \beta^{r+a-1}\mathrm{e}^{-(b+t)\beta}h(\beta,\tilde{x})\mathrm{d}\beta.$$

The BE under the squared error loss (SEL) function, denoted by $\widehat{\beta}_S$, is the mean of the posterior distribution, i.e.

$$\widehat{\beta}_S=\frac{1}{C_0}\int_0^\infty \beta^{r+a}\mathrm{e}^{-(b+t)\beta}h(\beta,\tilde{x})\mathrm{d}\beta.$$

But it seems that the above integral can not be obtained explicitly. Therefore, we use the Markov chain Monte Carlo (MCMC) technique and the Gibbs sampler to compute $\widehat{\beta}_S$. Clearly, we have

$$\pi(\beta|\tilde{x})\propto g(\beta|\tilde{x})h(\beta,\tilde{x}),$$

where $g(\beta|\tilde{x})$ is the density of the gamma distribution with parameters $r+a$ and $b+t$. We may consider the following algorithm to compute the approximated Bayes estimate (ABE) of β.

Algorithm 1:

- Step 1: Generate β_1 from $g(\beta|\tilde{x})$.
- Step 2: Repeat Step 1, N times to find $\beta_1,\cdots,\beta_N$.
- Step 3: The ABE of β, denoted by $\widehat{\beta}_{MS}$, is given by

$$\widehat{\beta}_{MS}=\sum_{i=1}^{N}\beta_i w_i,$$

where

$$w_j=\frac{h(\beta_j,\tilde{x})}{\sum_{i=1}^{N}h(\beta_i,\tilde{x})},\quad \text{for}\quad j=1,\cdots,N.$$

3 A real data example

In this section, we consider the remission times (in months) of 128 bladder cancer patients taken from [3]. The data are as follows:

0.08, 0.20, 0.40, 0.50, 0.51, 0.81, 0.90, 1.05, 1.19, 1.26, 1.35, 1.40, 1.46, 1.76, 2.02, 2.02, 2.07, 2.09, 2.23, 2.26, 2.46, 2.54, 2.62, 2.64, 2.69, 2.69, 2.75, 2.83, 2.87, 3.02, 3.25, 3.31, 3.36, 3.36, 3.48, 3.52, 3.57, 3.64, 3.70, 3.82, 3.88, 4.18, 4.23, 4.26, 4.33, 4.34, 4.40, 4.50, 4.51, 4.87, 4.98, 5.06, 5.09, 5.17, 5.32, 5.32, 5.34, 5.41, 5.41, 5.49, 5.62, 5.71, 5.85, 6.25, 6.54, 6.76, 6.93, 6.94, 6.97, 7.09, 7.26, 7.28, 7.32, 7.39, 7.59, 7.62, 7.63, 7.66, 7.87, 7.93, 8.26, 8.37, 8.53, 8.65, 8.66, 9.02, 9.22, 9.47, 9.74, 10.06, 10.34, 10.66, 10.75, 11.25, 11.64, 11.79, 11.98, 12.02, 12.03, 12.07, 12.63, 13.11, 13.29, 13.80, 14.24, 14.76, 14.77, 14.83, 15.96, 16.62, 17.12, 17.14, 17.36, 18.10, 19.13, 20.28, 21.73, 22.69, 23.63, 25.74, 25.82, 26.31, 32.15, 34.26, 36.66, 43.01, 46.12, 79.05.

The adequacy of the fitness of the STH distribution with $\beta = 0.08445444$ to the above data set was checked by using the Kolmogorov-Smirnov (K-S) test. The value of K-S test statistic was found to be $D = 0.071297$ and the corresponding p-value equaled 0.5334. We used the goodness.fit function, contained in the AdequacyModel package [4], in R to fit $N(\beta)$. Therefore, $N(\beta)$ fits the data satisfactorily.

Let us assume that the true value of β is 0.08445444. Here, $n = 128$ and we considered 3 censoring schemes with $r = 88, 100$ and 108. Since we had no prior information, we used the non-informative prior density with $a = b = 0$ for the Bayesian estimation. The MLEs, Exact BEs (obtained using the integrate function in R) and ABEs (with $N = 10000$) as well as the absolute values of the differences between these estimates and the true value of β were computed and reported in Table 1. we used the nleqslv function, contained in the nleqslv package [2], in R to solve (A.1). All the other computations were also performed using R [5].

We can see from Table 1 that the ABEs are quite close to the exact BEs. Therefore, we may conclude that Algorithm 1 works satisfactorily. Moreover, the MLEs are closer to the true value than the exact BEs but note that the true value of β was assumed to be the MLE of β based on the complete sample.

Table 1: The MLEs, exact BEs and ABEs as well as the absolute values of the differences between these estimates and the true value of β in parentheses.

r	MLE	Exact BE	ABE
88	0.09404 (0.00958)	0.09603 (0.01158)	0.10446 (0.02001)
100	0.09341 (0.00896)	0.12650 (0.04204)	0.10222 (0.01776)
108	0.09209 (0.00764)	0.18969 (0.10524)	0.10076 (0.01631)

References

[1] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1992), *A First Course in Order Statistics*, John Wiley & Sons, New York.

[2] Hasselman, B. (2014), nleqslv: Solve systems of non linear equations. R package version 2.5. https://CRAN.R-project.org/package=nleqslv.

[3] Lee, E.T. and Wang, J.W. (2003), *Statistical Methods for Survival Data Analysis*, Third edition, John Wiley & Sons, Hoboken, New Jersey.

[4] Marinho, P.R.D., Bourguignon, M. and Dias, C.R.B. (2013), AdequacyModel: Adequacy of probabilistic models and generation of pseudo-random numbers. R package version 1.0.8. https://CRAN.R-project.org/package=AdequacyModel.

[5] R Development Core Team (2016), R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

[6] Sarhan, A.M., Tadj, L. and Hamilton, D.C. (2014), A new lifetime distribution and its power transformation, *Journal of Probability and Statistics* Volume 2014, Article ID 532024, 14 pages.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

A look at the upper and lower means in reliability framework

Mirzaei, S. [1]

Department of Statistics, Payame Noor University

Abstract

In this work, a relation between the some notions in reliability and the upper and lower means is established. The similarity between income inequality curves and the upper and lower means gives the idea of connecting economy and reliability. So, we also refer the relationship between the Zenga new inequality curve and some uncertainty measures. The result is illustrated with several models commonly used in informetrics, such as exponential, Pareto and lognormal.

Keywords: Mean residual life, Mean waiting time, Zenga curve, Upper and lower means.

1 Introduction

The mean residual life (MRL) and the mean waiting time (MWT) of a non-negative random variable X plays an important role in reliability theory, survival analysis and other branches of probability and statistics. These notions are also related to income inequality studies. In this case, the age of an item has been considered as an income. Chandra and Singpurwalla (1981) [1] and Pham and Turkkan (1994) [4] proved that the Lorenz curve and the Gini index are closely related to MRL of reliability theory. In this work, we briefly describe the concepts of reliability theory in terms of the upper and lower means of variable. We also address the connection between the Zenga new inequality curve and some uncertainty measures. The result is illustrated with several models commonly used in informetrics, such as exponential, Pareto and lognormal. Finally, relationships with some other reliability concepts are also presented.

[1] sh_mirzaee@pnu.ac.ir

2 Definitions, notations and relationships

In this context, let X be a continuous random variable denoting the age of an item. We consider the cumulative distribution function (cdf) of X by F and $F^{-1}(p) = \inf\{x \mid F(x) \geq p; \quad p \in [0,1]\}$ be the inverse function of F. Given a population and a random variable X evaluated on it, for each $p \in (0,1)$, the population can be splitted into two groups: the first one, called lower group that consists of the proportion p of people with the lowest values of X, and the second one called upper group composed by all the others. Once the population is splitted into the lower and the upper group, the means of X in these two groups can be computed, obtaining the lower and the upper mean. The two following definitions delineate these two means.

Definition 1. *Let X be a continuous random variable with distribution function F and support $[a,b]$, where $0 \leq a < b \leq \infty$. The lower mean of X is defined as*

$$\mu_p^- = E(X \mid X \leq x) = \frac{1}{p}\int_0^p F^{-1}(u)du,, \qquad \forall p \in (0,1].$$

Definition 2. *Let X be a continuous random variable with distribution function F and support $[a,b]$, where $0 \leq a < b \leq \infty$. The upper mean of X is defined as*

$$\mu_p^+ = E(X \mid X \geq x) = \frac{1}{1-p}\int_p^1 F^{-1}(u)du,, \qquad \forall p \in [0,1).$$

It is easy to verify that for a random variabile X with expected value μ, the following formula holds true:

$$\mu = p\mu_p^- + (1-p)\mu_p^+, \qquad \forall p \in [0,1]. \tag{A.1}$$

2.1 Some consepts of reliability theory

Two ageing measure used in reliability analysis are the mean residual life

$$\begin{aligned} MRL(x) &= E(X - x \mid X > x), \\ &= \frac{1}{1-F(x)}\int_x^\infty (u-x)dF(u), \\ &= \frac{1}{1-F(x)}\int_x^\infty udF(u) - x, \end{aligned}$$

and the mean waiting time

$$\begin{aligned} MWT(x) &= E(x - X \mid X < x), \\ &= x - \frac{1}{F(x)}\int_0^x udF(u). \end{aligned}$$

2.2 Income inequalityt Curves

In the following, the most used inequality curves are considered. It is worth noting that they can be defined through the lower and upper group means.

The Lorenz curve introduced by Lorenz (1905) is the most well-known inequality curve used in the literature and several indices of income inequality are directly derived from this curve.

Definition 3. *Let X be a non-negative continuous random variable, with finite and positive expected value μ. The Lorenz curve of X is defined as:*

$$L(p) = \frac{\int_0^p F^{-1}(u)du}{\int_0^1 F^{-1}(u)du,} \qquad p \in [0,1],$$
$$= \frac{1}{\mu}\int_0^p F^{-1}(u)du, \qquad p \in [0,1].$$

The relation between the $L(p)$ curve, μ_p^- and μ_p^+ can be written as:

$$\mu_p^- = \frac{\mu L(p)}{p},$$
$$\mu_p^+ = \frac{\mu(1-L(p))}{1-p}.$$

Zenga (2007) [5] introduced a new inequality curve $Z(p)$, based on the ratio between the lower mean μ_p^- and the upper mean μ_p^+ of non-negative random variable by

$$Z(p) = 1 - \frac{\mu_p^-}{\mu_p^+}, \qquad p \in (0,1),$$

The link between the $Z(p)$ curve and the $L(p)$ curve can be obtained as:

$$Z(p) = \frac{p - L(p)}{p[1-L(p)]}. \tag{A.2}$$

The above equation shows that the Lorenz curve can be transformed into the Zenga curve and vice versa. From such $Z(p)$ curve, the related inequality index Z is defined as

$$Z = \int_0^1 Z(p)dp. \tag{A.3}$$

2.3 Relationships

The main result, which provides a simple explicit relation between the lower mean, upper mean, Zenga curve, mean residual life time and the mean waited time is presented in the following theorem.

Theorem 2.1. *Let X be a non-negative continuous random variable with finite and positive expected value. Then,*

$$MRL(x) = \mu_p^+ - x,$$
$$MWT(x) = x - \mu_p^-,$$

and

$$Z(p) = \frac{MRL(x) + MWT(x)}{MRL(x) + x},$$
$$= \frac{1}{p}(1 - \frac{\mu}{x + MRL(x)}), \qquad p \in (0, 1).$$

3 Some examples

In this section, we apply the main theorem to several relevant models of common use in informetrics.

3.1 Classical exponential distribution

Let X be a random variable with exponential distribution $F(x) = 1 - e^{-\frac{x}{\lambda}} I_{(x>0)}$, where $\lambda > 0$ is its scale parameter distribution. In this case, we have

$$\int_x^\infty [1 - F(t)]dt = \lambda \exp(-\frac{x}{\lambda})$$

and we find

$$MRL(x) = \lambda, \qquad x \geq 0.$$

Then, using (2.1) we obtain,

$$Z(p) = \frac{1}{p[1 - (\ln(1-p))^{-1}]}.$$

It is important to note that the scale parameter λ is not an inequality indicator, in fact the Zenga curve, Lorenz curve and inequality measures derived from them don't depend on λ (see Figure 1).

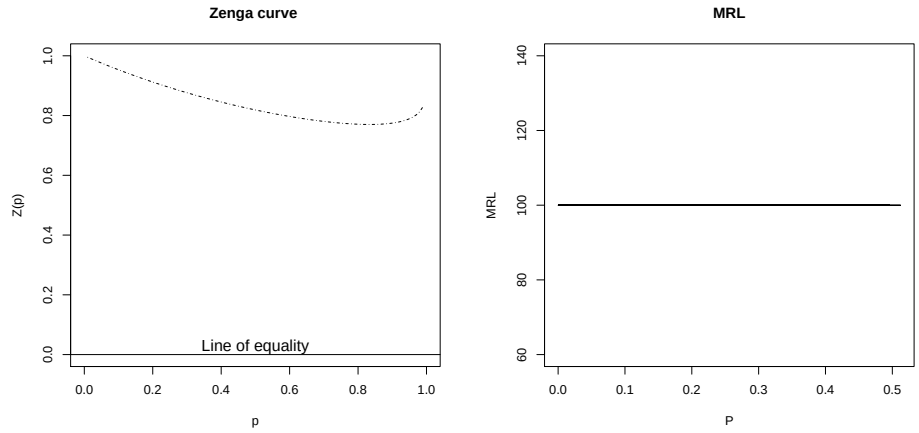


Figure 1: Zenga curve and mean residual life time in exponential distribution.

3.2 Power Distribution

Let X be a non-negative random income variable with power distribution and corresponding cdf,

$$F(x) = x^{\alpha}, \qquad 0 < x < 1$$

where $\alpha > 0$. In this case,

$$MRL(x) = \frac{\alpha(1 - x^{\alpha+1})}{(1 - x^{\alpha})(\alpha + 1)} - x, \qquad x > 0$$

and

$$Z(p) = \frac{1 - p^{\frac{1}{\alpha}}}{1 - p^{\frac{1}{\alpha}+1}}, \qquad p \in (0, 1)$$

In Figure 2 some MRL and $Z(p)$ curves are shown with different values of the distribution parameter α. It is evident that the value of α is an inverse inequality indicator. In the other words, if α increases, then $Z(p)$ curves decreases.

3.3 Pareto model

A random variable X follows a Pareto distribution if its distribution function is

$$F(x) = 1 - (\frac{x}{x_0})^{-\theta} I_{(x>x_0)},$$

where $x_0 > 0$ and $\theta > 1$. In this case,

$$MRL(x) = \frac{x}{x_0(\theta - 1)}, \qquad x > 0$$

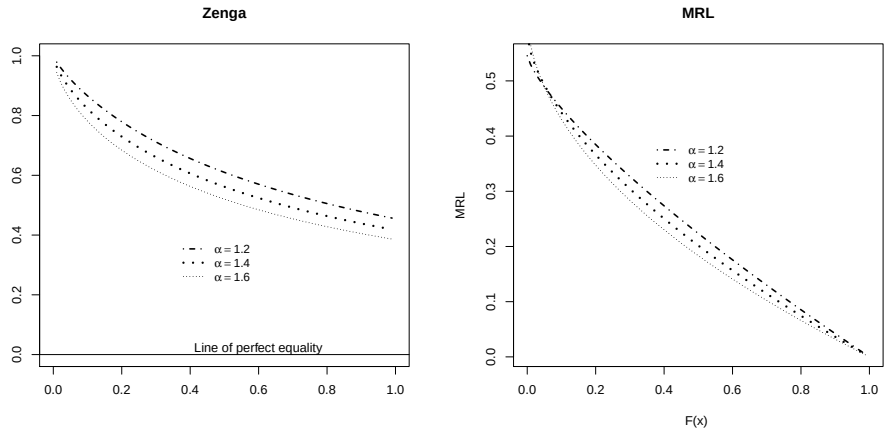


Figure 2: Zenga curve and mean residual life time in power distribution.

and from that, the Zenga curve can be obtained as

$$Z(p) = \frac{1-(1-p)^{\frac{1}{\theta}}}{p}, \quad p \in (0,1]. \tag{A.1}$$

In Figure 3 some MRL and $Z(p)$ curves are shown with different values of the distribution parameter θ. It is notable that the value of θ is an inverse inequality indicator and the scale parameter x_0 is not an inequality indicator.

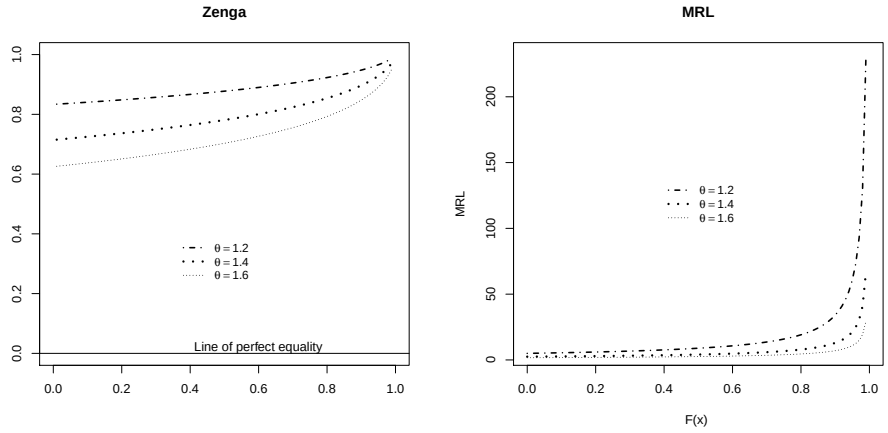


Figure 3: Zenga curve and mean residual life time in Pareto distribution.

3.4 Lognormal distribution

The classical lognormal distribution with the corresponding cdf is

$$F(x) = \Phi(\frac{\log(x) - \mu}{\sigma}), \qquad x > 0, \tag{A.2}$$

where $-\infty < \mu < \infty$, $\sigma > 0$ and $\Phi(\cdot)$ denotes the cdf of the standard normal distribution. Then, in this distribution

$$MRL(x) = \frac{\mu[1 - \Phi(\frac{\log(x)-\mu}{\sigma} - \sigma^2)]}{1 - \Phi(\frac{\log(x)-\mu}{\sigma})} - x, \qquad x \geq 0.$$

where $\mu = E(X) = e^{\frac{\mu+\sigma^2}{2}}$ and

$$Z(p) = \frac{p - \Phi(\Phi^{-1}(p) - \sigma^2)}{p[1 - \Phi(\Phi^{-1}(p) - \sigma^2)]}, \qquad p \in (0, 1).$$

4 Empirical illustration

As an example, the formulae presented in this section can be applied to the data - set represents the lifetime's data relating to relief times (in minutes) of 20 patients receiving an analgesic and reported by [2]. The data are as follows:

1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.5, 1.2, 1.4, 3.0, 1.7, 2.3, 1.6, 2.0

In Figure 4 empirical curves MRL, MWT, L(p), Z(p), lower and upper mean plots are drawn.

The four inequality curves are drawn together in the unitary square. The values of the related indexes corresponding to the areas below the curves can be computed as income inequality indices. The lower and upper mean plots are increasing. The Lorenz curve is convex and increasing. The Zenga curve is convex but at first decreases then increases.

5 Conclusion

In this paper, the relation between some notions in economic theory and reliability based on upper and lower means of variable has been shown. The considered notions are the mean waiting time, mean residual life and the new Zenga inequality curve. I hope that this discussion may attract wider applications in economic and applied statistics.

References

[1] Chandra, M. and Singpurwalla, N. D. (1981), Relationships between some notions which are common to reliability theory and economics, *Mathematics of Operations Research*, **5**, 113–121.

[2] Gross, A.J. and Clark, V.A. (1975), *Survival Distributions: Reliability Applications in the Biometrical Sciences*, John Wiley, New York.

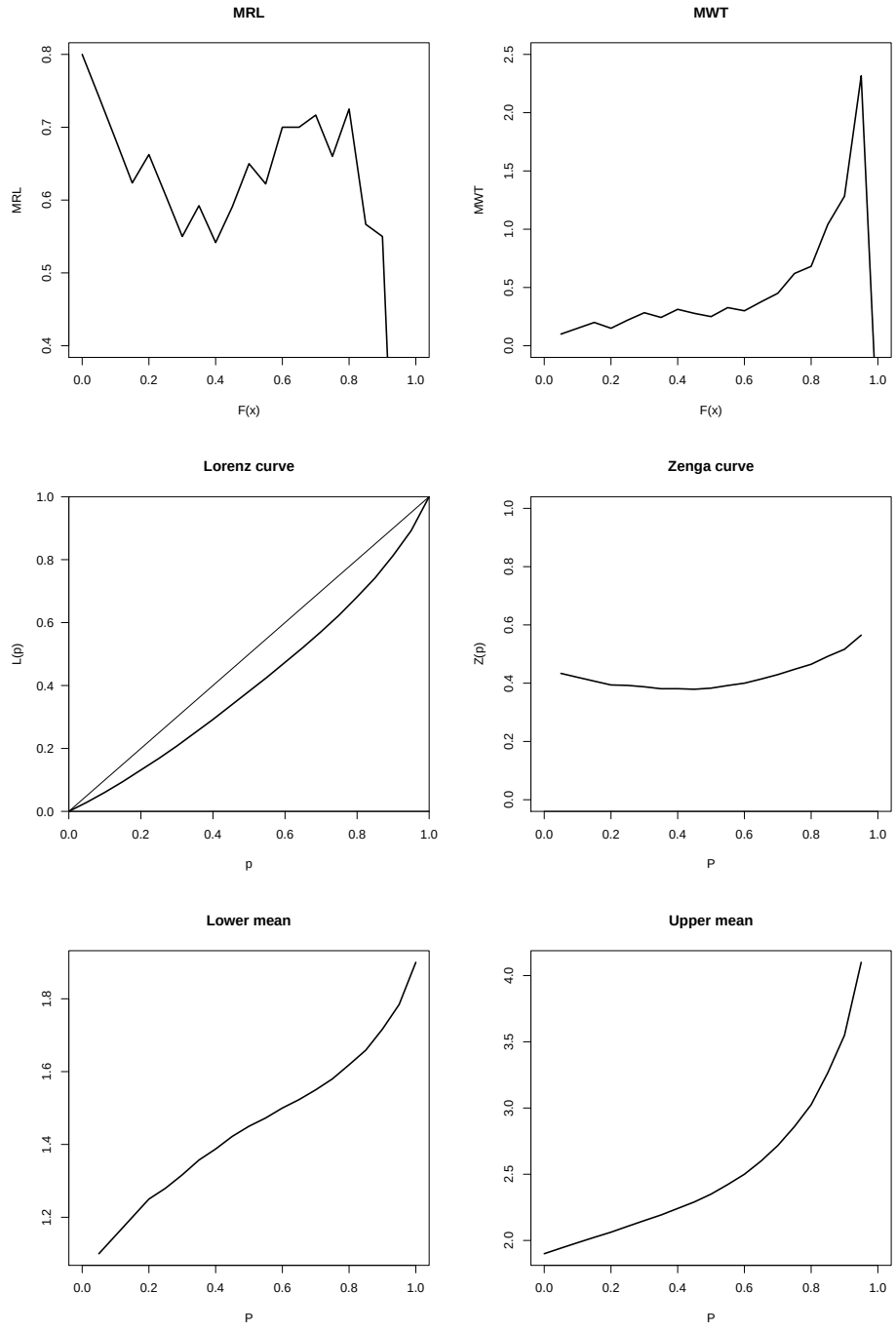


Figure 4: Inequality curves and uncertainly measures in lifetime's data.

[3] Lorenz, M. O. (1905), Methods of measuring the concentration of wealth, *Journal of the American Statistical Association*, **9** (70), 209-219.

[4] Pham, T. G. and Turkkan, N. (1994), The Lorenz and the scaled total-time-on-test transform curves: a unified approach, *IEEE Transactions on Reliability*, **43**, 76-84.

[5] Zenga, M. (2007), Inequality curve and inequality index based on the ratios between lower and upper arithmetic means, *Statistica Applicazioni*, **5** (1), 3-28.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Reliability of a k-out-of-n:F degradation system

Nezakati, E. [1] and Razmkhah, M. [2]

[1,2] Department of Statistics, Ferdowsi University of Mashhad

Abstract

In this paper, a k-out-of-n:F system with degrading components is considered. Likelihood and reliability function is derived based on degradations and the effect of presence a cold standby component is investigated on the reliability and parameter estimates. The gamma process is assumed for degradation of active and standby components. Finally the results are illustrated via a real data set.

Keywords: Cold standby component, Gamma process, Likelihood function, Reliability.

1 Introduction

For systems with high reliability, it is a difficult work to assess reliability with lifetime data, because a failure dosn't occur during short time at normal conditions. In this cases, degradation data contain more useful information than lifetime data about system reliability. So far, the degradation data have not been used to analyze the coherent systems. In this paper, we consider a k-out-of-n:F system and determine the failure times of the consisting components based on their deteriorations. As known in the literature of reliability, a k-out-of-n:F system consists of n components which fails if and only if at least k of its components fail. Such systems have various applications in engineering. For more details, we refer to Asadi and Bayramoglu (2006). On the other hand, there are different methods of redundancy to increase system reliability. One of them is to equip the system with cold

[1]nezakati-e@mail.um.ac.ir
[2]razmkhah_m@um.ac.ir

standby units which do not fail while they are waiting. Eryilmaz (2012) investigated the behavior of three different versions of mean residual life function corresponding to k-out-of-n system equipped with a cold standby unit. For recent studies, see also, Franko et al. (2015) and Wang et al. (2015). Here, we consider degradation performance of a k-out-of-n:F system with a single cold standby component and study the reliability of the system based on deterioration of consisting components. Toward this end, a Gamma process is considered for degradation data over time.

The rest of the paper is organized as follows. In Section 2, some preliminaries are presented regarding the Gamma process and time to failure distribution. Section 3 focuses on model description. In this section, some various likelihood functions are derived when either a cold standby component is used or not. In Section 4, the reliability function of a k-out-of-n:F system is investigated. A real data set is used to illustrate the proposed procedure in Section 5. Some conclusions are stated in Section 6.

2 Preliminaries

It is worthwhile to note that environment of an experiment, measurement errors, sample material, etc. are observed at random. So, degradation measurement is a random variable such that the parameters of its distribution change over time. Thus, it is recommended to model degradation performance in terms of a stochastic process. Since, degradation is monotone and status of samples is irreversible, Gamma process is much suitable for describing components degradation. Gamma process is a stochastic process with independent, non-negative and real-valued increments having a Gamma distribution. Let $X(t)$ denote the deterioration at time t, $t \geq 0$, which follow a Gamma process with the following probability density function

$$f_{X(t)}(x;\alpha(t),\beta(t)) = \frac{1}{\Gamma(\alpha(t))\beta(t)^{\alpha(t)}} x^{\alpha(t)-1} e^{-\frac{x}{\beta(t)}}; \quad x > 0, \tag{A.1}$$

where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}dx$ stands for the complete Gamma function. Note that the expectation and variance of $X(t)$ are

$$E(X(t)) = \alpha(t)\beta(t), \quad Var(X(t)) = \alpha(t)\beta^2(t),$$

where $\alpha(t)$ and $\beta(t)$ denote the shape and scale parameters at time t, respectively. Such process is denoted by $G(\alpha(t),\beta(t))$. For more details about gamma process, see Bagdonavicius and Nikulin (2000). Here we assume that $\alpha(t)$ is a nondecreasing, right-continuous and real-valued function for $t \geq 0$, and $\beta(t)$ is constant β, unchanging over time. Thus, process $\{X(t);\ t \geq 0\}$ has the following properties

(a) $X(0) = 0$ with probability one;

(b) $X(t)$ has independent increments;

(c) $X(t) - X(s) \sim G(\alpha(t) - \alpha(s), \beta)$ for all $t > s \geq 0$.

To determine the failure time of a component, we assume that it fails when its degradation reaches or exceeds a certain boundary level. In fact, for a given threshold level d, the lifetime T of the product is defined as the instant at which the degradation $X(t)$ reaches or exceeds d for the first time, i.e.,

$$T = \inf\{t \geq 0; \quad X(t) \geq d\}. \tag{A.2}$$

Using (A.1) and (A.1), the cdf and pdf of T are given by

$$F_T(t) = P(T \leq t) = P(X(t) \geq d) = \int_d^{\infty} \frac{1}{\Gamma(\alpha(t))\beta^{\alpha(t)}} x^{\alpha(t)-1} e^{-\frac{x}{\beta}} dx$$

and

$$f_T(t) = \alpha'(t) \int_{\frac{d}{\beta}}^{\infty} \big(\log y - \Psi(\alpha(t))\big) \frac{1}{\Gamma(\alpha(t))} y^{\alpha(t)-1} e^{-y} dy, \tag{A.3}$$

respectively, where $\alpha'(t)$ is the first derivative of $\alpha(t)$ with respect to t and $\Psi(t) = \frac{d}{dt}\log\Gamma(t)$ stands for the Digamma function.

3 Likelihood Functions

Let $X_1(t), ..., X_n(t)$ denote the degradations of n units of a k-out-of-n:F system at time t. Also, suppose that at the failure time of such system, a cold standby component with degradation $Y(t)$ is put into operation. The following assumptions are also considered throughout the paper:

(1) All active and standby components are statistically independent,

(2) Degradations of n active components are identically distributed with the Gamma process with the cdf $F(.; \alpha(t), \beta)$ at time t ,

(3) Degradation of standby component $Y(t)$ obeys the Gamma process with the cdf $F_s(.; \alpha_s(t), \beta_s)$ in active state.

(4) All components have increasing degradation paths and the ith component fails when $X_i(t)$ reaches or exceeds the given threshold value d.

(5) The degradations are measured at time points $0 \leq t_0 < t_1 < ... < t_m$.

As previously mentioned, a k-out-of-n:F system fails when the kth failure occurs. Suppose that the ν_ith $(i = 1, ..., k)$ component is the ith failed component for which its failure time is t_{r_i}. It is obvious that $\nu_1 \neq \nu_2 \neq ... \neq \nu_k$ and $1 \leq r_1 < r_2 < ... < r_k \leq m$. Denoting the parameters of Gamma process by $\boldsymbol{\theta}$, the associated likelihood function based on the degradation increments $\Delta_i(t_j) = X_i(t_j) - X_i(t_{j-1})$ is

$$L_1(\boldsymbol{\theta}) = \prod_{i=1}^{k-1} \prod_{j=1}^{r_i} f_{\Delta_{\nu_i}(t_j)}(\delta_{\nu_i,j}; \boldsymbol{\theta}) \prod_{\substack{h=1 \\ h \neq \nu_1,...,\nu_{k-1}}}^{n} \prod_{j=1}^{r_k} f_{\Delta_h(t_j)}(\delta_{h,j}; \boldsymbol{\theta}), \tag{A.1}$$

where $\delta_{i,j}$ is the observed value of $\Delta_i(t_j)$ with the pdf $f_{\Delta_i(t_j)}(\cdot;\boldsymbol{\theta})$. Moreover, $\prod_{i=1}^{0} a_i = 1$.

In the presence of a cold standby component, assume that a k-out-of-n:F system is failed at the ℓth inspection. That is, t_ℓ is the first inspection time after t_{r_k} for which the degradation of one of active components reaches or exceeds the threshold d, i.e., $r_k < \ell \leq m$. Denoting the parameters of standby components by $\boldsymbol{\theta}_s$, the likelihood function based on degradation increments $\Delta_i(t_j)$ and standby degradation increments $\Delta^{'}(t_j) = Y(t_j - t_{r_k}) - Y(t_{j-1} - t_{r_k})$ in active state, is given by

$$L_2(\boldsymbol{\theta},\boldsymbol{\theta}_s) = \prod_{i=1}^{k}\prod_{j=1}^{r_i} f_{\Delta_{\nu_i}(t_j)}(\delta_{\nu_i,j};\boldsymbol{\theta}) \prod_{\substack{h=1 \\ h\neq\nu_1,\ldots,\nu_k}}^{n}\prod_{j=1}^{\ell} f_{\Delta_h(t_j)}(\delta_{h,j};\boldsymbol{\theta}) \prod_{j=r_k+1}^{\ell} f_{\Delta^{'}(t_j)}(\delta_j^{'};\boldsymbol{\theta}_s), \tag{A.2}$$

where $\delta_j^{'}$ is the observed value of $\Delta^{'}(t_j)$. From (A.2), it is deduced that inferences about $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_s$ may be made based on the separate likelihood functions

$$L_{21}(\boldsymbol{\theta}) = \prod_{i=1}^{k}\prod_{j=1}^{r_i} f_{\Delta_{\nu_i}(t_j)}(\delta_{\nu_i,j};\boldsymbol{\theta}) \prod_{\substack{h=1 \\ h\neq\nu_1,\ldots,\nu_k}}^{n}\prod_{j=1}^{\ell} f_{\Delta_h(t_j)}(\delta_{h,j};\boldsymbol{\theta}) \tag{A.3}$$

and

$$L_{22}(\boldsymbol{\theta}_s) = \prod_{j=r_k+1}^{\ell} f_{\Delta^{'}(t_j)}(\delta_j^{'};\boldsymbol{\theta}_s), \tag{A.4}$$

respectively. Furthermore, when the active and standby components come from the same processes, i.e., $\boldsymbol{\theta} = \boldsymbol{\theta}_s$, the likelihood function

$$L_3(\boldsymbol{\theta}) = L_3(\boldsymbol{\theta},\boldsymbol{\theta}) \tag{A.5}$$

may be used to make inference about $\boldsymbol{\theta}$.

4 System reliability

Here, we study the reliability of a k-out-of-n:F system when either a cold standby component is used or not. First of all, let us define the lifetime of such a system without standby component as follows

$$T_{k:n} = \inf\{t > 0; \quad X_{i_1}(t) \geq d, ..., X_{i_k}(t) \geq d\}, \tag{A.1}$$

where $\{i_1, ..., i_k\}$ is a permutation of $\{1, ..., n\}$. Reliability function of a k-out-of-n:F system at time x is calculated as

$$R(x) = P(T_{k:n} > x) = \sum_{i=0}^{k-1} \binom{n}{i} \bar{F}^i(d;\alpha(x),\beta) F^{n-i}(d;\alpha(x),\beta). \tag{A.2}$$

When a cold standby component is used, the failure time of a k-out-of-n:F system may be defined as

$$T_{k:n}^s = \inf\{t > T_{k:n}; \quad X_{j_1}(t) \geq d \;\; or \;\; Y(t - T_{k:n}) \geq d\},$$

where $j_1 = 1, ..., n \setminus \{i_1, ..., i_k\}$ and $T_{k:n}$ is the failure time of the system without standby component defined in (A.2).
Reliability function of a k-out-of-n:F system equipped with a cold standby component at time point x, can obtain with small changes in corollary 1 of Eryilmaz (2012) as follows

$$R^s(x) = R(x) + \binom{n-1}{k-1} F^{n-k}(d; \alpha(x), \beta) \times \int_0^x \bar{F}^{k-1}(d; \alpha(u), \beta) F_s(d; \alpha_s(x-u), \beta_s) f_{T_1}(u) du, \quad \text{(A.3)}$$

where $R(x)$ is as defined in (A.2), F and F_s stand for the cdfs of $X(t)$ and $Y(t)$, respectively.

5 Application to a real data set

Here, we use the fatigue crack growth data to illustrate the proposed methodology in this paper. This data set, which consist of 21 sample paths, was provided by Hudak et al. (1978) and also analyzed by Lu and Meeker (1993). These data are measured at $t_0 = 0, t_1 = 0.01, ..., t_{12} = 0.12$ million cycles and presented in Table 1. It is observed that initial crack length at shorten test time $t_0 = 0$ is equal to 0.9. Let us denote the observed measures by $D_i(t_j), i = 1, ..., 21, j = 1, ..., 12$. Figure 1 is a plot of crack lengths $D_i(t_j)$ versus time (in million cycles). It is obvious that the paths follow an exponential regularity. So, we consider a Gamma process for $X_i(t_j) = D_i(t_j) - 0.9$ with shape parameter e^{bt_j} and constant scale parameter a. With such parameters, the average of crack lengths has exponential growth, i.e., $E(X_i(t_j)) = ae^{bt_j}$.

Table 1. Fatigue crack growth data

	Cycles (in million)												
path	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.12
1	0.90	0.95	1.00	1.05	1.12	1.19	1.27	1.35	1.48	1.64			
2	0.90	0.94	0.98	1.03	1.08	1.14	1.21	1.28	1.37	1.47	1.60		
3	0.90	0.94	0.98	1.03	1.08	1.13	1.19	1.26	1.35	1.46	1.58	1.77	
4	0.90	0.94	0.98	1.03	1.07	1.12	1.19	1.25	1.34	1.43	1.55	1.73	
5	0.90	0.94	0.98	1.03	1.07	1.12	1.19	1.24	1.34	1.43	1.55	1.71	
6	0.90	0.94	0.98	1.03	1.07	1.12	1.18	1.23	1.33	1.41	1.51	1.68	
7	0.90	0.94	0.98	1.02	1.07	1.11	1.17	1.23	1.32	1.41	1.52	1.66	
8	0.90	0.93	0.97	1.00	1.06	1.11	1.17	1.23	1.30	1.39	1.49	1.62	
9	0.90	0.92	0.97	1.01	1.05	1.09	1.15	1.21	1.28	1.36	1.44	1.55	1.72
10	0.90	0.92	0.96	1.00	1.04	1.08	1.13	1.19	1.26	1.34	1.42	1.52	1.67
11	0.90	0.93	0.96	1.00	1.04	1.08	1.13	1.18	1.24	1.31	1.39	1.49	1.65
12	0.90	0.93	0.97	1.00	1.03	1.07	1.10	1.16	1.22	1.29	1.37	1.48	1.64
13	0.90	0.92	0.97	0.99	1.03	1.06	1.10	1.14	1.20	1.26	1.31	1.40	1.52
14	0.90	0.93	0.96	1.00	1.03	1.07	1.12	1.16	1.20	1.26	1.30	1.37	1.45
15	0.90	0.92	0.96	0.99	1.03	1.06	1.10	1.16	1.21	1.27	1.33	1.40	1.49
16	0.90	0.92	0.95	0.97	1.00	1.03	1.07	1.11	1.16	1.22	1.26	1.33	1.40
17	0.90	0.93	0.96	0.97	1.00	1.05	1.08	1.11	1.16	1.20	1.24	1.32	1.38
18	0.90	0.92	0.94	0.97	1.01	1.04	1.07	1.09	1.14	1.19	1.23	1.28	1.35
19	0.90	0.92	0.94	0.97	0.99	1.02	1.05	1.08	1.12	1.16	1.20	1.25	1.31
20	0.90	0.92	0.94	0.97	0.99	1.02	1.05	1.08	1.12	1.16	1.19	1.24	1.29
21	0.90	0.92	0.94	0.97	0.99	1.02	1.04	1.07	1.11	1.14	1.18	1.22	1.27

Now, we assume that the data are actually the sample paths of 21 components with a k-out-of-n:F structure and present the results for $k = 1, 2$. Also, as defined by Lu and Meeker (1993), a component is considered to be failed when its crack length reaches or exceeds the threshold $d = 1.6$ inches.

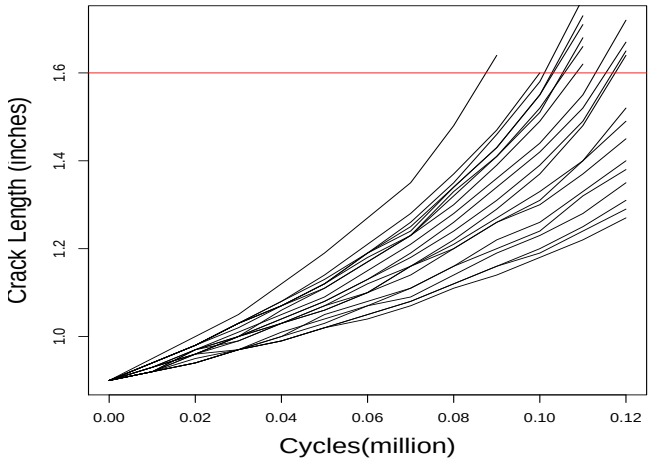


Figure 1: Fatigue crack length versus time

Using (A.1) and the entries of Table 1, the MLEs of a and b in a k-out-of-21:F system are numerically derived for $k = 1, 2$. Note that the values of $r_1 = 9$ and $r_2 = 10$ are observed from Table 1. Estimated values $(\hat{a}, \hat{b})$ are drived (0.027,31.271) and (0.0301,28.7021) for $k = 1$ and $k = 2$, respectively. Substituting the estimated values in (A.2), the reliability function may be estimated.

To obtain reliability function of the k-out-of-21:F system with a cold standby component, in addition to the degradations of active components, we need to have degradation of standby component when it is activated. Toward this end, we generate 1000 pseudo observations from $G(e^{b_s t}, a_s)$. For more investigation about the effect of the parameters of standby component, we use two different pairs of (a_s, b_s) such that the expected degradation of standby component is equal or less than the active components corresponded with the identically or non-identically distributed cases, respectively. Precisely, we use $(a_s, b_s) = (0.027, 31.271), (0.017, 21)$ for the case of $k = 1$ and $(a_s, b_s) = (0.0301, 28.7021), (0.02, 18)$ when $k = 2$. Using (A.3), the MLEs of a and b may be derived for different values of k in the presence of a cold standby component. Moreover, assuming $a_s = a$ and $b_s = b$, the MLEs may be obtained from (A.5). The results are presented in Table 2 for given values of (a_s, b_s), when $k = 1, 2$.

Table 2. Values of MLE in the presence of cold standby component.

	$k=1$		$k=2$	
(a_s, b_s)	$(0.027, 31.271)$	$(0.017, 21)$	$(0.0301, 28.7021)$	$(0.02, 18)$
$\hat{a}$	0.0304	0.0301	0.0315	0.0314
$\hat{b}$	28.6645	28.7021	27.4005	27.4015

The reliability of the k-out-of-21:F system may also be estimated from (A.3), when standby component is used. The estimated reliability functions are plotted in Figure 2 for $k = 1$ and 2. In this figure, the solid lines show the reliability for the systems without standby component. The dots

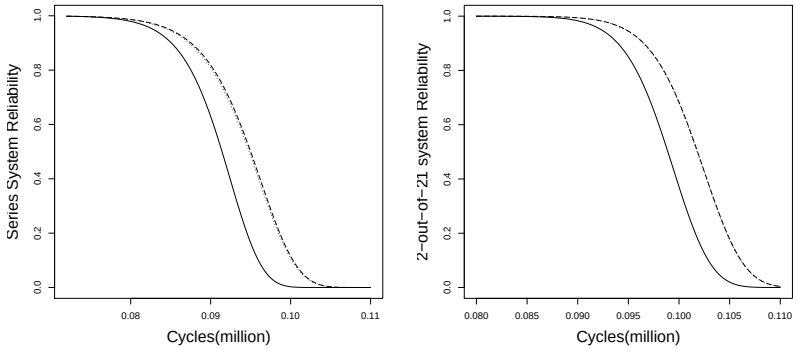


Figure 2: Reliability of a k-out-of-21:F system.

and dashed lines show the reliability of the systems in the presence of one standby component when its degradation is identical and non-identical with the active components, respectively. It is seen that the reliability increases significantly when a cold standby component is used, however, varying the degradation parameters of standby component have no significant effect on the reliability. Specially, when $k = 2$, maximum difference of reliability is 4.399926×10^{-3} such that the dots and dashed lines overlap approximately. Of course, the results are valid for the current data and they may be different for other data sets.

6 Conclusion

In this paper, degradation perormance of a k-out-of-n:F system was considered. Likelihood function was derived based on degradtions. Moreover, the effect of cold standby was investigated on reliability function and parameter estimates. It was shown in a real data set that the reliability of the system increases significantly when a cold standby component is used.

References

[1] Asadi, M. and Bayramoglu, I. (2006). The mean residual life function of a k-out-of-n structure at the system level. *IEEE Trans. Reliab.*, **55**(2), 314-318.

[2] Bagdonavicius, V. and Nikulin, M.S. (2000). Estimation in degradation models with explanatory variables. *Lifetime Data Anal*, **7**, 85-103.

[3] Eryilmaz, S. (2012). On the mean residual life of a k-out-of-n:G system with a single cold standby component. *Eur. J. Oper. Res.*, **222**, 273-277.

[4] Franko, C., Ozkut, M. and Kan, C. (2015). Reliability of coherent systems with a single cold standby component. *J. Comput. Appl. Math.*, **281**, 230-238.

[5] Hudak, S.J., Saxena, A., Bucci, R.J. and Malcolm, R.C. (1978). Development of standard methods of testing and analyzing fatigue crack growth rate data. *Tech. Rep.*, AFML-TR-78-40, Westinghouse R. and D. Center, Westing-House Electric Corporation, Pittsburg, PA.

[6] Lu, C.J. and W.O. Meeker. (1993). Using degradation measures to estimate a time-to-failure distribution. *Technometrics*, **35**(2), 161-174.

[7] Wang, W., Xiong, J. and Xie, M. (2015). Cold standby redundancy allocation problem with degrading components. *Int. J. Gen. Syst.*, **44**, 876-888.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

The stress-strength reliability under the environmental factors

Pakdaman, Z. [1] and Ahmadi,J. [2]

[1,2] Department of Statistics, Ferdowsi University of Mashhad

Abstract

Environmental factors such as temperature, humidity, dust, fumes or corrosive agents often have considerable influence on system reliability characteristics. In this paper, we assume that the effect of the environmental factors on the system can be modelled by a distortion function and under this assumption the stress-strength reliability is studied. Also, the dynamic stress-strength reliability and the number of working components of static stress-strength system are considered.

Keywords: Stress-strength reliability, Distortion function, Distorted probability distribution.

1 Introduction

There are various definitions for reliability indices in the literature. One of them is the probability of satisfactory performance of a system at a given time which works under certain conditions. In some experiments, it is desired to evaluate strength components under the other stress variables, which is called "stress-strength model" in reliability analysis. One issue of interest in the reliability engineering is the study of the probability that a random variable can dominate the other, that is the probability $P(X < Y)$ where X and Y are two non-negative random variables, defined on a same probability space $(\Omega, \mathscr{F}, P)$. If X and Y represent the lifetimes of two systems, then

[1]pakdaman.zohreh@stu.um.ac.ir
[2]ahmadi-j@um.ac.ir

$P(X < Y)$ is known as the probability that system with the lifetime X fails before another. The stress-strength reliability is defined by $P(X < Y)$. In many situation in reliability, due to exposure of the environmental factors such as, temperature, pressure, humidity, dust and etc, distribution of the system under these factors may be changed. For example, consider the designing of a bridge in the cities or an oil offshore platform on the sea. If the random variable Y represent strength of leg of bridge or oil platform and the random variable X is represented the stress of weight of the bridge or oil offshore platform on its legs. With the passing of the time and under the temperature, pressure or humidity the the random variables X and Y transform to the random variables X^* and Y^*, respectively, then $P(X^* < Y^*)$ is the probability of successful bridge design or oil platform during a mission. In this paper, we consider a stress-strength model under the environmental factors. We assume that these factors have almost equal effect, with the distortion function, on the original distribution of a stress-strength system. Distorted probability distributions are special cases of the more general theory of monotone set functions and non-additive measures, see, Denneberg (1994) for more details. The stress-strength reliability in static modeling is defined as the probability that the units strength is greater than the stress, that is, $P(Y > X)$, where Y is the random strength of the unit and X is the random stress placed on it. Dynamic modeling of stress-strength interference might offer more realistic applications to real-life reliability studies than static modeling and it enables us to investigate the time-dependent (dynamic) reliability properties of the system. The rest of this paper is organized as follows. Section 2 contains definition and auxiliary results. Section 3 deals with the investigation of the dynamic stress-strength model under the distortion functions. Exact expressions are obtained and the conditions to increase (or decrease) the reliability by concavity or convexity of distortion function are discussed. The number of working strength components under the environmental factors has been studied in Section 4.

2 Definition and auxiliary results

A distortion function, $\Phi(u)$, is a continuous and non-decreasing function on the interval $[0, 1]$ such that $\Phi(0) = 0$ and $\Phi(1) = 1$ and expect for a finite number of points, $\varphi(u) = \frac{d}{du}\Phi(u)$ exists on the interval $(0, 1)$. For a given distortion function $\Phi(\cdot)$, a distorted probability distribution can be defined by $F^* = \Phi(F)$. Suppose that under the environmental factors the random lifetimes of X and Y with distributions F and G transform to $\Phi(F)$ and $\Phi(G)$, respectively. Let us denote by X^* and Y^* any random variables that are distributed according to $F^* = \Phi(F)$ and $G^* = \Phi(G)$, respectively, which are also called the distorted random variables induced by Φ. If $\Phi(\cdot)$ is concave (or convex) distortion function we have $X^* \leqslant_{st} X$ (or $X \leqslant_{st} X^*$). For example, take $F(x) = 1 - e^{-x}$, $\Phi_1(u) = u^2$ and $\Phi_2(u) = u^{\frac{1}{2}}$. Then Figure 1 shows that the lifetime of the system has been reduced due to concavity of the distortion function.

Navarro et al. (2015) studied the preservation of stochastic orders under the formation of the generalized distorted distributions. The next result shows that some stochastic orders preserved under the distortion.

Result 1. *Let Φ be a continuous and strictly increasing distortion function, then*
(i) $X \leqslant_c Y$ if, and only if, $X^* \leqslant_c Y^*$.

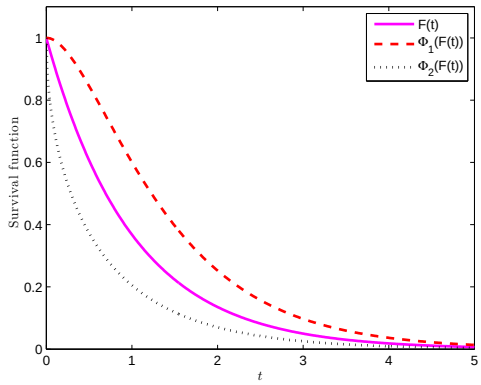


Figure 1: Survival function of distorted random variable

(ii) $X \leqslant_* Y$ *if, and only if,* $X^* \leqslant_* Y^*$.
(iii) $X \leqslant_{su} Y$ *if, and only if,* $X^* \leqslant_{su} Y^*$.
(iv) $X \leqslant_{disp} Y$ *if, and only if,* $X^* \leqslant_{disp} Y^*$.

Proof. By assumptions, we have

$$G^{*-1}(F^*(x)) = G^{*-1}(\Phi(F(x))) = G^{-1}(F(x)). \tag{A.1}$$

The results follow from (A.1) and Definition 2. □

3 Dynamic stress-strength reliability under the distortion function

Assume that the strength system consists of n i.i.d components and the deteriorating strength of the ith component at time t is denoted by the process $Y_i(t)$ for $i = 1, \ldots, n$ with cumulative distribution function G_t and with signature vector $\mathbf{s} = (s_1, \cdots, s_n)$. The random variable X represents the stress system with cumulative distribution function F and remains fixed over time (static). Let T_t be the system's lifetime at time t then (see, Samaniego, 2007)

$$P(T_t > u) = \sum_{i=1}^{n} s_i P(Y_{i:n}(t) > u). \tag{A.1}$$

The stress-strength reliability of this system is given by

$$R(t) = P(T_t > X) = \sum_{k=0}^{n-1} \bar{\mathbf{s}}_{\mathbf{k}} \binom{n}{k} \int_0^{\infty} (G_t(x))^k (\bar{G}_t(x))^{n-k} dF(x), \tag{A.2}$$

where $\bar{\mathbf{s}}_{\mathbf{k}} = \sum_{j=k+1}^{n} s_j$. Under the distortion function Φ, the stress-strength reliability in (A.2) is given by

$$R^*(t) = P(T_t^* > X) = \sum_{k=0}^{n-1} \bar{\mathbf{s}}_{\mathbf{k}} \binom{n}{k} \int_0^{\infty} (\Phi(G_t(x)))^k (\bar{\Phi}G_t(x))^{n-k} dF(x), \tag{A.3}$$

where T_t^* is lifetime of strength system under the environmental factors with component $Y_i^*(t)$ which has distribution function $G_t^* = \Phi(G_t)$. we have the following results for comparing $R^*(t)$ and $R(t)$.

Result 2. **(i)** *If* $\mathbf{s} = (0, \cdots, 1)$ *and* Φ *is concave (convex) then* $R^*(t) \leq (\geq) R(t)$.
(ii) *If* $\mathbf{s} = (1, \cdots, 0)$ *and* Φ *is concave (convex) then*$R^*(t) \leq (\geq) R(t)$.
(iii) *For* $n = 1$ *and* Φ *is concave (convex) then* $R^*(t) \leq (\geq) R(t)$.

In order to obtain the effect of environmental factor we evaluate $R^*(t) - R(t)$ under the two distortion function and for some special signature vectors. From Eryilmaz (2013), let $Y_i(t)$ be defined by

$$Y_i(t) = e^{-Ct},\ t \geq 0,\ i = 1, \ldots, n, \tag{A.4}$$

where C follows Pareto distribution with distribution function $F_C(x) = 1 - (\frac{\mu}{x})^2,\ x > \mu$, then the distribution function of $Y_i(t)$ for $i = 1, \ldots, n$ is

$$G_t(x) = \frac{\mu^2 t^2}{(\ln x)^2},\ 0 < x < e^{-\mu t}. \tag{A.5}$$

Let us take the following notations for two special signature vectors

$$\begin{aligned} C_1 &= \{\mathbf{s} | s_i = \frac{6i^2}{n(n+1)(2n+1)}, i = 1, \ldots, n\}, \\ C_2 &= \{\mathbf{s} | s_i = \frac{(\frac{1}{2})^i}{1 - (\frac{1}{2})^n}, i = 1 \ldots, n\} \end{aligned}$$

Assume that the random stress X has power distribution function $F(x) = x^\theta$, $0 < x < 1$, $\theta > 0$. In the following, under two distortion functions the expression $R^*(t) - R(t)$ has been studied.

• **Power distortion**: It is defined as

$$\Phi(u) = u^p,\ 0 \leq u \leq 1,\ p > 0. \tag{A.6}$$

It is obvious that $\Phi(u)$ is concave (convex) if $p \leq 1$ ($p \geq 1$). From (A.2) and (A.3) we obtain

$$\begin{aligned} R^*(t) - R(t) &= \sum_{k=0}^{n-1} \bar{\mathbf{s}}_\mathbf{k} \binom{n}{k} \int_0^{e^{-\mu t}} \Big[\big(\frac{\mu^2 t^2}{(\ln x)^2}\big)^{kp} \big(1 - \big(\frac{\mu^2 t^2}{(\ln x)^2}\big)^p\big)^{n-k} \\ &- \big(\frac{\mu^2 t^2}{(\ln x)^2}\big)^k \big(1 - \frac{\mu^2 t^2}{(\ln x)^2}\big)^{n-k} \Big]. \end{aligned} \tag{A.7}$$

For some selected values of θ and p, we have computed (A.7) as a function of t. The results are presented in the Figure 2 and it is seen that for two signature vectors and for $n = 5$ and $\mu = 0.5$, $R^*(t) - R(t) > 0$ ($R^*(t) - R(t) < 0$), if the distortion function be convex (concave). Note that from Figure 2 for large value of t we have $R^*(t) \simeq R(t)$.

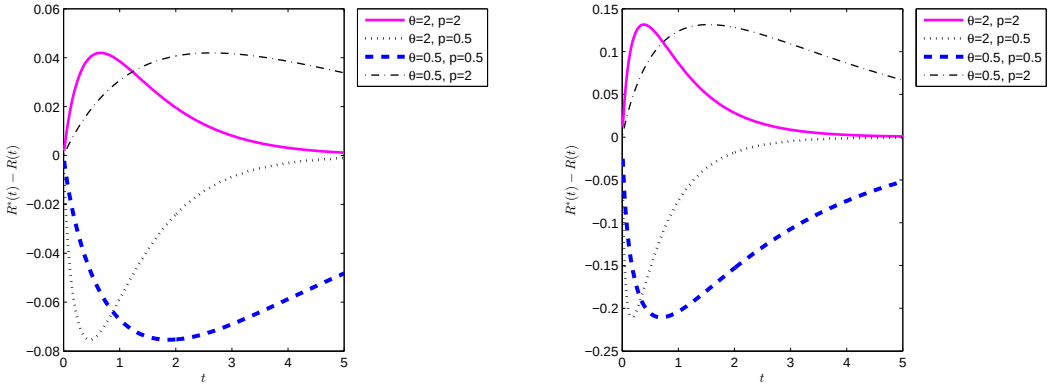


Figure 2: The plot of $R^*(t) - R(t)$ under the power distortion function for signature vectors C_1 and C_2 in left and right panel, respectively.

• **Dual Power distortion**: It is defined as

$$\Phi(u) = 1 - (1-u)^p, \ 0 \leq u \leq 1, \ p > 0. \tag{A.8}$$

It is obvious that $\Phi(u)$ is concave (convex) if $p \geq 1$ ($p \leq 1$). From (A.2) and (A.3) we obtain

$$\begin{aligned} R^*(t) - R(t) &= \sum_{k=0}^{n-1} \bar{\mathbf{s}}_{\mathbf{k}} \binom{n}{k} \int_0^{e^{-\mu t}} \Big[\big(1 - (1 - \frac{\mu^2 t^2}{(\ln x)^2})^p\big)^k \big(1 - \frac{\mu^2 t^2}{(\ln x)^2}\big)^{p(n-k)} \\ &- \big(\frac{\mu^2 t^2}{(\ln x)^2}\big)^k \big(1 - \frac{\mu^2 t^2}{(\ln x)^2}\big)^{n-k} \Big]. \end{aligned} \tag{A.9}$$

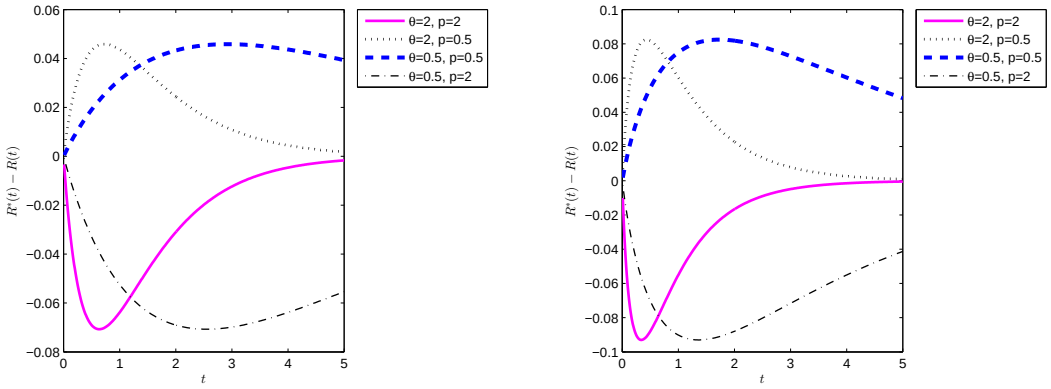


Figure 3: The plot of $R^*(t) - R(t)$ under the dual power distortion function for signature vectors C_1 and C_2 in left and right panel, respectively.

For some selected values of θ and p, we have computed (A.9) as a function of t. The results are presented in the Figure 3 and it is seen that for two signature vectors and for $n = 5$ and $\mu = 0.5$, $R^*(t) - R(t) < 0$ ($R^*(t) - R(t) > 0$), if the distortion function be concave (convex).

4 Stochastic comparison for the number of working strength components in static stress-strength model

Consider the stress-strength system consisting of n independent components whose random strengths are indicated by $Y_1, \cdots, Y_n$ with survival function $\bar{G}_1, \cdots, \bar{G}_n$, respectively. Assume that the components are subjected to a common stress X with distribution function F, and let X and Y_i for $i = 1, \cdots, n$ be independent random variables. Describe the random variable $N(l, X, Y)$ to be the random number of active components of the system who endured the stress X up to level strength l. Then we have

$$N(l, X, Y) = \sum_{i=1}^{n} I(Y_i - X > l), \tag{A.1}$$

where $I(A) = 1$ if event A occurs, and $I(A) = 0$, otherwise. The random variable defined by (A.1) may be useful for understanding the behaviour of a system over time, see Ling and Li (2013). The distribution of $N(l, X, Y)$ can be obtained by Theorem 3.8 of Ling and Li (2013) which is stated in the next result.

Result 3. *Let the random variable $N(l, X, Y)$ be as in (A.1), then for $0 \leq k \leq n$*

$$P(N(l, X, Y) = k) = \sum_{S} \prod_{u=1}^{k} \int_0^{\infty} \bar{G}_{j_u}(l + x) dF(x) \prod_{u=k+1}^{n} \int_0^{\infty} G_{j_u}(l + x) dF(x), \tag{A.2}$$

where the summation S expand over all permutation $(j_1, \ldots, j_n)$ of $1, \ldots, n$ for which $j_1, \cdots, j_k$ and $j_k, \cdots, j_n$ (Note that $\prod_{u=1}^{0} = 1$ and $\prod_{u=n+1}^{n} = 1$).

Proof. We have

$$P(N(l, X, Y) = k) = \sum_{S} P(Y_{j_1} - X > t, \cdots, Y_{j_k} - X > t, Y_{j_{k+1}} - X > t, \cdots, Y_{j_n} - X > t). \tag{A.3}$$

By equation (A.3) and $P(Y_{j_k} - X > t) = \int_0^{\infty} \bar{G}_{j_k}(t + x) dF(x)$ the proof has been completed. □

Let the aforementioned stress-strength system under the environmental factors has been changed and the distribution of random stress (or strength) has been re-weighted. Assume that the environmental factors have the effect on the original stress (or strength) distributions as a distortion function. Suppose that under the environmental factors the random lifetime of X (or Y) with distributions F (or G_i) transform to $\Phi(F)$ (or $\Phi(G_i)$). In the following, we give sufficient conditions for the stochastic orders between $N(l, X, Y)$ and $N(l, X^*, Y)$ (or $N(l, X, Y^*)$).

Result 4. *Assume that the stress-strength system with random stress X^* and random strength components $Y_1, \cdots, Y_n$.*
(i) If $\Phi(.)$ is a concave (or convex) distortion function, then $N(l, X, Y) \leqslant_{st} (\geqslant_{st}) N(l, X^*, Y)$.
(ii) If $\Phi(.)$ is a concave (or convex) and differentiable distortion function , then $N(l, X, Y) \leqslant_{lr} (\geqslant_{lr}) N(l, X^*, Y)$.

Similar Result 4, st and lr orders hold if the distribution of strength components changed. There are stated in the next result.

Result 5. *Assume that the stress-strength system with random stress X and random strength components $Y_1^*, \cdots, Y_n^*$.*
(i) If $\Phi(.)$ is a concave (or convex) distortion function, then $N(l, X, Y) \geqslant_{st} (\leqslant_{st}) N(l, X, Y^*)$.
(ii) If $\Phi(.)$ is a concave (or convex) and differentiable distortion function , then $N(l, X, Y) \geqslant_{lr} (\leqslant_{lr}) N(l, X, Y^*)$.

References

[1] Denneberg, D. (1994). *Non-Additive Measure and Integral*, Springer, Netherlands.

[2] Eryilmaz, S. (2013). On stress-strength reliability with a time-dependent strength. *Journal of Quality and Reliability Engineering.*

[3] Ling, X. and Li, p. (2013). Stochastic comparisons for the number of working components of a system in random environment, *Metrika*, **76**, 1017–1030.

[4] Navarro, J., del Águila, Y., Sordo, M. A. and Suárez-Llorens, A. (2016). Preservation of stochastic orders under the formation of generalized distorted distributions. Applications to coherent systems. *Methodology and Computing in Applied Probability*, **18**, 529–545.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

On Gini-type index applications in reliability analysis

Parsa, M. [1] Di Crescenzo, A. [2] and Jabbari, H. [3]

[1,3] Department of Statistics, Ferdowsi University of Mashhad
[2] Dipartimento di Matematica, Università degli Studi di Salerno

Abstract

The reliability scientists have always investigated the ageing properties of different systems structures and for this reason utilizing applicable tools is vitally important. The Gini-type index is an applicable diagnostic tool to realise the ageing properties of lifetime variables. In this paper Gini-type and its properties are studied. Besides different illustrative examples are given for further intuition. Later, a new stochastic order in terms of Gini-type index is introduced to compare the speed of ageing of components and systems.

Keywords: Reliability, Ageing properties, Stochastic comparisons.

1 Introduction

So far many researchers have been comparing different strategies under various assumptions specially when the systems elements are assumed to be dependent, such as Spizzichino [11], Navarro et al. [7], Khaledi and Shaked [3], Pellerey [9] and Navarro et al. [5], [6] and Borgonovo et al. [1].

Along with the previous mentioned investigations we propose to apply the Gini-type index defined by Kaminskiy and Krivtsoz [4] as a novel tool to gain useful information on the ageing characteristics of reliability systems.

[1]parsa.motahareh@mail.um.ac.ir
[2]adicrescenzo@unisa.it
[3]jabbarinh@um.ac.ir

Some extensions to multidimensional case of Gini-type index are also thoroughly investigated, by which the multivariate and conditional ageing properties of the lifetime variables are accessible. In particular, we show that the ageing properties of a component lifetime can vary when the other (dependent) components are working or have already failed.

This paper is organized as follows. In Section 2 the Gini-type index is introduced and its application in reliability theory is expressed. Section 3 defines a new stochastic order in terms of Gini-type index and different properties of this order are discussed.

Note that throughout the paper, 'log' means natural logarithm, and prime denotes derivative.

2 Gini-type index

The ageing behaviour of repairable or non-repairable systems is vitally important for maintenance strategies. Kaminskiy and Krivtsoz [4] introduced a simple index which could help to assess the degree of ageing or rejuvenation of repairable or non-repairable systems. Let X be a non-negative absolutely continuous (a.c.) random variable denoting the lifetime of a component or a system. Let $\bar{F}(t) = \mathbb{P}(X > t)$ and $H_X(t) = -\log \bar{F}(t)$ represent its survival function and cumulative hazard rate function, respectively. We recall that $h_X(t) = \frac{d}{dt} H_X(t)$ is the hazard rate of X. Assuming that

$$D_X^1 := \{t > 0 : \ 0 < \bar{F}(t) < 1\}, \tag{A.1}$$

the Gini-type (GT) index is introduced for all $t \in D_X^1$ as follows (see [4]).

Definition 1. *The GT index for a non-negative a.c. random variable X in time interval $(0, t]$ is*

$$GT_X(t) = 1 - \frac{2}{t\, H_X(t)} \int_0^t H_X(u)\, \mathrm{d}u, \qquad t \in D_X^1. \tag{A.2}$$

It is shown that GT index satisfies the inequality

$$-1 < GT_X(t) < 1 \quad \text{for all } t \in D_X^1.$$

We point out that the GT index can be seen a measure of the ageing property of the underlying random variable. Indeed, since $H_X(0) = 0$ and X is absolutely continuous, the following result holds (see [4]).

Proposition 2.1. *We have that*

(i) $GT_X(t) \geq (\leq)\ 0$ *for all* $t \in D_X^1$ *if and only if* X *is IFR (DFR);*

(ii) $GT_X(t) = 0$ *for all* $t \in D_X^1$ *if and only if* X *is CFR (constant failure rate), i.e.* X *has exponential distribution.*

Clearly, the GT index changes sign when the hazard rate is non-monotonic. For instance, if $h_X(t) = t(t-1)^2$, $t \geq 0$, then $GT_X(t) = \frac{3}{5} + \frac{4(t-2)}{5(6+t(3t-8)}$, $t \geq 0$, which is first positive, then negative, and finally positive when t increases.

As stated in [4], the index introduced in Definition 1 is defined similarly as the 'Gini coefficient, which is used in macroeconomics for analysing income distributions. In the reliability analysis of repairable systems it is of large interest to establish if the point process of the failure times is close to or far from the homogeneous Poisson process. The analysis of the GT index is thus useful to determine if the system is stable, or is improving, or is deteriorating.

Let us now investigate if the GT index can be a constant for further cases than the exponential distribution. This allows us to give a characterization result for the Weibull distribution in terms of the GT index, which extends case (ii) of Proposition 2.1.

Theorem 2.2. *The non-negative a.c. random variable X has Weibull distribution if and only if the corresponding GT index is constant.*

Proof. The proof of one side is straightforward. Thus, suppose that GT index of X is constant, i.e. $GT_X(t) = r$ for $t \in D_X^1$, with $-1 < r < 1$. Hence, from (A.1) we have

$$\frac{1}{t\,H_X(t)} \int_0^t H_X(u)\,\mathrm{d}u = \frac{1-r}{2}, \qquad t \in D_X^1.$$

Differentiating both sides with respect to t, since $H_X(t)$ is differentiable, we obtain

$$H_X'(t) - \frac{1+r}{1-r}\,\frac{H_X(t)}{t} = 0. \tag{A.3}$$

By solving the differential equation in (A.3), with $H_X(0) = 0$, and using (A.1) one attains

$$\bar{F}(t) = \exp\left\{-c\,t^{\frac{1+r}{1-r}}\right\}, \qquad -1 < r < 1, \quad t \in D_X^1,$$

with $c > 0$, this being the survival function of the Weibull distribution. □

The following example presents a number of distributions, where $D_X^1 = (0, \infty)$, with the corresponding GT index and the related limit behaviour.

Example 2.3. Consider the following survival functions, having support $(0, \infty)$:

(i) (Lomax distribution) $\bar{F}(t) = (1 + \frac{t}{\beta})^{-\alpha}$, $\alpha > 0$ and $\beta > 0$;

(ii) (Gompertz Makeham distribution) $\bar{F}(t) = \exp\{\alpha(1 - e^{\beta t})\}$, $\alpha > 0$ and $\beta > 0$;

(iii) (Log-logistic distribution) $\bar{F}(t) = \frac{1}{1+(t\beta)^{\alpha}}$, $\alpha > 0$ and $\beta > 0$;

(iv) (A bathtub-shaped hazard rate distribution) $\bar{F}(t) = \exp\{\alpha\frac{t^3}{3} - \alpha\beta t^2 + (\beta^2 + \lambda)t\}$, $\alpha > 0$, $\lambda > 0$ and $\beta > 0$, having the bathtub-shaped hazard rate $h_X(t) = \alpha(t - \beta)^2 + \lambda$.

Table 2.3 gives the corresponding GT indexes and their limits. In case (iii), Φ denotes the Lerch transcendence function, defined as

$$\Phi(z, s, a) = \sum_{k=0}^{\infty} \frac{z^k}{(k+a)^s}. \tag{A.4}$$

For a better intuition about the GT indexes in Table 2.3, see Figures 1–4.

	$GT_X(t)$	$t \to 0$	$t \to \infty$
(i)	$1 + 2\dfrac{(\beta + t)\log(1 + \frac{\beta}{t}) - t}{t\log(1 + \frac{\beta}{t})}$	0	-1
(ii)	$1 + 2\dfrac{-1 + e^{\beta t} - t\beta}{\beta t(1 - e^{\beta t})}$	0	1
(iii)	$1 + \dfrac{2(-(t\beta)^{\alpha}\Phi(-(t\beta)^{\alpha}, 1, 1 + \frac{1}{\alpha}) + \log(1 + (t\beta)^{\alpha}))}{-\log(1 + (t\beta)^{\alpha})}$		-1
(iv)	$\dfrac{\frac{\alpha}{6}t^2 - \frac{\alpha\beta}{3}t}{\frac{\alpha}{3}t^2 - \alpha\beta t + \lambda + \alpha\beta^2}$	0	$\dfrac{1}{2}$

Table 1: GT indexes and their limits for the distributions of Example 2.2.

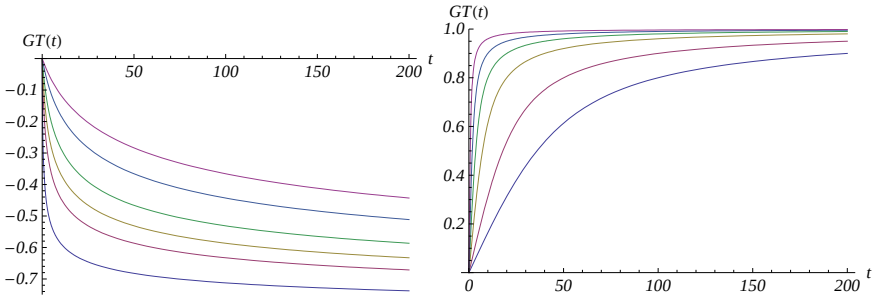


Figure 1: The GT index of the Lomax distribution (i) for $\beta = 0.1, 0.5, 1, 2, 5, 10$ from bottom to top (left) and the GT index of the Gompertz Makeham distribution (ii) for $\beta = 0.1, 0.5, 1, 2, 5, 10$ from bottom to top (right).

3 A GT index-based stochastic order

Let us now define a new stochastic order in terms of GT index, called 'GT order' for short. This is useful to compare the ageing properties of lifetime random variables.

Definition 2. *Let X and Y be two random lifetimes having GT indexes GT_X and GT_Y, respectively. We say that X is less than Y in GT index, and write $X \leq_{GT} Y$, if*

$$GT_X(t) \leq GT_Y(t) \quad \text{for all } t \in D^1 := D_X^1 \cap D_Y^1.$$

An equivalent condition is stated hereafter.

Proposition 3.1. *We have that $X \leq_{GT} Y$ if and only if*

$$\frac{\int_0^t H_Y(u)\,\mathrm{d}u}{\int_0^t H_X(u)\,\mathrm{d}u} \qquad \text{is non-decreasing in } t \in D^1. \tag{A.1}$$

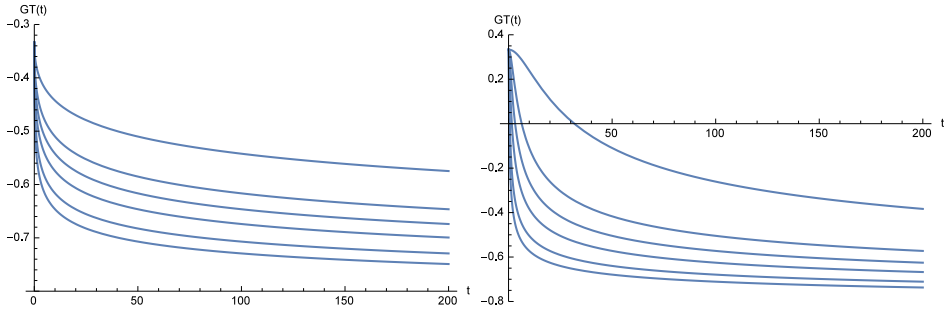


Figure 2: The GT index of the Log-logistic distribution (iii) for $\alpha = 0.5$ (left) and $\alpha = 2$ (right) where $\beta = 0.1, 0.5, 1, 2, 5, 10$ (from bottom to top).

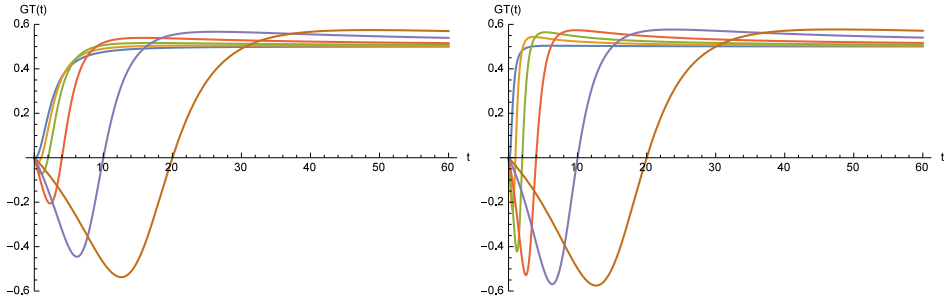


Figure 3: The GT index of the bathtub-shaped hazard rate distribution (iv) for $\alpha = 0.5$ (left) and $\alpha = 10$ (right) where $\lambda = 1$ and $\beta = 0.1, 0.5, 1, 2, 5, 10$ (from left to right).

Proof. By Definition 2 and Eq. (A.1), one concludes that $X \leq_{GT} Y$ if and only if the following condition is fulfilled for all $t \in D^1$:

$$H_Y(t) \int_0^t H_X(u)\,\mathrm{d}u - H_X(t) \int_0^t H_Y(u)\,\mathrm{d}u \geq 0, \tag{A.2}$$

Therefore, it is not hard to see that inequality (A.2) is equivalent to $\frac{d}{dt}(\int_0^t H_Y(u)\,\mathrm{d}u / \int_0^t H_X(u)\,\mathrm{d}u) \geq 0$, and then the stated result holds. □

Example 3.2. Let $X(a)$, $a > 0$, be a family of non-negative a.c. random variables having survival functions $\bar{F}_{X(a)}(t) = t^a$, $0 \leq t \leq 1$. The GT index of $X(a)$ is given by

$$GT_{X(a)}(t) = 1 - \frac{2}{\log(1-t^a)}\left[t^a \Phi\left(t^a, 1, 1+\frac{1}{a}\right) + \log(1-t^a)\right], \quad 0 < t < 1,$$

where Φ is the Lerch function defined in (A.4). It can be seen that $X(a)$ is increasing in a in the GT order, since $GT_{X(a)}(t) < GT_{X(b)}(t)$ for all a and b such that $0 < a < b < 1$ and all $t \in (0, 1)$.

In the following proposition we show that the random lifetimes X and Y having common support D^1 satisfy the proportional hazard rate model

$$\bar{F}_Y(t) = \left[\bar{F}_X(t)\right]^\alpha, \qquad t \in D^1, \quad \alpha > 0, \tag{A.3}$$

if and only if they have the same GT indexes.

Proposition 3.3. *The non-negative random variables X and Y having common support D^1 satisfy the proportional hazard rate property as in (A.3), if and only if*

$$GT_X(t) = GT_Y(t), \qquad \forall t \in D^1. \tag{A.4}$$

Proof. Let $H_X(t)$ and $H_Y(t)$ represent the cumulative hazard rate functions for X and Y, respectively. By the proportional hazard rate property (A.3) we have $H_Y(t) = \alpha\, H_X(t)$, $t \in D^1$, $\alpha > 0$. Thus, (A.4) holds trivially.

On the other hand when (A.4) holds, according to (A.2) one concludes that $\int_0^t H_Y(u)\,\mathrm{d}u / \int_0^t H_X(u)\,\mathrm{d}u = \alpha$, where $\alpha > 0$, which implies the proportional hazard rate property. □

Let us now discuss some further properties of the GT order.

Proposition 3.4. *Let X, Y and Z be non-negative random variables. The following properties hold:*

(i) *(reflexivity)* $X \le_{GT} X$.

(ii) *(transitivity) If* $X \le_{GT} Y$ *and* $Y \le_{GT} Z$ *then* $X \le_{GT} Z$.

(iii) $X \le_{GT} Y \iff aX + b \le_{GT} aY + b$ *for* $a, b \in \mathbb{R}^+$.

Proof. The proof is trivial by Proposition 3.1. □

Remark 3.5. Due to Proposition 3.3, the antisymmetry property of GT order does not hold in strict sense. Indeed, $X \le_{GT} Y$ and $Y \le_{GT} X$ are achieved simultaneously if and only if X and Y satisfy the proportional hazard rate model.

Sengupta and Deshpande [10] introduced the following partial orderings dealing with ageing properties. Hereafter we slightly modify their definitions in order to have more consistent notions.

Definition 3. *Given two non-negative random lifetimes X and Y, we say that X is*

- *ageing slower than Y, and write $X \le_c Y$, if $h_Y(t)/h_X(t)$ is non-decreasing in $t \in D^1$, or equivalently if $Z = H_Y(X)$ is IFR;*
- *ageing slower than Y in average, and write $X \le_\star Y$, $H_Y(t)/H_X(t)$ is non-decreasing in $t \in D^1$, or equivalently if $Z = H_Y(X)$ is IFRA.*

And therefore we have

$$X \leq_c Y \quad \Longrightarrow \quad X \leq_\star Y.$$

We stress that the inequalities given in Definition 3 have been inverted with respect to the corresponding ordering concepts in [10].

The following proposition expresses the relation between the GT order and the aforementioned lifetime orders.

Proposition 3.6. *Suppose that at least one of the non-negative random variables X and Y has continuous and strictly increasing distribution. Then we have*

$$X \leq_\star Y \quad \Longrightarrow \quad X \leq_{GT} Y. \tag{A.5}$$

Moreover, if ψ is a strictly increasing positive function passing through $(0,0)$ then

$$X \leq_\star Y \quad \Longrightarrow \quad \psi(X) \leq_{GT} \psi(Y). \tag{A.6}$$

Proof. According to Proposition 2.3 in [10] we have $X \leq_\star Y$ if and only if $H_Y(t)/H_X(t)$ is non-decreasing in $t \in D^1$. Since $H_X(u)$ and $H_Y(u)$ are non-negative non-decreasing functions of u, then it is not hard to see that the condition (A.1) in Proposition 3.1 is held and we have $X \leq_{GT} Y$.

Theorem 2.1 in [10] states that $X \leq_\star Y$ if and only if $\psi(X) \leq \psi(Y)$ for every strictly increasing positive function ψ passing through $(0,0)$. Hence, the validity of (A.6) follows straightforwardly from (A.5). □

The stochastic ordering introduced in Definition 2 is useful to assess the ageing properties of random lifetimes, indeed, due to the above results, condition $X \leq_{GT} Y$ means that X is ageing slower than Y in a broad sense. As example, hereafter we analyse the problem of improving a coherent system by introducing a redundant component by means of comparisons based on the GT order.

Conclusion

The Gini-type index is an applicable tool to investigate the ageing properties of lifetime variables. Specifically, when one needs to study complicated systems properties this index could be employed as a diagnostic gadget. According to Gini-type index we have introduced a new stochastic order by which lifetime comparisons are studied as well. Also, the relationship of this stochastic order with other orders in the literature has been expressed.

References

[1] Borgonovo, E., Aliee H., Glass, M., Teich, J. (2016) A new time-independent reliability importance measure. *European J. Oper. Res.* **254**, 427–442.

[2] Di Crescenzo, A., Pellerey, F. (2011) Improving series and parallel systems through mixtures of duplicated dependent components. *Naval Res. Logist.* **58**, 411–418.

[3] Khaledi, B., Shaked, M. (2007) Ordering conditional lifetimes of coherent systems. *IIE Transactions* **137,4**, 1173–1184.

[4] Kaminskiy, M.P., Krivtsoz, V.V. (2010) A Gini-type index for ageing/rejuvenating objects In: *Mathematical and Statistical Models and Methods in Reliability: Applications to Medicine, Finance, and Quality Control*. Statistics for Industry and Technology, Springer, pp. 133–140.

[5] Navarro, J., del Águila, Y., Sordo, M.A., Suárez-Llorens, A. (2013) Stochastic ordering properties for systems with dependent identically distributed components. *Appl. Stoch. Mod. Bus. Ind.* **29**, 264–278.

[6] Navarro, J., Pellerey, F., Di Crescenzo, A. (2015) Orderings of coherent systems with randomized dependent components. *European J. Oper. Res.* **240**, 127–139.

[7] Navarro, J., Ruiz, J.M., Sandoval, C.J. (2005) A note on comparisons among coherent systems with dependent components using signatures. *Stat. Prob. Lett.* **27**, 179–185.

[8] Navarro, J., Spizzichino, F. (2010) Comparisons of series and parallel systems with components sharing the same copula. *Appl. Stoch. Mod. Bus. Ind.* **26**, 775–791.

[9] Pellerey, F. (2008) On univariate and bivariate ageing for dependent lifetimes with archimedean survival copulas. *Kybernetika* **44**, 795–806.

[10] Sengupta, D., Deshpande, J.V. (1994) Some results on the relative ageing of two life distributions. *J. Appl. Prob.* **31**, 991–1003.

[11] Spizzichino, F. (2001) *Subjective probability models for lifetimes*. Monographs on Statistics and Applied Probability, 91. Chapman & Hall/CRC, Boca Raton, FL.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Optimum maintenance policy for a system with three types of failures

Safaei, F. [1] and Ahmadi, J. [2]

[1,2] Department of Statistics, Ferdowsi University of Mashhad

Abstract

Consider a system subject to three types of failures with different rate functions and different costs in which repairs and replacement. The type I and II failures arrive according to a non-homogeneous Poisson process where type I can be repair $(k-1)$ times by a minimal repair policy. We suppose that type III of failure is catastrophic failure and the system should be replaced with a new one. We are interested in determining an optimal planned replacement time which minimizes the expected discounted costs.

Keywords: Minimal repair, Non-homogeneous Poisson process, Optimal replacement.

1 Introduction

The majority systems deteriorate with age and usage are influenced by stochastic failures during operation. Failures of systems incur high costs. Some of failures are repairable and some of failures should be replaced. Finding an optimal replacement policy for a system becomes a major problem in a reliability studies. Barlow and Proschan (1965) presented the traditional age-replacement maintenance policy which a system is replaced at failure time or at age T, whichever occurs first. Several extensions of this policy have been investigated by researchers, such as Nakagawa and Kowada (1983) and Sheu and Chang (2009). Furthermore, Boland and Proschan (1982) considered

[1]safaei_fa123@yahoo.com
[2]ahmadi-j@um.ac.ir

the case of periodic replacement at times $kT(k = 1, 2, \ldots)$ and minimal repair if the system fails otherwise. This model has been extended by Nakagawa (1981), Boland and Proschan (1983) and Sheu (1996). Sheu et al. (2015) introduced optimal preventive maintenance and repair policies for multi-state systems. Lai and Chen (2016) proposed a bivariate replacement policy (n, T) for a cumulative shock damage process under cumulative repair cost limit. Aven and Casrto (2008) considered a system subject two types of failure, the system is minimally repaired and replacement. This paper presents an extension model for determining the optimal replacement policy based on the results of Aven and Casrto (2008).

2 Model Description

Minimal-repair and replacement are often used as practical maintenance activities of real reliability systems. A minimal repair is the maintenance activity to repair the failed system so that its function is recovered, without changing its age, while a replacement restores the entire system into the new condition so that it behaves as a new system. Further, replacement is classified into preventive replacement or failure (or corrective) replacement according as whether the system is in operation or in failure.

Consider a system subject to three types of failures with different rate functions in which repairs and replacement take place according to the following scheme:

A$_1$. The type I failures arrive according to a non-homogeneous Poisson process $\{N_1(t); t \geq 0\}$ with intensity function $r_1(t)$. This type of failure can be repair $(k-1)$ times by a minimal repair policy and at the time of the k-th failure system is replace with a new one. The process of replacement in type I failure denotes by $\{N_{1,k}(t); t \geq 0\}$ with intensity function $r_{1,k}(t)$.

A$_2$. The type II failures arrive according to a non-homogeneous Poisson process $\{N_2(t); t \geq 0\}$ with intensity function $r_2(t)$. When type II failure occurs at time t, the system is minimally repaired with probability $p(t)$ and replaced by a new one with probability $1 - p(t)$, where $0 \leq p(t) \leq 1$, $t > 0$. Therefore the process of minimal repair is non-homogeneous Poisson process $\{N_{2,m}(t); t \geq 0\}$ with intensity function $p(t)r_2(t)$ and the process of replacement is non-homogeneous Poisson process $\{N_{2,r}(t); t \geq 0\}$ with intensity function $(1 - p(t))r_2(t)$.

A$_3$. The type III failures arrive according to a non-homogeneous Poisson process $\{N_3(t); t \geq 0\}$ with intensity function $r_3(t)$. This type of failure is catastrophic failure, because when type III failure occurs the system should be replaced with a new one.

Further it is assumed that three types of failures are independent, also the intensity functions $r_i(t)$, $i = 1, 2, 3$ are continuous.
The system is replaced at a constant time T $(T > 0)$, $k-$th failure of type I failure, a non-repairable type II failure or at a type III failure, whichever occurs first. Replacement of type I, II and III failures are non-planned and replacement at age T is planned replacement.
The costs of minimal repair and replace for a type I failure at the k-the failure are c_1 and c_k, respectively. The costs as to the minimal repair and the replacement for a type II failure are $c_{2,m}$ and $c_{2,r}$, respectively. And the cost of replacement for a type III failure and planned replacement

of the system at age T are c_3 and c_r, respectively. Also, it is supposed that $c_r < \min\{c_k, c_{2,r}, c_3\}$. That is, the cost of the planned replacement is less than the costs of the non-planned replacements. All costs are positive real-values.

3 Main results

Let X_T, X_M and X_r be the time to replacement of the system, the time to a non-repairable type II failure and the time to a type III failure, respectively. Also let X_k be the time when the k-th failure of type I occur. Then, we have $X_T = \min\{X_k, X_M, X_r, T\}$. Denote the survival functions of X_k, X_M and X_r by $\bar{F}_k(t)$, $\bar{F}_M(t)$ and $\bar{F}_r(t)$, respectively, then for $T \geq t$, the survival function of X_T is given by

$$\bar{F}_{X_T}(t) = \bar{F}_k(t)\bar{F}_M(t)\bar{F}_r(t). \tag{A.1}$$

Under the assumption A_1, the process $N_1(t)$ is non-homogeneous Poisson processes, therefore

$$\begin{aligned}\bar{F}_k(t) = P(X_k > t) = P(N_1(t) < k) &= \sum_{i=0}^{k-1} \exp\left\{-\int_0^t r_1(u)du\right\} \frac{\left(\int_0^t r_1(u)du\right)^i}{i!}. \\ &= \int_t^\infty \frac{(\lambda_1(u))^{k-1}}{(k-1)!} r_1(t) e^{-\lambda_1(u)} du,\end{aligned}$$

where is $\lambda_1(t)$ cumulative failure intensity function for $N_1(t)$. From assumptions A_2 and A_3 we can write

$$\begin{aligned}\bar{F}_M(t) &= \exp\left\{-\int_0^t r_2(u)(1-p(u))du\right\}, \\ \bar{F}_r(t) &= \exp\left\{-\int_0^t r_3(u)du\right\}.\end{aligned}$$

The expected discounted cost associated for minimal repair for type I failure with cost c_1 during one cycle, by doing same steps as in Aven and Casrto (2008), is given by

$$E\left(\int_0^{X_T} c_1 e^{-\alpha t} d(N_1(t))\right) = c_1 \int_0^T e^{-\alpha t} \bar{F}_k(t)\bar{F}_M(t)\bar{F}_r(t) r_1(t) dt.$$

The expected discounted cost associated for the k-th failure in type I failure (replace) with cost c_k during one cycle is given by

$$E\left(\int_0^{X_T} c_k e^{-\alpha t} d(N_{1,k}(t))\right) = \int_0^{X_T} c_k e^{-\alpha t} \bar{F}_k(t)\bar{F}_M(t)\bar{F}_r(t) r_{1,k}(t) dt,$$

where $r_{1,k}(t)$ is intensity function of $N_{1,k}(t)$. By assumption A_1 we have

$$P(N_{1,k}(t) = r) = \sum_{j=rk}^{(r+1)k-1} e^{-\lambda_1(t)} \frac{(\lambda_1(t))^j}{j!}.$$

Then, by using the results in ages 7-8 of Aven and Jensen (1999), we can write

$$\begin{aligned} r_{1,k}(t) &= \lim_{h\to 0} \frac{E(N_{1,k}(t+h)) - E(N_{1,k}(t))}{h} \\ &= \lim_{h\to 0} \frac{1}{h} \sum_{r=0}^{\infty} r \sum_{j=rk}^{(r+1)k-1} \frac{e^{-\lambda_1(t+h)}(\lambda_1(t+h))^j - e^{-\lambda_1(t)}(\lambda_1(t))^j}{j!}. \end{aligned}$$

For repairable type II failures, non-repairable type II failures and type III failures with cost $c_{2,m}$, $c_{2,r}$ and c_3 during one cycle, the expected discounted cost are given by, respectively

$$E\left(\int_0^{X_T} c_{2,m} e^{-\alpha t} d(N_{2,m}(t))\right) = c_{2,m} \int_0^T e^{-\alpha t} \bar{F}_k(t) \bar{F}_M(t) \bar{F}_r(t) p(t) r_2(t) dt,$$

$$E\left(\int_0^{X_T} c_{2,r} e^{-\alpha t} d(N_{2,r}(t))\right) = c_{2,r} \int_0^T e^{-\alpha t} \bar{F}_k(t) \bar{F}_M(t) \bar{F}_r(t) (1-p(t)) r_2(t) dt,$$

and

$$E\left(\int_0^{X_T} c_3 e^{-\alpha t} d(N_3(t))\right) = c_3 \int_0^T e^{-\alpha t} \bar{F}_k(t) \bar{F}_M(t) \bar{F}_r(t) r_3(t) dt.$$

The expected cost for preventive replace (replace in time T) is $c_r e^{-\alpha T} \bar{F}_k(T) \bar{F}_M(T) \bar{F}_r(T)$. Let $C(T)$ be the total expected discounted cost for one cycle i.e. $C(T) = \dfrac{E(C(\alpha))}{E(1-e^{-\alpha T})}$. Then, by the use of the above results we have

$$\begin{aligned} E(1-e^{-\alpha X_T})C(T) &= c_1 \int_0^T e^{-\alpha t} \bar{F}_k(t) \bar{F}_M(t) \bar{F}_r(t) r_1(t) dt + c_k \int_0^T e^{-\alpha t} \bar{F}_k(t) \bar{F}_M(t) \bar{F}_r(t) r_{1,k}(t) dt \\ &+ c_{2,m} \int_0^T e^{-\alpha t} \bar{F}_k(t) \bar{F}_M(t) \bar{F}_r(t) p(t) r_2(t) dt \\ &+ c_{2,r} \int_0^T e^{-\alpha t} \bar{F}_k(t) \bar{F}_M(t) \bar{F}_r(t) (1-p(t)) r_2(t) dt \\ &+ c_3 \int_0^T e^{-\alpha t} \bar{F}_k(t) \bar{F}_M(t) \bar{F}_r(t) r_3(t) dt + c_r e^{-\alpha T} \bar{F}_k(T) \bar{F}_M(T) \bar{F}_r(T). \end{aligned} \tag{A.2}$$

Using (A.1), we have

$$E(1-e^{-\alpha X_T}) = \int_0^T \alpha e^{-\alpha t} \bar{F}_k(t) \bar{F}_M(t) \bar{F}_r(t) dt. \tag{A.3}$$

The right hand side of (**??**) can be reexpressed as

$$\int_0^T Q(t) dt + 2c_r, \tag{A.4}$$

where

$$\begin{aligned}Q(t) =&(c_1r_1(t) + c_kr_{1,k}(t) + c_{2,r}r_2(t)(1-p(t)) + c_{2,m}p(t)r_2(t) + c_3r_3(t)\\ &+ c_r(r_2(t)(1-p(t)) + r_3(t) - \alpha))\bar{F}_k(t)\bar{F}_M(t)\bar{F}_r(t)\\ &- (c_r(f_k(t) + 2r_2(t)(1-p(t)) + 2r_3(t)))\bar{F}_M(t)\bar{F}_r(t) - 2(c_rr_3(t))\bar{F}_k(t)\bar{F}_r(t)\\ &- 2(c_rr_2(t)(1-p(t)))\bar{F}_k(t)\bar{F}_M(t) + (-2f_k(t) + r_3(t))\bar{F}_r(t)\\ &+ 2(r_2(t)(1-p(t)) - f_k(t))\bar{F}_M(t) + 2f_k(t).\end{aligned}$$

By using (A.3) and (A.4), $C(T)$ can be rewritten as a simple expression

$$C(T) = \frac{\int_0^T Q(t)dt + 2c_r}{\int_0^T \alpha e^{-\alpha t}\bar{F}_k(t)\bar{F}_M(t)\bar{F}_r(t)dt} = \frac{\int_0^T Q(t)dt + 2c_r}{\int_0^T A(t)dt}. \tag{A.5}$$

In order to find T^* so that $C(T)$ should be minimized at $T = T^*$, we derive

$$\frac{d}{dT}C(T) = \frac{Q(T)\int_0^T A(t)dt - A(T)\left(\int_0^T Q(t)dt + 2c_r\right)}{\left(\int_0^T A(t)dt\right)^2}.$$

Take

$$h(T) = Q(T)\int_0^T A(t)dt - A(T)\left(\int_0^T Q(t)dt + 2c_r\right). \tag{A.6}$$

Therefore $\frac{d}{dT}C(T)$ equals 0, is greater than 0, or is less than 0, if the function in (A.6) equals 0, is greater than 0, or is less than 0, respectively. Hence, an optimal T^* satisfies $h(T^*) = 0$. On the other hand

$$\lim_{T\to 0} h(T) = -2c_r < 0,$$

and

$$\frac{d}{dT}h(T) = \left(\int_0^T A(t)dt\right)\frac{d}{dT}Q(T) - \left(\int_0^T Q(t)dt + 2c_r\right)\frac{d}{dT}A(T). \tag{A.7}$$

Since $A(T) = \int_0^T \alpha e^{-\alpha t}\bar{F}_k(t)\bar{F}_M(t)\bar{F}_r(t)dt$ is decreasing in T, thus, if $Q(T)$ is increasing then from (A.7) $h(T)$ is increasing. From (A.6) $h(T)$ can be re-expressed as

$$h(T) = (Q(T) - C(T)A(T))\int_0^T A(t)dt. \tag{A.8}$$

By using (A.8), we have the next result.

Proposition 3.1.

1. *If* $Q(\infty) > C(\infty)A(\infty)$ *then* $h(\infty) > 0$ *and there exist finite* T^* *that* $\frac{d}{dT}C(T)\big|_{T=T^*} = 0$ *and* T^* *minimizing* $C(T)$.
2. *If* $Q(t)$ *is strictly increasing then* $h(t)$ *is strictly increasing and* $Q(\infty) > C(\infty)A(\infty)$, *there exists a unique and finite minimum* T^*.
3. *If* $Q(t)$ *is non-decreasing then* $h(t)$ *is non-decreasing and* $Q(\infty) < C(\infty)A(\infty)$, *then* $T = \infty$ *is optimal.*

4 Numerical Example

In order to illustrate the obtained results, let us take $r_i(t) = \lambda_i\beta_i(\lambda_i t)^{\beta_i-1}$, $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3, \beta_1 = 3, \beta_2 = 2, \beta_3 = 2, \alpha = 0.1$ and $p(t) = \frac{1}{1+t^2}$. Also, suppose $c_1 = 3, c_k = 12, c_{2,m} = 15, c_{2,r} = 25, c_3 = 27, c_r = 10$ and $k = 5$. Then from assumptions, the cumulative failure intensity functions are $\lambda_1(t) = t^3, \lambda_2(t) = 4t^2$ and $\lambda_3(t) = 9t^2$. By calculating $h(T)$ in (A.6) and solve $h(T) = 0$, we can find the optimum T. The optimum T that minimize $C(T)$ is $T^* = 1.44$ and $C(T^*) = 1085.7$. Therefore, if the system doesn't have any failure need to replace until $T = 1.44$, then the system must do planned replacement in $T = 1.44$ in order to minimize the total expected discounted cost. To investigate the effect of c_r and k on optimal time for planned replacement, Figures 1 and 2 are plotted. In Figure 1 cost function is plotted versus T per different c_r. In this figure the change of optimal point of T, which is T^*, can be seen per different c_r's. It is also seen that T^* increases with planned replacement cost increases. That is, when c_r increases, planned replacement should be done in a later time.

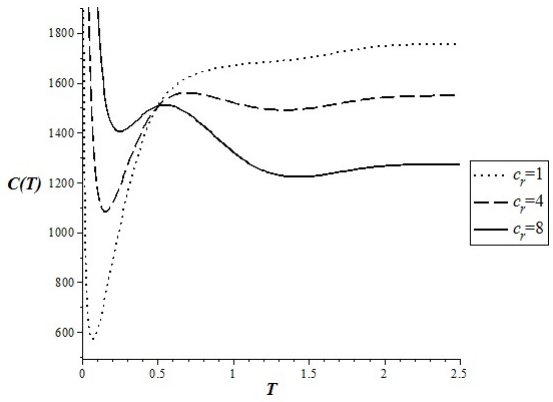


Figure 1: Cost function versus T per different c_r.

In Figure 2 cost function is plotted versus T per different k. In this figure the change of optimal point of T, which is T^*, can be seen per different k's. It is also seen that T^* increases with number of minimal repair before replacement in type I failure increases. That is, when k increases, planned replacement should be done in a later time. This figure shows that cost function decreasing with number of minimal repair before replacement in type I failure increases. That is, when k increases,

costs incurred to system decreases. Summing-up, the results observed from Figures 1 and 2 are as we intuitively expected.

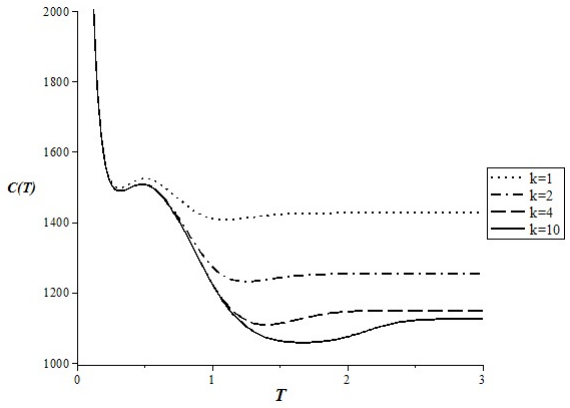


Figure 2: Cost function versus T per different k.

References

[1] Aven, T. and Castro, I.T (2008). A minimal repair replacement model with two types of failure and a safety constraint. *European Journal of Operational Research*, **188**, 506-515.

[2] Aven, T. and Jensen, U. (1999). *Stochastic Models in Reliability*, Springer-Verlag, London.

[3] Barlow, R. E. and Proschan, F. (1965). *Mathematical Theory of Reliability*, John Wiley and Sons, New York.

[4] Lai, M. and Chen, S. (2016). A bivariate optimal replacement policy with cumulative repair cost limit under cumulative damage model. *Sadhana*, **41**, 497-505.

[5] Nakagawa, T. and Kowada, M. (1983). Analysis of a system with minimal repair and its application to replacement policy. *European Journal of Operation Research*, **12**, 176-182.

[6] Sheu, S., Chang, C. C. (2009). An extended periodic preventive maintenance model with age-dependent failure type. *IEEE Transformation on Reliability*, **85**, 397-405.

[7] Sheu, S., Chang, C. C., Chen, Y. and Zhang, Z. (2015). Optimal preventive maintenance and repair policies for multi-state systems. *Reliability Engineering and System Safety*, **140**, 78-87.

[8] Boland, P. J. and Proschan, F. (1982). Periodic replacement with increasing minimal repair costs at failure. *Operations Research*, **30**, 1183-1189.

[9] Nakagawa, T. (1981). Generalized models for determining optimal number of minimal repairs before replacement. *European Journal of Operation Research*, **24**, 325-337.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Reliability of random complex networks based on percolation theory

Sajadi, F. [1]

Department of Statistics, University of Isfahan

Abstract

In this paper we evaluate the reliability of a complex network using concepts of percolation theory by regarding the network failure as a percolation process. The critical threshold of percolation can be used as network failure criterion. In particular we consider complex systems which can be modeled as a general inhomogeneous random graphs.

Keywords: Complex network, Random graph, Percolation threshold, Network reliability.

1 Introduction

Complex networks can be used to model different systems of the real world where nodes represent the elements of the system and edges/links depict the interactions patterns between them. The performance of such systems depends on the interconnection between the components of the system. An important property of such connection patterns is their robustness. The focus of robustness in complex networks is the response of the network to the removal of nodes or links. The mathematical model of such a process can be thought of as a percolation process. Percolation is a model of statistical physics. It models the process of randomly removing either nodes (site percolation) or edges (bond percolation) that are randomly occupied with probability p, from a graph. In percolation, the random network considered naturally lives on an infinite graph. A common way to evaluate the robustness of a graph is that of studying the size of the giant

[1] f.sajadi@sci.ui.ac.ir

component after the removal of a fraction of nodes/edges. By removing nodes/edges, this size decreases. Using percolation theory, one can predicts the presence of a threshold value, above it the giant component disappears, i.e. the network is fragmented in many disconnected components of very small sizes. Any infinite graph has a site percolation critical probability and a bond percolation critical probability.

In network reliability, many studies focused on the terminal reliability between pair of nodes in the network. In these studies, only the connectivity between two nodes, namely the origin and the destination are considered. The removal of a single node has only limited impact on a network's integrity. The removal of several nodes, however, can break a network into several isolated components. Obviously, the more nodes we remove, the higher are the chances that we damage a network, prompting us to ask: what fraction of nodes must be removed so that the network turns into smaller clusters that are unable to communicate with each other? Or equivalently we can ask with how many operating components, still a system works.

In this paper, we consider a random graph model for a system of components. We use a statistical physics description of the failure process for this system and based on percolation theory, we calculate the reliability of network. In [4], the authors considered the same node-reliability for each node in the random graph and based on site percolation, computed the network reliability at time t. In fact they considered the same lifetime probability for nodes. Here we assume that the lifetime of edges are not identical and we use the results of bond percolation theory to compute the global network reliability.

2 Preliminary Definitions

This paper deals with a mathematical concept related to percolation theory known as percolation threshold. Before we address percolation threshold, however, some definitions of basic concepts in random graph theory will be necessary. We give some definitions of random graph-theoretical terms used throughout the paper.

Definition 1. *Let G be a graph with vertex set $[n] = \{1, 2, ..., n\}$. Given a real number p, $0 \leq p \leq 1$, the binomial random graph, denoted by $\mathcal{G}_{n,p}$, is defined by taking as Ω the set of all graphs on vertex set $[n]$ and setting*

$$\mathbb{P}(G) = p^{e_G}(1-p)^{\binom{n}{2}-e_G},$$

where e_G stands for the number of edges of G.

Definition 2. *Given an integer $M, 0 \leq M \leq \binom{n}{2}$, the uniform random graph, denoted by $\mathcal{G}_{n,M}$, is defined by taking as Ω the family of all graphs on the vertex set $[n]$ with exactly M edges, and as $\mathbb{P}$ the uniform probability on Ω,*

$$\mathbb{P}(G) = \binom{\binom{n}{2}}{M}^{-1} \quad G \in \Omega$$

Definition 3. *A general inhomogeneous random graph $\mathcal{G}_{n,\boldsymbol{p}}$, where $\boldsymbol{p} = \{p_{ij}\}_{1 \leq i < j \leq n}$ is such that p_{ij} is the probability that the edge $ij = (i, j)$ is occupied, and where different edges are independent.*

If $p_{ij} = p$ for all ij, then we get the Erdos-Renyi random graph $\mathcal{G}_{n,p}$. Note that the distributions of $\mathcal{G}_{n,p}$ and $\cup_{i=1}^{n}\mathcal{G}_{n,p_i}$ are identical, (Staged exposure,[5]).

Definition 4. *Let Γ be a finite set, $|\Gamma| = N$, let $0 \leq p \leq 1$. Then the random subset Γ_p of Γ is obtained by flipping a coin, with probability p of success, for each element of Γ to determine whether the element is to be included in Γ_p; the distribution of Γ_p is the probability distribution on $\Omega = 2^{\Gamma}$ given by $\mathbb{P}(F) = p^{|F|}(l-p)^{|\Gamma|-|F|}$ for $F \subseteq \Gamma$.*

For an arbitrary set X and an integer k, let $[X]^k$ stand for the family of all k-element subsets of X. If $X = \{1, 2, ..., n\}$, we will simplify this notation to $[n]^k$. Taking $\Gamma = [n]^2$, we obtain the model of random graphs $\mathcal{G}_{n,p}$, [5].

Definition 5. *A family of subsets $\mathcal{Q} \subseteq 2^{\Gamma}$ is called increasing if $A \subseteq B$ and $A \in \mathcal{Q}$ imply that $B \in \mathcal{Q}$.*

For example "being connected", is an increasing graph property. For many graph properties the limiting probability that a random graph possesses them jumps from 0 to 1 (or vice versa) very rapidly, that is, with a rather small increase in the (expected) number of edges. It is known that every monotone property has a threshold, [5].

Definition 6. *For an increasing property $\mathcal{Q}$, a sequence $\hat{p} = \hat{p}(n)$ is called a threshold if $\mathbb{P}(\Gamma_p \in \mathcal{Q})$ converges to 1 for $p >> \hat{p}$ and converges to 0 for $p << \hat{p}$.*

It is easy to show that $\hat{p}$ is a threshold for a monotone property if and only if $\hat{M} = \hat{p}|\Gamma|$.

Definition 7. *Denote by $C(v)$ the set of nodes which can be reached from a node v by occupied edges. Define the percolation function $p \longrightarrow \theta(p)$ by*

$$\theta(p) = \mathbb{P}_p(|C(v)| = \infty)\,.$$

The critical value is

$$p_c = \inf\{p : \theta(p) > 0\}\,.$$

It is known that for any infinite graph, the number of infinite connected components is in $\{0, 1, \infty\}$ a.s., [2].

3 Reliability of random network

The correct function of a real complex network, relies on its structural properties: the resilience to damage and the robustness to external attacks, cascade failures, or to collapses due to traffic jams and overloading, [1]. Let $G(V, E)$ be an undirected graph on n nodes. Let $R_s(t)$ be the network reliability of the system. In this section by considering the lifetimes of network edges, we study the global network lifetime with respect to the critical point of the network percolation process. Given the time t, let $R_{ij}(t)$ be the probability that edge between nodes i and j is functional

at time t, (the reliability of edge i, j at time t). We ignore the direction of the edges, therefore $R_i(t) = R_j(t) = R_{ij}(t)$. Let $1 - p = 1 - R(t)$ where

$$R(t) = 1 - \prod_{i=1}^{n} [1 - R_i(t)] \ .$$

According to the reliability of edges, a fraction $1-p$ of them fail and as the failure process proceeds, cluster of connected nodes will form as they are fall of from the giant cluster. When the number of failed edges increases, the network breaks into small clusters. If $R(t)$ is less than a critical point p_c, the giant cluster does not exist anymore. Define the instant at which this occurs as the lifetime of the network.

It is known that [5], the binomial (uniform) random graph loses its connectivity when the number of failed nodes reaches $n-[n*p_c]$ (when the number of edges in the uniform random graph reaches $\binom{n}{2} - [M^*]$, where M^* is the number of remaining edges at critical point).

Now the network reliability at time t, $R_s(t)$ is given by the following equation:

$$R_s(t) = \sum_{i=[M^*]+1}^{N} \sum_{A \in [N]^i} \prod_{j \in A} R_j(t) \prod_{j \in A^c} [1 - R_j(t)] \,, \tag{A.1}$$

where $N = \binom{n}{2}$. In case when, all of $R_i(t)$'s are equal to R(t), we have

$$R_s(t) = \sum_{i=[M^*]+1}^{N} \binom{N}{i} R(t)^i [1 - R(t)]^{N-i} \,, \tag{A.2}$$

The lifetime distribution of the network based on $R_s(t)$ is

$$f_s(t) = \frac{d(1 - R_s(t))}{dt} \,,$$

and the network lifetime, T_s, is given by

$$T_s = \int_0^\infty R_s(t) dt$$

In the following, since the node percolation and edge percolation have the same properties, we consider node percolation for simplicity and give an example. Let $R_i(t)$ be the probability that node i is functional at time t, $i = 1, 2, ..., n$. The expected minimal number of functional nodes to keep the network connected is $[n * p_c]$. Hence,

$$R_s(t) = \sum_{i=[n*p_c]+1}^{n} \sum_{A \in [n]^i} \prod_{j \in A} R_j(t) \prod_{j \in A^c} [1 - R_j(t)] \,. \tag{A.3}$$

The computing of $R_s(t)$ in (A.1) by enumerating all elements in $[n]^i$ is not practicable, even when n is small. Thus we need to use approximation methods for computing $R_s(t)$. One can think

of Poisson approximation method, which is used for Poisson binomial distribution(the distribution of the sum of independent and non-identical random indicators). Fix $t > 0$. Let $p_i(t) = R_i(t)$, $\mu(t) = \sum_{i=1}^{n} p_i(t)$. Then by Poisson approximation method,

$$R_s(t) \approx \sum_{i=[n*p_c]+1}^{n} \frac{\mu(t)^i \mathrm{e}^{-\mu(t)}}{i!}. \tag{A.4}$$

By Le Cam's theorem [3], the approximation error for the Poisson approximation method is

$$\sum_{k=0}^{n} | \sum_{A\in[n]^k} \prod_{j\in A} R_j(t) \prod_{j\in A^c} [1 - R_j(t)] - \frac{\mu(t)^k \mathrm{e}^{-\mu(t)}}{k!} | < 2 \sum_{j=1}^{n} p_j(t)^2$$

Another approximation method is the normal approximation method which is based on the central limit theorem.

Example 3.1. Suppose lifetime distribution of node $i, i = 1, 2, ..., n$ follows an Exponential distribution with mean $\frac{1}{\lambda_i}$, i.e.,

$$R_i(t) = \mathrm{e}^{-\lambda_i t}.$$

Denote the average value of degree (sum of node's degree divided by the number of nodes in the network) by $< k >$. For a random network, the percolation threshold can be calculated as $p_c = \frac{1}{<k>}$. Therefore,

$$R_s(t) = \sum_{i=[n*p_c]+1}^{n} \sum_{A\in[n]^i} \prod_{j\in A} \mathrm{e}^{-\lambda_j t} \prod_{j\in A^c} [1 - \mathrm{e}^{-\lambda_j t}]. \tag{A.5}$$

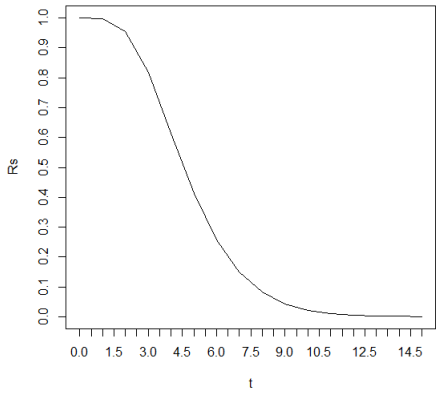
If $\lambda_j = \lambda$, for all $j = 1, ..., n$, then

$$R_s(t) = \sum_{i=[n*p_c]+1}^{n} \mathrm{e}^{-i\lambda t} [1 - \mathrm{e}^{-\lambda t}]^{n-i}. \tag{A.6}$$

Let $n = 8$, $< k >= 4$ and $\lambda_i \sim U(0.249, 0.251), i = 1, 2, ..., 8$. The plot of $R_s(t)$ verses $t, 0 \le t \le 15$ is shown in the following figure.

4 Future Works

It would be interesting to study the case when the edges are not independent of each other, i.e., the presence of an edge in one position is correlated with the presence of an edge in a different position. For that we will study the random geometric graphs. Also, in future, we will study the failure rate of system based on failure rate of components, using percolation threshold for underlying graph and for that a comparison between theoretical and simulation results will be considered.



References

[1] Callaway,D. S., Newman, M. E. J., Strogatz, S. H. and Watts, D. J., (2000), Network robustness and fragility: percolation on random graphs, *Phys. Rev. Lett.* . **85**, 5468-5471.

[2] Grimmett, G. (1999) , *Percolation*, vol. 321 of Grundlehren der Mathematischen Wissenschaften, Springer-Verlag, Berlin, second ed., .

[3] Le Cam, L. ,(1960), An approximation theorem for the Poisson binomial distribution, *Pacific Journal of Mathematics*. **10**, 1181-1197.

[4] Li, D., Zhang, Q., Zio, E., Havlin, S. and Kang, R., (2015), Network reliability analysis based on percolation theory, Reliability. *Reliability Engineering and System Safety*. **142**.

[5] Luczak, T. and Rucinski, A. (2000), *Random Graphs*, Wiley, New York. .

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Non-parametric estimation of bivariate MRL and QRL functions

Shafaei Noughabi, M. [1]

Department of Mathematics and Statistics, University of Gonabad

Abstract

When we deal with dependent pairs of lifetimes, estimation of bivariate mean residual life (BMRL) and bivariate α-quantile residual life (α-BQRL) functions may play a key role in survival analysis. We present the empirical non-parametric estimators of them based on complete data and discuss their main asymptotic properties. We have a real data example to illustrate their applications.

Keywords: Bivariate mean residual life , Bivariate α-quantile residual life, Empirical estimator.

1 Introduction

Many statisticians studied the problem of estimating the BMRL function. Among them we can refer to Zahedi (1982), Kulkarni and Rattihalli (2002) and Rojo and Ghebremichael (2006). However, the BMRL may be useful in survival analysis, but it may encounter some shortcomings, e.g., when the distribution is strongly skewed or heavy-tailed in which a long-term survivor may have a marked effect on the estimator. This fact that the list of common heavy-tailed distribution includes Pareto, Weibull, log-normal and log-gamma distributions supports this idea. To overcome such shortcomings, we can consider the α-BQRL function as a good competitor for BMRL function. Here, we review a non-parametric estimator of the BMRL function and it's key attributes briefly. Then, a non-parametric estimator for α-BQRL has been proposed and studied. Finally, two aforementioned measures have been computed for a real bivariate lifetime data set.

[1] mohamad.shafaee@gmail.com

2 Bivariate mean residual life function

Let the non-negative random vector $\mathbf{X} = (X_1, X_2)$ represent lifetimes of two possibly dependent components and it follows absolutely continuous distribution F in the first quadrant of $\mathbb{R}^2$, $Q = \{(x_1, x_2); x_i \geq 0\}$. The well-known reliability function is $\bar{F}(x_1, x_2) = P(X_1 > x_1, X_2 > x_2)$. Nair and Nair (1989) showed that the bivariate mean residual life function

$$m_i(x_1, x_2) = \begin{cases} \frac{1}{\bar{F}(x_1,x_2)} \int_{x_1}^{\infty} \bar{F}(u, x_2)du, & i = 1, \\ \frac{1}{\bar{F}(x_1,x_2)} \int_{x_2}^{\infty} \bar{F}(x_1, u)du, & i = 2, \end{cases}$$

characterizes the underlying distribution uniquely.
Let (X_{1i}, X_{2i}), $i = 1, 2$ be n independent and identically distributed pairs of random lifetimes following reliability function $\bar{F}$. Kulkarni and Rattihalli (2002) proposed the empirical estimator of $\big(m_1(x_1, x_2), m_2(x_1, x_2)\big)$ measure by

$$\hat{m}_i(x_1, x_2) = \begin{cases} \frac{\sum_{i=1}^{n}(X_{1i}-x_1)I[X_{1i}>x_1, X_{2i}>x_2]}{\sum_{i=1}^{n} I[X_{1i}>x_1, X_{2i}>x_2]}, & i = 1, \\ \\ \frac{\sum_{i=1}^{n}(X_{2i}-x_2)I[X_{1i}>x_1, X_{2i}>x_2]}{\sum_{i=1}^{n} I[X_{1i}>x_1, X_{2i}>x_2]}, & i = 2, \end{cases} \tag{A.1}$$

Let $||(x, y)|| = \max\{|x|, |y|\}$. The following theorem holds for every (b_1, b_2) with $\bar{F}(b_1, b_2) > 0$ and $D = [0, b_1] \times [0, b_2]$ for such points.

Theorem 2.1. *The estimator $(\hat{m}_1, \hat{m}_2)$ is asymptotically unbiased, and uniformly strongly consistent on D, i.e.,*

$$\sup_D ||(\hat{m}_1, \hat{m}_2) - (m_1, m_2)|| \to 0.$$

They showed that $(\hat{m}_1, \hat{m}_2)$ weakly converges to a zero-mean bivariate gaussian process.
Rojo and Ghebremichael (2006) studied the estimation of the bivariate mean residual life when it is bounded above by another known or unknown ones.

3 Bivariate quantile residual life function

Recently Shafaei and Kayid (2017) proposed the α-BQRL function as a vector $\big(q_{1,\alpha}(x_1, x_2), q_{2,\alpha}(x_1, x_2)\big)$ where

$$q_{i,\alpha}(x_1, x_2) = \bar{F}_i^{-1}(\bar{\alpha}\bar{F}(x_1, x_2); x_{3-i}) - x_i, \quad i = 1, 2, \tag{A.1}$$

in which $\bar{F}_1^{-1}(p; x_2) = \inf\{x_1 : \ \bar{F}(x_1, x_2) \leq p\}$ and $\bar{\alpha} = 1 - \alpha$. $\bar{F}_2^{-1}(p; x_1)$ is defined in a similar way.
Applying the empirical estimator of $\bar{F}$, a non-parametric estimator of it can be introduced as

$$\begin{cases} \hat{q}_{1,\alpha}(\mathbf{x}) = \hat{\bar{F}}_1^{-1}(\bar{\alpha}\hat{\bar{F}}(\mathbf{x}); x_2) - x_1, \\ \hat{q}_{2,\alpha}(\mathbf{x}) = \hat{\bar{F}}_2^{-1}(\bar{\alpha}\hat{\bar{F}}(\mathbf{x}); x_1) - x_2, \end{cases}$$

Table 1: Survival times in months.

Patient (i)	1	2	3	4	5	6	7	8	9
X_{1i}	6.9	1.63	13.83	35.53	14.8	6.2	22	1.7	43.03
X_{2i}	20.17	10.27	5.67	5.90	33.9	1.73	30.2	1.7	1.77
Patient (i)	10	11	12	13	14	15	16	17	18
X_{1i}	6.53	42.17	48.43	9.6	7.6	1.8	9.9	13.77	0.83
X_{2i}	18.7	42.17	14.3	13.33	14.27	34.57	21.57	13.77	10.33
Patient (i)	19	20	21	22	23	24			
X_{1i}	1.97	11.3	30.4	19	5.43	46.63			
X_{2i}	11.07	2.1	13.97	13.80	13.57	42.43			

where

$$\hat{\bar{F}}(\mathbf{x}) = n^{-1}\sum_{i=1}^{n} I(X_{1i} > x_1, X_{2i} > x_2),$$

is the empirical reliability function and

$$\hat{\bar{F}}_1^{-1}(p; x_2) = \inf\{x : \ \hat{\bar{F}}(x, x_2) \leq p\}, \quad p \leq \hat{\bar{F}}(0, x_2),$$

shows the inverse of the empirical reliability. $I(X_{1i} > x_1, X_{2i} > x_2)$ is 1 when $X_{1i} > x_1, X_{2i} > x_2$ and 0 otherwise. $\hat{\bar{F}}_2^{-1}(p; x_1)$ can be defined analogously.

Theorem 3.1. *If $\bar{F}_1^{-1}(p; x_2)$ and $\bar{F}_2^{-1}(p; x_1)$ be continuous at $p = \bar{\alpha}\bar{F}(x_1, x_2)$ then the estimator $(\hat{q}_{1,\alpha}, \hat{q}_{2,\alpha})$ is strongly consistent, i.e.,*

$$\Big(\hat{q}_{1,\alpha}(x_1, x_2), \hat{q}_{2,\alpha}(x_1, x_2)\Big)^T \to \Big(q_{1,\alpha}(x_1, x_2), q_{2,\alpha}(x_1, x_2)\Big)^T, \quad a.e.,$$

Recently, Shafaei and Kayid (2017) investigated its weakly convergence to a bivariate gaussian process under proper normalization.

4 Real data example

Here we consider a data set related to a diabetic retinopathy (DR) study reported from Rojo and Ghebremichael (2006). DR is a major cause of visual loss in many countries and the leading cause of blindness in patients under 60 years old in the USA. One main goal is to study the effect of laser photocoagulation in delaying blindness in DR patients. Some patients with DR at both eyes and at least 20/100 clarity of vision in both eyes were taken part in the study. One eye of each patient was randomly selected to receive treatment. The interesting event for each eye is the time from initiation of treatment to the time when visual acuity dropped below 5/200 two visits in a row. Such low vision is called "blindness". The survival times have been presented in Table 1. These

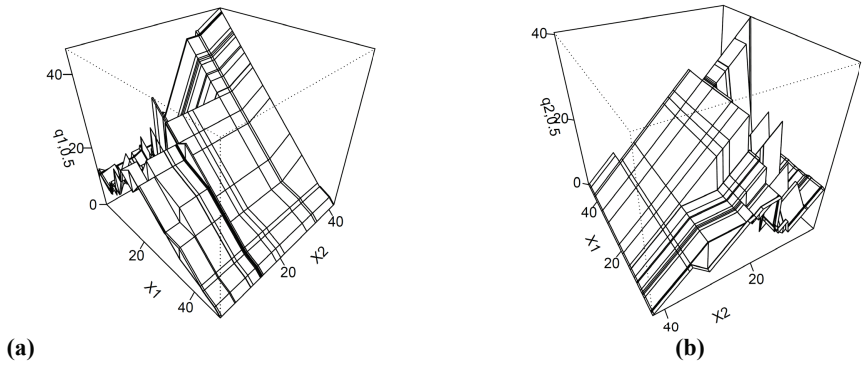


Figure 1: (a): Surface plot of $\hat{q}_{1,0.5}(x_1, x_2)$ for the data set of Table 1. (b): Surface plot of $\hat{q}_{2,0.5}(x_1, x_2)$ for this data set.

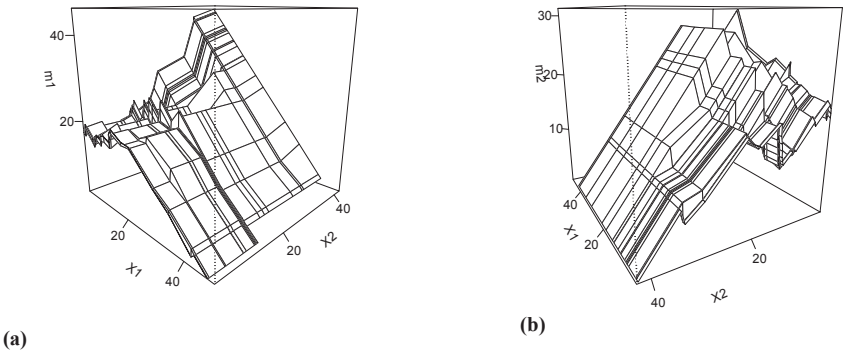


Figure 2: (a): Surface plot of $\hat{m}_1(x_1, x_2)$ for the data set of Table 1. (b): Surface plot of $\hat{m}_2(x_1, x_2)$ for this data set.

survival times are time to blindness minus 6.5 months which is minimum possible time to events. Here, X_{1i} and X_{2i} measure time to event for treated and untreated eyes respectively.
Figures 1 (a) and (b) present surface plots of $\hat{q}_{1,0.5}(x_1, x_2)$ and $\hat{q}_{2,0.5}(x_1, x_2)$ respectively. Moreover, surface plots related to $\hat{m}_1(x_1, x_2)$ and $\hat{m}_2(x_1, x_2)$ have been drawn in Figures 2 (a) and (b) respectively. Both $\hat{q}_{1,0.5}(x_1, x_2)$ and $\hat{m}_1(x_1, x_2)$ show decreasing in x_1 except for small values of x_1. Similarly, $\hat{q}_{2,0.5}(x_1, x_2)$ and $\hat{m}_2(x_1, x_2)$ are decreasing in x_2 except small values of x_2.

References

[1] Kulkarni, H. V., and Rattihalli, R. N., (2002). Nonparametric Estimation of a Bivariate Mean Residual Life Function. *Journal of the American Statistical Association*, **97**, 907-917.

[2] Nair, K. R. and Nair, N. U. (1989). Bivariate mean residual life. *IEEE Transactions on Reliability*, **38**, 362-364.

[3] Rojo, J. and Ghebremichael, M. (2006). Estimation of two ordered bivariate mean residual life functions. *Journal of Multivariate Analysis*, **97**, 431-454.

[4] Shafaei N. M. and M. Kayid (2017). Bivariate quantile residual life: a characterization theorem and statistical properties. *Statistical Papers, to appear.*

[5] Shafaei N. M. and M. Kayid (2017). Further results involving the bivariate quantile residual life model with applications. *Submitted for possible publications.*

[6] Zahedi, H. (1982). Multivariate mean remaining life functions, *Ph.D. Dissertation,* University of California, Riverside, CA.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Some results on comparing the mean residual life functions of two groups

Sharafi, M. [1]

Department of Statistics, Razi University

Abstract

The mean residual life (MRL) function or remaining life expectancy function at age t is defined to be the expected remaining life given survival to age t. In this paper, we first propose two new tests for comparing mean residual life functions of two populations. We also obtain some results on limit theorems under null and fixed hypothesis. Finally, some examples is presented for illustration and is discussed about comparing powers of proposed test statistics.

Keywords: Asymptotic distribution, Brownian motion process, Convergence, Mean residual order, Nonparametric test.

1 Introduction

The Mean Residual Life function (MRL) of a subject given that the subject has survived uo to a given time point. This concept play an important and a signican role in reliability as well as insurance studies. For a detailed discussion and statistical applications of the MRL function and and empricall process, you can refer to [1,3,4,5] and [2,6] respectively . In this paper, we interest to compare the mean residual life functions from two populations or treatment groups. Suppose that F and G are two distribution functions on $\mathbb{R}^+$ with finite means μ_F and μ_G, and let e_F and e_G denote the corresponding mean residual life functions: i.e.

$$e_F(x) = E\{X - x | X > x\} = \frac{\int_x^\infty \overline{F}(y)dy}{\overline{F}(x)}$$

[1] m.sharafi@razi.ac.ir

where $\overline{F}(x) \equiv 1 - F(x)$. Consider testing $H: \ e_F = e_G$ versus $K: \ e_G(t) \geq e_F(t)$ for all $t \in \mathbb{R}^+$ and with strict inequality for some $t > 0$. Note that K holds if and only if

$$\delta_{F,G}(t) \equiv \delta(t) \equiv \overline{F}(t) \int_t^\infty \overline{G}(v)dv - \overline{G}(t) \int_t^\infty \overline{F}(v)dv \geq 0 \tag{A.1}$$

for all $t \geq 0$, and with strict inequality for some t. Note that taking $t = 0$ yields $\mu_G \geq \mu_F$. We also note that we can write

$$\begin{aligned} \delta_{F,G}(t) &= \overline{F}(t)\overline{G}(t)e_G(t) - \overline{F}(t)\overline{G}(t)e_F(t) \\ &= \overline{F}(t)\overline{G}(t)\left\{e_G(t) - e_F(t)\right\}. \end{aligned} \tag{A.2}$$

With $H_\lambda(x) \equiv \lambda F(x) + (1-\lambda)G(x)$ for some $\lambda \in (0,1)$ we define

$$\Delta_\lambda(F,G) \equiv \int_0^\infty \delta_{F,G}(t)dH_\lambda(t). \tag{A.3}$$

Note that $\Delta_\lambda(F,G) \geq 0$ with equality holding under the null hypothesis H (or $F = G$), while $\Delta_\lambda(F,G) > 0$ if the set of t's for which strict inequality holds in (A.1) has positive H_λ measure.

2 Some statistics

2.1 Integral type test statistics

Suppose that we observe $X_1, \ldots, X_m$ i.i.d. F and $Y_1, \ldots, Y_n$ i.i.d. G, and let $\mathbb{F}_m$ and $\mathbb{G}_n$ denote the empirical distribution functions of the X's and Y's respectively. We also let $N \equiv m/N \equiv \lambda_N$ and let $\mathbb{H}_N \equiv \lambda_N \mathbb{F}_m + (1-\lambda_N)\mathbb{G}_n$ denote the pooled empirical measure. Based on the development leading to (A.2) and (A.3) it is natural to define

$$\hat{\delta}_{m,n}(t) \equiv \overline{\mathbb{F}}_m(t)\overline{\mathbb{G}}_n(t)\left\{\hat{e}_{\mathbb{G}_n}(t) - \hat{e}_{\mathbb{F}_m}(t)\right\},$$

and to consider integral type test statistics of the form

$$T_{m,n} \equiv \sqrt{\frac{mn}{N}} \int_0^\infty \hat{\delta}_{m,n}(t)d\mathbb{H}_N(t) = \int_0^\infty \mathbb{Z}_{m,n}(t)d\mathbb{H}_N(t) \tag{A.1}$$

where

$$\mathbb{Z}_{m,n}(t) \equiv \sqrt{\frac{mn}{N}}(\hat{\delta}_{m,n}(t) - \delta(t)) \tag{A.2}$$

and the second equality in (A.1) holds under the null hypothesis H.

We will show in Section 3 that under the null hypothesis $e_G = e_F$ we have

$$\mathbb{Z}_{m,n}(t) \Rightarrow \overline{F}(t)\mathbb{B}(\overline{F}(t)\sigma_F^2(t)).$$

Thus the processes $\overline{\mathbb{Z}}_{m,n} \equiv \mathbb{Z}_{m,n}/\overline{\mathbb{H}}_N$ satisfy (modulo some possible difficulties for small values of $\overline{F}$)

$$\overline{\mathbb{Z}}_{m,n}(t) \Rightarrow \mathbb{B}(\overline{F}(t)\sigma_F^2(t)). \tag{A.3}$$

This leads to integral type statistics based on the processes $\overline{\mathbb{Z}}_{m,n}$: we define

$$\overline{T}_{m,n} \equiv \int_0^\infty \overline{\mathbb{Z}}_{m,n}(t) d\mathbb{H}_N(t).$$

3 Limit theory under the null hypothesis

To treat the asymptotic theory of the test statistics $T_{m,n}$, we first consider the sequence of processes $\{\mathbb{Z}_{m,n}(t) : 0 < t < \infty\}_{m,n=1}^\infty$ defined by A.2 Now suppose that the null hypothesis holds: thus $F = G$ and $e_F = e_G$. We will also assume that $m/N \to \lambda \in [0,1]$. These considerations lead naturally to the following theorem:

Theorem 3.1. *Suppose that the null hypothesis H holds (so that $e_F = e_G$ and $\delta_{F,G} = 0$). If $E_F(X^2) < \infty$ and $m \wedge n \to \infty$, then*

$$\mathbb{Z}_{m,n} \Rightarrow \mathbb{Z}_0 \quad \textit{in} \quad \ell^\infty(\mathbb{R}^+)$$

where

$$\mathbb{Z}_0(t) \equiv \overline{F}(t)\mathbb{B}(\overline{F}(t)\sigma_F^2(t))$$

for a standard Brownian motion process $\mathbb{B}$ and $\sigma_F^2(x) \equiv Var_F(X - x|X > x)$.

Now we let $H_\lambda = \lambda F + (1-\lambda)G$. Under the null hypothesis $H_\lambda = F$, and the above calculations suggest that

$$T_{m,n} \equiv \int_0^\infty \mathbb{Z}_{m,n}(t) dH_\lambda(t) \to_d \int_0^\infty \overline{F}(t)\mathbb{B}(\overline{F}(t)\sigma_F^2(t)) dF(t) \sim N(0, A^2)$$

where

$$\begin{aligned} A^2 &= \int_0^\infty \int_0^\infty \overline{F}(s)\overline{F}(t) \left\{ \overline{F}(t)\sigma_F^2(t) 1\{t \geq s\} + \overline{F}(s)\sigma_F^2(s) 1\{t \leq s\} \right\} dF(s) dF(t) \\ &= 2\int_0^\infty \left(\int_t^\infty \overline{F}^2(s)\sigma_F^2(s) dF(s) \right) \overline{F}(t) dF(t). \end{aligned} \tag{A.1}$$

Similar considerations apply to the statistics $\overline{T}_{m,n}$: integrating the processes $\overline{\mathbb{Z}}_{m,n}$ yields

$$\overline{T}_{m,n} \equiv \int_0^\infty \overline{\mathbb{Z}}_{m,n}(t) d\mathbb{H}_N(t).$$

Based on the developments above it is not hard to see that under the null hypothesis $e_G = e_F$ we should have

$$\overline{T}_{m,n} \to_d \int_0^\infty \mathbb{B}\left(\overline{F}(t)\sigma_F^2(t)\right) dF(t) \sim N(0, \overline{A}^2)$$

where now

$$\begin{aligned} \overline{A}^2 &= \int_0^\infty \int_0^\infty \{\overline{F}(t)\sigma_F^2(t)1\{t \geq s\} + \overline{F}(s)\sigma_F^2(s)1\{s \geq t\}\} dF(s)dF(t) \\ &= 2\int_0^\infty \left(\int_t^\infty \overline{F}(s)\sigma_F^2(s)dF(s)\right) dF(t). \end{aligned} \tag{A.2}$$

This leads to the following theorem concerning $T_{m,n}$ and $\overline{T}_{m,n}$ under the null hypothesis:

Theorem 3.2. *Suppose that the null hypothesis H holds. If $E_F(X^2) < \infty$ and $m \wedge n \to \infty$, then:*
A. The statistics $\{T_{m,n} : m, n \geq 1\}$ satisfy

$$T_{m,n} \equiv \int_0^\infty \mathbb{Z}_{m,n}(t)d\mathbb{H}_N(t) \to_d \int_0^\infty \mathbb{Z}_0(t)dF(t) \sim N(0, A^2)$$

where A^2 given by (A.1).
B. The statistics $\{\overline{T}_{m,n} : m, n \geq 1\}$ satisfy

$$\overline{T}_{m,n} \equiv \int_0^\infty \overline{\mathbb{Z}}_{m,n}(t)d\mathbb{H}_N(t) \to_d \int_0^\infty \overline{\mathbb{Z}}_0(t)dF(t) \sim N(0, \overline{A}^2)$$

where $\overline{A}^2$ is given by (A.2) and $\overline{\mathbb{Z}}_0(t) \stackrel{d}{=} \mathbb{B}(\overline{F}(t)\sigma_F^2(t))$.

Suppose that the null hypothesis H holds. If $E_F(X^2) < \infty$ and $m \wedge n \to \infty$, then:
A. The statistics $\{T_{m,n} : m, n \geq 1\}$ satisfy

$$T_{m,n} \equiv \int_0^\infty \mathbb{Z}_{m,n}(t)d\mathbb{H}_N(t) \to_d \int_0^\infty \mathbb{Z}_0(t)dF(t) \sim N(0, A^2)$$

where A^2 given by (A.1).
B. The statistics $\{\overline{T}_{m,n} : m, n \geq 1\}$ satisfy

$$\overline{T}_{m,n} \equiv \int_0^\infty \overline{\mathbb{Z}}_{m,n}(t)d\mathbb{H}_N(t) \to_d \int_0^\infty \overline{\mathbb{Z}}_0(t)dF(t) \sim N(0, \overline{A}^2)$$

where $\overline{A}^2$ is given by (A.2) and $\overline{\mathbb{Z}}_0(t) \stackrel{d}{=} \mathbb{B}(\overline{F}(t)\sigma_F^2(t))$.

4 Limit theory under a fixed alternative hypothesis, $e_F \neq e_G$

Here our goal is to prove the following theorem concerning the processes $\{\mathbb{Z}_{m,n} : m, n \geq 1\}$ under a fixed alternative, $e_G \neq e_F$.

Theorem 4. Suppose that F and G satisfy $E_F X^2 < \infty$ and $E_G Y^2 < \infty$, and assume that $\lambda_N \to \lambda \in [0,1]$ as $m \wedge n \to \infty$. Then: (a)

$$\mathbb{Z}_{m,n} \Rightarrow \mathbb{Z} \text{ in } \ell^\infty(\mathbb{R})$$

where $\mathbb{Z}$ is given by

$$\begin{aligned} \mathbb{Z}(t) &\equiv \sqrt{\lambda}\overline{F}(t)\mathbb{Z}_G(t)\overline{G}(t) - \sqrt{1-\lambda}\overline{G}(t)\mathbb{Z}_F(t)\overline{F}(t) \\ &\quad - (e_G - e_F)(t)\left\{\overline{G}(t)\sqrt{1-\lambda}\mathbb{U}(F(t)) + \overline{F}(t)\sqrt{\lambda}\tilde{\mathbb{U}}(G(t))\right\} \\ &\equiv \mathbb{Z}_1(t) - \mathbb{Z}_2(t). \end{aligned} \tag{A.1}$$

(b) Under the same hypotheses(?) and with $\overline{\mathbb{Z}}_{m,n} \equiv \mathbb{Z}_{m,n}/\overline{H}_N$,

$$\overline{\mathbb{Z}}_{m,n} \Rightarrow \overline{\mathbb{Z}} \text{ in } \ell^\infty(\mathbb{R})$$

where $\overline{\mathbb{Z}} \equiv \mathbb{Z}/\overline{H}_\lambda$ and $H_\lambda \equiv \lambda F + (1-\lambda)G$.
(c) Furthermore,

$$\begin{aligned} Cov\left[\mathbb{Z}_1(s), \mathbb{Z}_1(t)\right] &= (1-\lambda)\overline{G}(s)\overline{G}(t)\left(\overline{F}\sigma_F^2\right)(s \vee t) + \lambda\overline{F}(s)\overline{F}(t)\left(\overline{G}\sigma_G^2\right)(s \vee t), \\ Cov\left[\mathbb{Z}_2(s), \mathbb{Z}_2(t)\right] &= (e_G - e_F)(s)(e_G - e_F)(t) \\ &\quad \cdot \left\{(1-\lambda)\overline{G}(s)\overline{G}(t)(F(s \wedge t) - F(s)F(t)) + \lambda\overline{F}(s)\overline{F}(t)(G(s \wedge t) - G(s)G(t))\right\}, \\ Cov\left[\mathbb{Z}_1(s), \mathbb{Z}_2(t)\right] &= (e_G - e_F)(t)\left\{\lambda\overline{F}(s)\overline{F}(t)Cov\left[\widetilde{\mathbb{Z}}_G(s)\overline{G}(s), \tilde{\mathbb{U}}(G(t))\right]\right. \\ &\quad \left. + (1-\lambda)\overline{G}(s)\overline{G}(t)Cov\left[\mathbb{Z}_F(s)\overline{F}(s),\ \mathbb{U}(F(t))\right]\right\}, \end{aligned}$$

where

$$\begin{aligned} Cov\left[\mathbb{Z}_F(s)\overline{F}(s), \mathbb{U}(F(t))\right] &= -1\{s \leq t\}\overline{F}(t)\left\{e_F(t) - e_F(s) - \overline{F}(t)(t-s)\right\}, \\ Cov\left[\widetilde{\mathbb{Z}}_G(s)\overline{G}(s), \widetilde{\mathbb{U}}(G(t))\right] &= -1\{s \leq t\}\overline{G}(t)\left\{e_G(t) - e_G(s) - \overline{G}(t)(t-s)\right\}. \end{aligned}$$

In particular, by taking $s = t$ we find that $Var\left[\mathbb{Z}(t)\right]$ is given by

$$\begin{aligned} Var\left[\mathbb{Z}(t)\right] &= (1-\lambda)\overline{G}^2(t)\left(\overline{F}\sigma_F^2\right)(t) + \lambda\overline{F}^2(t)\left(\overline{G}\sigma_G^2\right)(t) \\ &\quad + (e_G - e_F)^2(t) \cdot \left\{(1-\lambda)\overline{G}^2(t)F(t)\overline{F}(t) + \lambda\overline{F}^2(t)G(t)\overline{G}(t)\right\}. \end{aligned}$$

5 Studentization: estimating the asymptotic variances

First, recall the expression A^2 for the asymptotic variance of $T_{m,n}$ under the null hypothesis H given in (A.1), and the expression $\overline{A}^2$ for the asymptotic variance of $\overline{T}_{m,n}$ under H given in (A.2). Here we propose two different estimators of the asymptotic variances A^2 and $\overline{A}^2$ under the null hypothesis H.

The first estimator uses the null hypothesis very strongly. Since the limiting distribution under the null hypothesis is $N(0, A^2)$ with

$$A^2 = 2\int_0^\infty \left(\int_t^\infty \overline{F}^2(s)\sigma_F^2(s)dF(s)\right)\overline{F}(t)dF(t)$$

and $\mathbb{H}_N$ satisfies $\|\mathbb{H}_N - F\|_\infty \to_{a.s.} 0$ under H, one possible estimator is given by

$$\widehat{A}^2_{(1)} \equiv 2\int_0^\infty \left(\int_t^\infty \widehat{\left(\overline{F}^2(s)\sigma_F^2(s)\right)}d\mathbb{H}_N(s)\right)\overline{\mathbb{H}}_N(t)d\mathbb{H}_N(t)$$

where

$$\widehat{\left(\overline{F}^2\sigma_F^2\right)}(t) \equiv \overline{\mathbb{H}}_N(t)\int_t^\infty (x-t)^2 d\mathbb{H}_N(x) - \left(\int_t^\infty (x-t)d\mathbb{H}_N(x)\right)^2.$$

This last display is based on the fact (correcting a minor typographical error in the display below (2.4) on page 5 of [**?**]) that

$$\sigma_F^2(t) \equiv Var\left[X - t | X > t\right] = \frac{\int_t^\infty (x-t)^2 dF(x)}{\overline{F}(t)} - e_F^2(t).$$

The second estimator of A^2 is based on the decomposition of the limit process $\mathbb{Z}$ under fixed alternatives as given in (A.1). Note that the process $\mathbb{Z}_2$ is identically zero under the null hypothesis, and that the process $\mathbb{Z}_1$ has covariance function given in (c) of Theorem 2:

$$Cov\left[\mathbb{Z}_1(s), \mathbb{Z}_1(t)\right] = \lambda\overline{F}(s)\overline{F}(t)\left(\overline{G}\sigma_G^2\right)(s\vee t) + (1-\lambda)\overline{G}(s)\overline{G}(t)\left(\overline{F}\sigma_F^2\right)(s\vee t).$$

Furthermore, we also have (under the null hypothesis H)

$$T_{m,n} \to_d \int_0^\infty \mathbb{Z}_1(t)dH_\lambda(t) \sim N(0, A_\lambda^2)$$

where

$$\begin{aligned} A_\lambda^2 &= E\left(\int_0^\infty \mathbb{Z}_1(s)dH_\lambda(s)\cdot\int_0^\infty \mathbb{Z}_1(t)dH_\lambda(t)\right) \\ &= \int_0^\infty\int_0^\infty \left\{\lambda\overline{F}(s)\overline{F}(t)(\overline{G}\sigma_G^2)(t\vee s) + (1-\lambda)\overline{G}(s)\overline{G}(t)(\overline{F}\sigma_F^2)(t\vee s)\right\}dH_\lambda(s)dH_\lambda(t). \end{aligned}$$

Thus our second proposed estimator of A^2 is

$$\widehat{A}^2_{(2)} \equiv \int_0^\infty\int_0^\infty \left\{\lambda_N\overline{\mathbb{F}}_m(s)\overline{\mathbb{F}}_m(t)\widehat{(\overline{G}\sigma_G^2)}(t\vee s) + (1-\lambda_N)\overline{\mathbb{G}}_n(s)\overline{\mathbb{G}}_n(t)\widehat{(\overline{F}\sigma_F^2)}(t\vee s)\right\}d\mathbb{H}_N(s)d\mathbb{H}_N(t)$$

where

$$\widehat{(\overline{F}\sigma_F^2)}(t) \equiv \int_t^\infty (x-t)^2 d\mathbb{F}_m(x) - \overline{\mathbb{F}}_m(t) e^2_{\mathbb{F}_m}(t), \tag{A.1}$$

$$\widehat{(\overline{G}\sigma_G^2)}(t) \equiv \int_t^\infty (x-t)^2 d\mathbb{G}_n(x) - \overline{\mathbb{G}}_n(t) e^2_{\mathbb{G}_n}(t). \tag{A.2}$$

It seems apparent that this second estimator relies less strongly on the null hypothesis H. With these variance estimators in hand, the studentized test statistics based on $T_{m,n}$ are

$$\hat{T}^{(1)}_{m,n} \equiv \frac{T_{m,n}}{\widehat{A}_{(1)}} \quad \text{and} \quad \hat{T}^{(2)}_{m,n} \equiv \frac{T_{m,n}}{\widehat{A}_{(2)}}. \tag{A.3}$$

The corresponding estimators of $\overline{A}^2$ are as follows: the first (simple) estimator is

$$\widehat{\overline{A}}^2_{(1)} \equiv 2 \int_0^\infty \left(\int_t^\infty \widehat{(\overline{F}\sigma_F^2)}(s) d\mathbb{H}_N(s) \right) d\mathbb{H}_N(t)$$

where

$$\widehat{(\overline{F}\sigma_F^2)}(t) \equiv \left\{ \int_t^\infty (x-t)^2 d\mathbb{H}_N(x) - \frac{\left(\int_t^\infty (x-t) d\mathbb{H}_N(x)\right)^2}{\overline{\mathbb{H}}_N(t)} \right\} 1_{[0, X_{(m)} \vee Y_{(n)}]}(t).$$

On the other hand, the second, somewhat more complicated, estimator is

$$\widehat{\overline{A}}^2_{(2)} \equiv \int_0^\infty \int_0^\infty \left\{ \lambda_N \widehat{(\overline{G}\sigma_G^2)}(t \vee s) + (1-\lambda_N) \widehat{(\overline{F}\sigma_F^2)}(t \vee s) \right\} d\mathbb{H}_N(s) d\mathbb{H}_N(t)$$

where $\widehat{(\overline{F}\sigma_F^2)}$ and $\widehat{(\overline{G}\sigma_G^2)}$ are given by (A.1) and (A.2). Then the two studentized versions of $\overline{T}_{m,n}$ upon which we will base our tests are

$$\overline{T}^{(1)}_{m,n} \equiv \frac{\overline{T}_{m,n}}{\widehat{\overline{A}}_{(1)}} \quad \text{and} \quad \overline{T}^{(2)}_{m,n} \equiv \frac{\overline{T}_{m,n}}{\widehat{\overline{A}}_{(2)}}. \tag{A.4}$$

Theorem 5.1. *(Consistency of the the variance estimators). If the null hypothesis holds, then*

$$\widehat{A}^2_{(1)} \to_p A^2, \quad \widehat{A}^2_{(2)} \to_p A^2 \quad \textit{as} \quad m \wedge n \to \infty,$$

and

$$\widehat{\overline{A}}^2_{(1)} \to_p \overline{A}^2, \quad \widehat{\overline{A}}^2_{(2)} \to_p \overline{A}^2 \quad \textit{as} \quad m \wedge n \to \infty.$$

6 Limit theory under local alternatives

Now we study the test statistic(s) $T_{m,n}/\hat{A}_j$ and $\overline{T}_{m,n}/\hat{A}_j$, $j = 1, 2$, under local alternatives and study local asymptotic approximations to the power of our tests. Suppose that $\{f_m\}$ and $\{g_n\}$ are sequences of densities satisfying

$$\begin{aligned} &\|\sqrt{m}(\sqrt{f_m} - \sqrt{f}) - (1/2)\alpha\sqrt{f}\|_2 \to 0 \ \text{ as } \ m \to \infty, \\ &\|\sqrt{n}(\sqrt{g_n} - \sqrt{f}) - (1/2)\beta\sqrt{f}\|_2 \to 0 \ \text{ as } \ n \to \infty \end{aligned} \tag{A.1}$$

where $\alpha, \beta \in L_2(F)$: $\int \alpha^2 dF < \infty$ and $\int \beta^2 dF < \infty$. Then (see e.g. Shorack and Wellner (1986), Theorem 4.1.1, page 152; or, see BKRW, Example 1, pages 191-192)

$$\begin{aligned} &\|\sqrt{m}(F_m - F) - A\|_\infty \to 0, \quad \text{and} \\ &\|\sqrt{n}(G_n - G) - B\|_\infty \to 0 \end{aligned}$$

where

$$A(x) \equiv \int_{-\infty}^{x} \alpha(y) dF(y), \quad \text{and} \quad B(x) \equiv \int_{-\infty}^{x} \beta(y) dF(y).$$

The following proposition connects these with the corresponding sequences $\{e_{F_m}\}$ and $\{e_{G_n}\}$ of mean residual life functions.

Proposition 1. Suppose that the sequences $\{f_m\}$ and $\{g_n\}$ are as in (A.1) and that $E_f(X^2) < \infty$. Then

$$\begin{aligned} &\| \left[\sqrt{m}(e_{F_m} - e_F) - C\right] \overline{F}\|_\infty \to 0 \quad \text{as} \quad m \to \infty, \quad \text{and} \\ &\| \left[\sqrt{n}(e_{F_n} - e_F) - D\right] \overline{F}\|_\infty \to 0 \quad \text{as} \quad n \to \infty, \end{aligned}$$

where

$$\begin{aligned} C(x) \equiv C(x;\alpha) &= \frac{\int_x^\infty (y-x)\alpha(y)dF(y)}{\overline{F}(x)} - \frac{e_F(x)}{\overline{F}(x)} A(x), \\ D(x) \equiv D(x;\beta) &= \frac{\int_x^\infty (y-x)\beta(y)dF(y)}{\overline{F}(x)} - \frac{e_F(x)}{\overline{F}(x)} B(x). \end{aligned}$$

Now we can describe the behavior of the processes $\{\mathbb{Z}_{m,n} : \ m \geq 1, \ n \geq 1\}$ and $\{\overline{\mathbb{Z}}_{m,n} : \ m \geq 1, \ n \geq 1\}$ under the local alternatives $\{f_m\}$ and $\{g_n\}$.

Proposition 2. If $\{f_m\}$ and $\{g_n\}$ are as in Proposition 1, and $\lambda_N \equiv m/N \to \lambda \in (0,1)$. Then, with $\delta_{m,n} \equiv (e_{G_n} - e_{F_m})\overline{F}_m\overline{G}_n$,

$$\begin{aligned} \mathbb{Z}_{m,n} \quad &\equiv \quad \sqrt{\frac{mn}{N}}\hat{\delta}_{m,n} = \sqrt{\frac{mn}{N}}\left(\hat{\delta}_{m,n} - \delta_{m,n}\right) + \sqrt{\frac{mn}{N}}\delta_{m,n} \\ &\Rightarrow \quad \mathbb{Z} + (\sqrt{\lambda}D - \sqrt{1-\lambda}C)\overline{F}^2 \\ &\stackrel{d}{=} \quad \overline{F}\mathbb{B}(\overline{F}\sigma_F^2) + (\sqrt{\lambda}D - \sqrt{1-\lambda}C)\overline{F}^2. \end{aligned}$$

This, in turn, yields the following theorem concerning the behavior of the test statistics $\{T_{m,n}/\hat{A}_{(j)}\}$ and $\{\overline{T}_{m,n}/\hat{\overline{A}}_{(j)}\}$ for $j = 1, 2$.

Theorem 6. Suppose that $\{f_m\}$ and $\{g_n\}$ are as in Proposition 1, and $\lambda_N \equiv m/N \to \lambda \in (0,1)$. Then

$$T_{m,n} \to_d N(\Delta, A^2) \quad \text{and} \quad \overline{T}_{m,n} \to_d N(\overline{\Delta}, \overline{A}^2)$$

where

$$\Delta \equiv \int_0^\infty \overline{F}^2 \left(\sqrt{\lambda} D - \sqrt{1-\lambda} C\right) dF,$$
$$\overline{\Delta} \equiv \int_0^\infty \overline{F} \left(\sqrt{\lambda} D - \sqrt{1-\lambda} C\right) dF,$$

and where A^2 and $\overline{A}^2$ are given by (A.1) and (A.2) respectively. Consequently, under $\{f_m\}$ and $\{g_n\}$

$$T_{m,n}/\hat{A}_{(j)} \to_d N(\Delta/A, 1), \quad \text{and} \tag{A.2}$$
$$\overline{T}_{m,n}/\hat{\overline{A}}_{(j)} \to_d N(\overline{\Delta}/\overline{A}, 1), \tag{A.3}$$

and it follows that

$$\lim_{m,n\to\infty} P_{f_m,g_n}(T_{m,n}/\hat{A}_{(j)} \geq z_\alpha) = P(Z \geq z_\alpha - \Delta/A),$$
$$\lim_{m,n\to\infty} P_{f_m,g_n}(\overline{T}_{m,n}/\hat{\overline{A}}_{(j)} \geq z_\alpha) = P(Z \geq z_\alpha - \overline{\Delta}/\overline{A})$$

for $j = 1, 2$ where $Z \sim N(0,1)$, $P(Z \geq z_\alpha) = \alpha$.

In fact, in particular examples it will be easier to compute the functions C and D appearing in Δ and $\overline{\Delta}$ by consider certain parametric families $e_m \equiv e_{F_m}$ and $e_n \equiv e_{G_n}$.

7 Examples and special cases

Here we calculate explicitly in some particular examples.

Example 1. Suppose that $e_F(x) = (\mu - cx)_+$ and $e_G(x) = (\mu - dx)_+$ with $0 < d < c$ so that $e_G \geq e_F$ with equality only at $x = 0$; see Figure 1. We will take F to be fixed and consider $G_n \to G$ with μ fixed and $d = d_n \equiv c - vn^{-1/2}$. Then we find that

$$\sqrt{n}(e_{G_n} - e_F)(x) \to vx1_{[0,\mu/c]}(x) \equiv D(x).$$

Since $\overline{F}(x) = (1-bx)^{-1+1/c} 1_{[0,1/b)}(x)$ with $b \equiv c/\mu$, we find that

$$\begin{aligned} \Delta &= \int_0^{\mu/c} \overline{F}^2(t)\sqrt{\lambda} v t dF(t) \\ &= \sqrt{\lambda} v \mu \frac{1}{3(3-2c)}, \end{aligned}$$

and, with $\gamma^2 = (1-c)/(1+c)$,

$$A^2 = 2\gamma^2\mu^2\frac{(1/c-1)^2}{(3/c-1)(5/c-3)} = 2\mu^2\frac{(1-c)^3}{(1+c)(3-c)(5-3c)}.$$

Consequently

$$\frac{\Delta}{A} = \frac{\sqrt{\lambda}v\mu/(3(3-2c))}{\sqrt{2}\mu\sqrt{\frac{(1-c)^3}{(1+c)(3-c)(5-3c)}}} = \sqrt{\frac{\lambda}{2}}\frac{v}{3(3-2c)}\sqrt{\frac{(1+c)(3-c)(5-3c)}{(1-c)^3}}.$$

Furthermore, similar calculations yield

$$\frac{\overline{\Delta}}{\overline{A}} = \frac{\sqrt{\lambda}v\sqrt{(1+c)(3-c)}}{2(2-c)\sqrt{(1-c)^3}}.$$

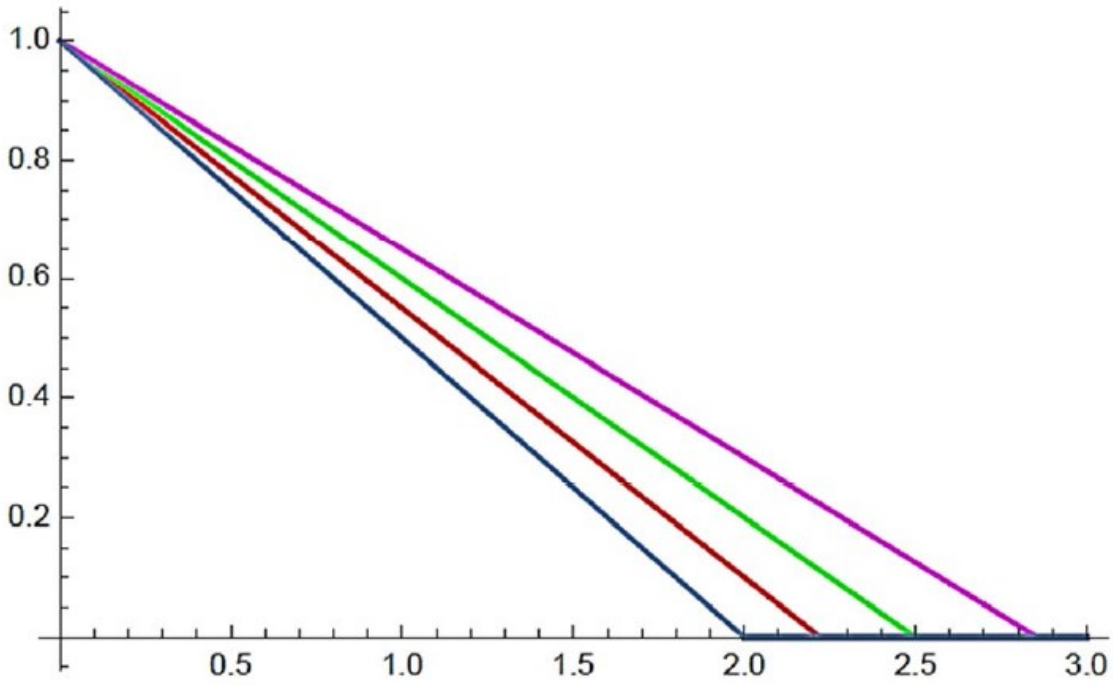


Figure 1: Pareto alternatives, negative slope, slope of e_G smaller than slope of e_F but equal means

Example 2. Suppose that $e_F(x) = (\mu - cx)_+$ and $e_G(x) = (\nu - cx)_+$ with $0 < \mu < \nu$ so that $e_G \geq e_F$. We take F to be fixed and consider $G_n \to G$ with $c > 0$ fixed and $\nu = \nu_n = \mu + vn^{-1/2}$. See Figure 2. Then

$$\sqrt{n}(e_{G_n} - e_F)(x) \to v1_{[0,\mu/c]}(x) \equiv D(x).$$

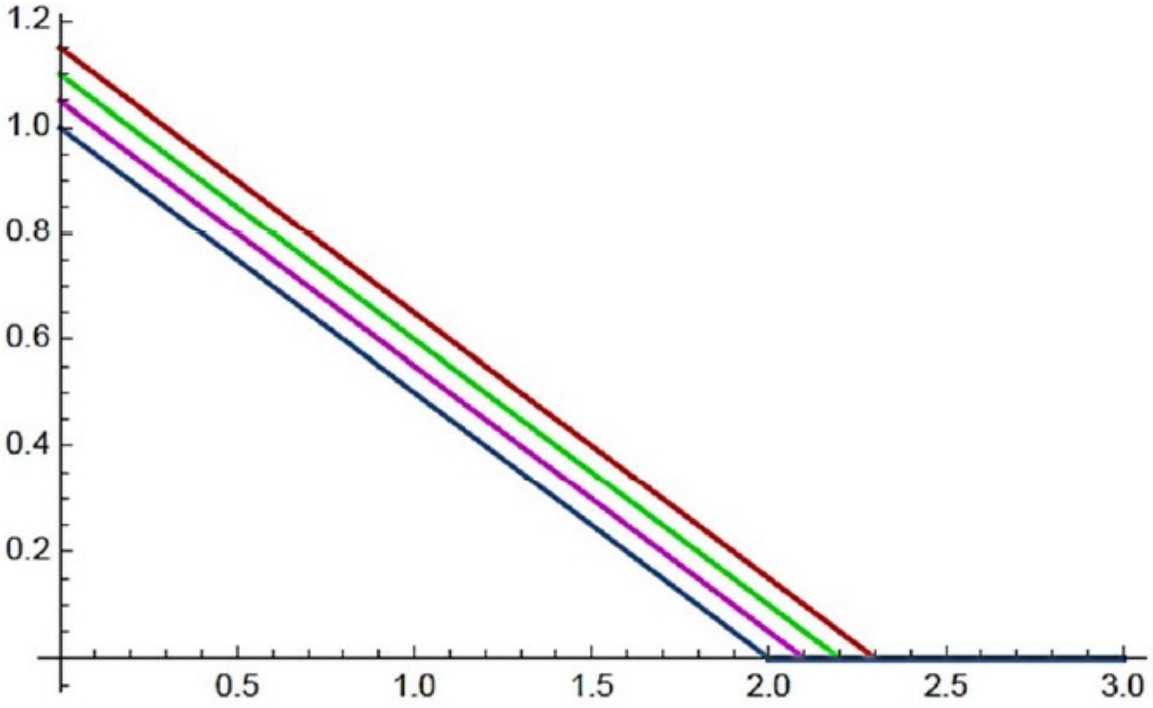


Figure 2: Pareto alternatives, negative slope, $e_G(0) = \mu_G$ greater than $e_F(0 = \mu_F$ but equal slopes

Since $\overline{F}(x) = (1 - bx)^{-1+1/c}$ with $b \equiv c/\mu$ we find that

$$\Delta = \int_0^{\mu/c} \overline{F}^2(t)\sqrt{\lambda}v dF(t) = \frac{1}{3}\sqrt{\lambda}v$$

while A^2 is exactly as in Example 1. This yields

$$\frac{\Delta}{A} = \frac{\frac{1}{3}\sqrt{\lambda}v}{\sqrt{2}\mu\sqrt{\frac{(1-c)^3}{(1+c)(3-c)(5-3c)}}} = \sqrt{\frac{\lambda}{2}}\frac{v}{3\mu}\sqrt{\frac{(1+c)(3-c)(5-3c)}{(1-c)^3}}.$$

Moreover we find that $\overline{\Delta} = \frac{1}{2}\sqrt{\lambda}v$ while $\overline{A}^2$ is as in Example 1. Consequently

$$\frac{\overline{\Delta}}{\overline{A}} = \frac{\frac{\sqrt{\lambda}\nu}{2}}{\frac{\mu^2\sqrt{(1-c)^3}}{\sqrt{(3-c)(1+c)}}} = \frac{\nu\sqrt{\lambda(3-c)(1+c)}}{2\mu\sqrt{(1-c)^3}}.$$

This yields

$$\frac{\Delta/A}{\overline{\Delta}/\overline{A}} = \frac{2\sqrt{5-3c}}{3\sqrt{2}};$$

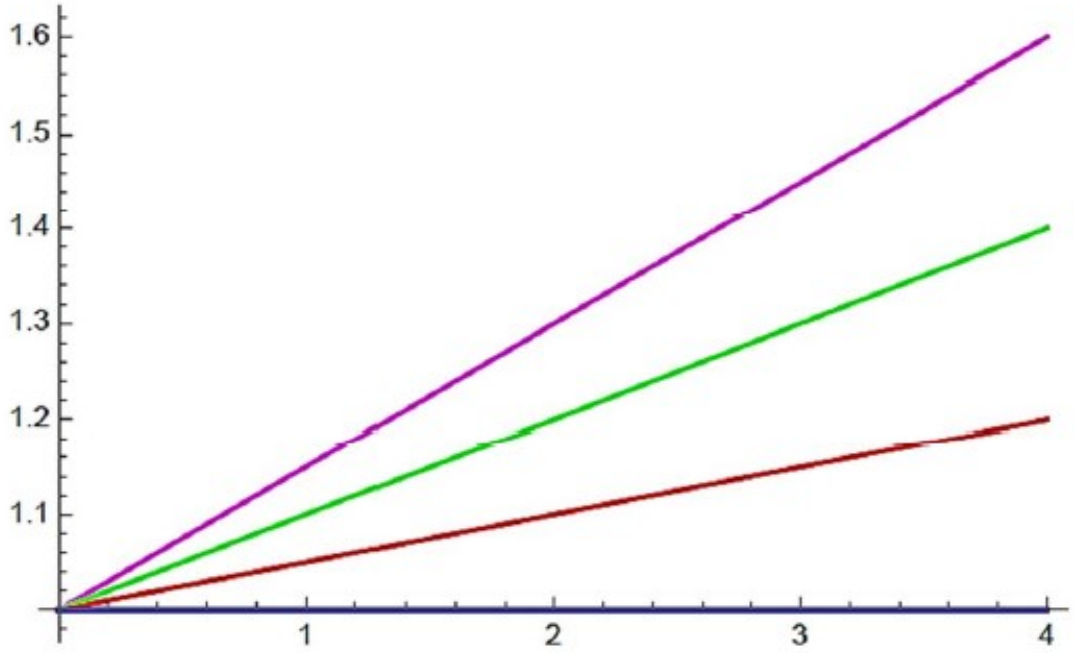


Figure 3: Pareto alternatives, positive slope, $e_G(0) = \mu = e_F(0)$ but greater slopes

this decreases from $(2/3)\sqrt{5/3} \approx 1.05$ at $c = 0$ to $(2/3) \approx 0.67$ at $c = 1$, and crosses 1 at $c \approx .17$ Thus we expect the test statistics $T_{m,n}/\hat{A}_{(j)}$ to be somewhat more powerful than the test statistics $\overline{T}_{m,n}/\hat{\overline{A}}_{(j)}$ for $c < 0.17$, but that the reverse will be true for $c > 0.17$. This is supported by simulations for finite m, n.

References

[1] Berger, R. L., Boos, D. D. and Guess, F. M. (1988).Tests and confidence sets for comparing two mean residual life functions. *Biometrics*, **44**, 103-115.

[2] Dudley, R. M.(1997). *Empirical processes and p-variation.* In Festschrift for Lucien Le Cam 219-233.Springer, New York.

[3] Hall, W. J. and Wellner, J. (1981). Mean residual life. *In Statistics and related topics (Ottawa, Ont.,1980)* ,169-184. North-Holland, Amsterdam.

[4] Hall, W. J. and Wellner, J. A.(2016). Estimation of mean residual life. In IMS Collections, Statistical

[5] Jeong, J.-H (2014). *Statistical inference on residual life*. Statistics for Biology and Health. Springer, NewYork.

[6] Shorack, G. R. and Wellner, J. A.(1986). *Empirical Processes with Applications to Statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York.

[7] Ansley, C.F. and Kohn, R. (1983), Exact likelihood of vector autoregressive, *Biometrika*, **70**, 275-278.

[8] Cressie, N. (1993), *Statistics for Spatial Data*, Revised edition, John Wiley, New York.

[9] Guan, Y. (2006), A composite likelihood approach in fitting spatial point process models. *Journal of the American Statistical Association*, **101**(476), 1502-1512.

[10] Møller, J. and Waagepetersen, R. (2003), *Statistical inference and simulation for spatial point processes*, Chapman and Hall/CRC Press.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Doubly truncated (interval) quantile Shannon entropy

Shokrani, A. [1] and Khorashadizadeh, M. [2]

[1,2] Department of Statistics, University of Birjand

Abstract

It is well known that Shannon's entropy plays an important role in the measurement of uncertainty of probability distributions. Also, in statistical modelling and analysis of data there is an equivalent and alternate approach, through the quantile function (QF), which is very important in exploratory data analysis and in many other areas of applied statistics. In the present paper, we introduce and study quantile version of the Shannon entropy functions in doubly truncated (interval) lifetime, which includes the residual and past lifetimes as special case.

Keywords: Shannon entropy, Quantile function, Generalized failure rate, Quantile doubly truncated Shannon entropy.

1 Introduction

The main measure of the uncertainty contained in random variable X is the Shannon entropy (Shannon, 1948) defined by,

$$H_X = -E(\log f(X)) = -\int_0^\infty f(x)\log f(x)dx.$$

Studying the measures based on residual life time random variable $X_t = X - t|X \geq t$ has a fundamental role in many areas such as reliability theory, survival analysis and information theory.

[1]afsaneshokrani@gmail.com
[2]M.Khorashadizadeh@Birjand.ac.ir

Ebrahimi and Pelerey (1995) and Ebrahimi (1996) defined the residual entropy (RE) based on the random variable X_t by

$$RE_X(t) = -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \log \frac{f(x)}{\overline{F}(t)} dx. \tag{A.1}$$

Di Crescenzo and Longobardi (2002) introduced an entropy-based measure of uncertainty in past life time distributions and called it past entropy (PE) which represents the uncertainty of the idle time (or inactivity time) of a component or system which is based on past life time random variable, $X_t^* = t - X|X \leq t$, and it is given by,

$$PE_X(t) = -\int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)} dx.$$

Sometimes, in many situations, we only have information between two points, so we should study the statistical measures (especially in information theory and reliability) under the condition of doubly truncated random variables. For example in reliability, if X denotes the lifetime of a unit, then the random variable $X_{t_1,t_2} = X - t_1|t_1 \leq X \leq t_2$ is called the doubly truncated (interval) residual lifetime, which is in special case of $t_2 \to \infty$ tends to residual lifetime random variable X_t. Also, we can use the doubly truncated (interval) past lifetime random variable $X^*_{t_1,t_2} = t_2 - X|t_1 \leq X \leq t_2$, which in special case $t_1 = 0$, it tends to the past lifetime random variable X_t^*. Another extension of shannon entropy is based on a doubly truncated (interval) random variable, which is defined by Misagh and Yari (2010, 2011, 2012) as follows,

$$H_X(t_1, t_2) = -\int_{t_1}^{t_2} \frac{f(x)}{F(t_2) - F(t_1)} \log \frac{f(x)}{F(t_2) - F(t_1)} dx,$$

where $(t_1, t_2) \in D = \{(t_1, t_2) : F(t_1) < F(t_2)\}$ and $H_X(0, \infty)$ is the Shannon entropy H_X and $H_X(t_1, \infty)$ is the residual entropy$RE_X(t_1)$ and also $H_X(0, t_2)$ is the past entropy $PE_X(t_2)$. Recently, many researchers have pay attention to use of quantile function (QF) instead of distribution function, which is defined for any random variables by:

$$Q(u) = F^{-1}(u) = \inf\{t|F(t) \geq u\}, \quad 0 \leq u \leq 1.$$

Since quantile functions are less influenced by extreme observation they are more convenient. For more details and recent studies on QF, its properties and usefulness in Reliability and information theories and also in the identification of models we refer to Gilchrist (2000), Nair and Sankaran (2009), Nair et al. (2008, 2011), Sankaran and Nair (2009), Sankaran et al. (2010), Sunoj and Sankaran (2012), Sunoj et al. (2013) and the references therein.

On the other hand, many of distributions in reliability are known by their quantile functions like Lambda distributions (van Staden and Loots, 2009), Power-Pareto distribution (Hankin and Lee, 2006) and Govindarajulu distribution (Nair et al., 2008), so using quantile version of reliability and information measures is the best way to analysis the data for such distributions.

From now on, we assume that F is continuous, so, $F(Q(u)) = FQ(u) = u$ and defining the density quantile function by $fQ(u) = f(Q(u))$ and quantile density function by $q(u) = \frac{\partial Q(u)}{\partial u}$, we have (see Parzen, 1979),

$$q(u)fQ(u) = 1.$$

From the perspective of quantile function, Sunoj and Sankaran (2012) introduced a quantile based Shannon entropy and residual entropy $RE_q(u)$ respectively by,

$$H_q = \int_0^1 \log q(u)du,$$

and

$$RE_q(u) = \log(1-u) + (1-u)^{-1} \int_u^1 \log q(p)dp.$$

Unlike the residual entropy of form (A.1), $RE_q(u)$ determine the QF uniquely. For more properties of $RE_q(u)$, one may refer to Sunoj and Sankaran (2012). Similarly, Sunoj et al. (2013) have studied the past entropy in terms of QF by,

$$PE_q(u) = \log(u) + u^{-1} \int_0^u \log q(p)dp.$$

The measure $PE_q(u)$ gives the expected uncertainty contained in the conditional density about the predictability of and out come of X until $100u\%$ point of distribution.

2 Interval quantile based Shannon entropy

In this section, we first define the doubly truncated version of quantile entropy and obtain some of it's properties. In similar way of previous measures, the doubly truncated (interval) quantile Shannon entropy function is defined by,

$$\begin{aligned} H_q(u_1,u_2) &= -\int_{u_1}^{u_2} \frac{f(Q(p))}{u_2-u_1} \log \frac{f(Q(p))}{u_2-u_1} dQ(p) \\ &= \log(u_2-u_1) + \frac{1}{u_2-u_1} \int_{u_1}^{u_2} \log(q(p))dp, \end{aligned} \tag{A.1}$$

where $(u_1,u_2) \in D = \{(u_1,u_2); Q(u_1) < Q(u_2)\}$.
In doubly truncation, Navarro and Ruiz (1996) defined the generalized failure rate (GFR) by,

$$\begin{cases} h_1(t_1,t_2) = \frac{f(t_1)}{\overline{F}(t_1)-\overline{F}(t_2)}, \\ h_2(t_1,t_2) = \frac{f(t_2)}{F(t_2)-F(t_1)}. \end{cases}$$

We can define the quantile GFR by,

$$\begin{aligned}\Lambda_1(u_1,u_2) &= \frac{1}{(u_2-u_1)q(u_1)}; (u_1,u_2)\in D,\\ \Lambda_2(u_1,u_2) &= \frac{1}{(u_2-u_1)q(u_2)}; (u_1,u_2)\in D.\end{aligned}$$

Definition 1. *The random variable X is said to have*

- *decreasing (increasing) interval quantile entropy DIQE (IIQE) property in term of u_1 if and only if for any fixed u_2, $H_q(u_1,u_2)$ is decreasing (increasing) with respect to u_1.*

- *decreasing (increasing) interval quantile entropy DIQE (IIQE) property in term of u_2 if and only if for any fixed u_1 ,$H_q(u_1,u_2)$ is decreasing (increasing) with respect to u_2.*

Now, differentiating Eq.(A.1) with respect u_1, we get:

$$\begin{aligned}\frac{\partial H_q(u_1,u_2)}{\partial u_1} &= \frac{-1}{u_2-u_1}+\frac{1}{(u_2-u_1)^2}\int_{u_1}^{u_2}\log q(p)dp-\frac{1}{u_2-u_1}\log q(u_1)\\ &= \frac{1}{u_2-u_1}\{H_q(u_1,u_2)-\log(u_2-u_1)-\log q(u_1)+1\} \qquad (A.2)\end{aligned}$$

So, $H_q(u_1,u_2)$ is increasing with respect to u_1 if and only if, for all $(u_1,u_2)\in D$,

$$H_q(u_1,u_2)\geq \log(u_2-u_1)+\log(q(u_1))+1,$$

or equivalently

$$H_q(u_1,u_2)\geq -\log(\Lambda_1(u_1,u_2))+1.$$

Similarly for u_2 we have,

$$\begin{aligned}\frac{\partial H_q(u_1,u_2)}{\partial u_2} &= \frac{1}{u_2-u_1}-\frac{1}{(u_2-u_1)^2}\int_{u_1}^{u_2}\log q(p)dp+\frac{1}{u_2-u_1}\log q(u_2)\\ &= \frac{1}{u_2-u_1}\{-H_q(u_1,u_2)+\log(u_2-u_1)+\log q(u_2)+1,\end{aligned}$$

Hence , $H_q(u_1,u_2)$ is increasing with respect to u_2 if and only if, for all $(u_1,u_2)\in D$,

$$H_q(u_1,u_2)\leq \log(u_2-u_1)+\log(q(u_2))+1$$

or equivalently

$$H_q(u_1,u_2)\leq -\log(\Lambda_2(u_1,u_2))+1.$$

The next theorem shows that $H_q(u_1,u_2)$ can determine the distribution uniquely.

Theorem 2.1. *If X be a non-negative absolutely continue random variable then the quantile density function is uniquely determined for fixed u_2 by interval quantile entropy via relation,*

$$q(u) = \exp\{1 - \log(u_2 - u) + H_q(u, u_2) - (u_2 - u)H_q'(u, u_2)\}.$$

Proof. Using the Eq. (A.2) the required results are follow. □

Example 2.2. In this example we obtain the interval quantile shannon entropy for some distributions.

• Exponential distribution

If X have the exponential distribution with parameter λ. Then we have, $Q(p) = \frac{1}{\lambda}(-\log(1-p))$, $q(p) = \frac{1}{\lambda(1-p)}$ and the interval quantile Shannon entropy is,

$$\begin{aligned} H_q(u_1, u_2) &= \log(u_2 - u_1) - \frac{1}{u_2 - u_1}\int_{u_1}^{u_2} \log(\lambda(1-p))dp \\ &= \log(u_2 - u_1) - \log\lambda - \frac{(1-u_1)}{(u_2-u_1)}\log(1-u_1) - \frac{(u_2-1)}{(u_2-u_1)}\log(1-u_2) + 1. \end{aligned}$$

In special case of $u_1 = u$ and $u_2 = 1$ we have quantile residual entropy,

$$RE_q(u) = 1 - \log\lambda,$$

and when $u_1 = 0$ and , $u_2 = u$ we have quantile past entropy,

$$PE_q(u) = 1 - \log\lambda + \log u + (\frac{1-u}{u})\log(1-u).$$

• Pareto I distribution

Let X have the Pareto I distribution with $Q(p)s$,$\sigma\,(1-p)^{-\frac{1}{\alpha}}$ and $q(p) = \frac{\sigma}{\alpha}\,(1-p)^{-1-\frac{1}{\alpha}}$.Also, we have

$$\begin{aligned} H_q(u_1, u_2) &= \log(u_2 - u_1) + \frac{1}{u_2 - u_1}\{\log(\frac{\sigma}{\alpha}(u_2 - u_1)) + \int_{u_1}^{u_2}(-1 - \frac{1}{\alpha})\log(1-p)dp \\ &= \log(u_2 - u_1) + \log\frac{\sigma}{\alpha} + \frac{(-1-\frac{1}{\alpha})(1-u_1)}{(u_2-u_1)}\log(1-u_1) \\ &+ \frac{(-1-\frac{1}{\alpha})(u_2-1)}{(u_2-u_1)}\log(1-u_2) + 1 + \frac{1}{\alpha}. \end{aligned}$$

In special case of $u_1 = u$ and $u_2 = 1$ we have quantile residual entropy,

$$RE_q(u) = \log(\frac{\sigma}{\alpha}) + (\frac{\alpha+1}{\alpha}) - \frac{1}{\alpha}\log(1-u),$$

and when $u_1 = 0$ and $u_2 = u$ we have quantile past entropy,

$$PE_q(u) = \log(\frac{\sigma}{\alpha}) + (\frac{\alpha+1}{\alpha}) + \log u + (\frac{\alpha+1}{\alpha})(\frac{1-u}{u}) \log(1-u).$$

• Uniform distribution
Let X have the uniform distribution with parameter (a, b) and $Q(p)$,$a + (b-a)p$ and $q(p) = b - a$.Then quantile doubly truncated Shannon entropy for this distribution is,

$$H_q(u_1, u_2) \quad = \quad \log(u_2 - u_1) + \log(b-a).$$

In special case of $u_1 = u$ and $u_2 = 1$ we have quantile residual entropy,

$$RE_q(u) = \log(1-u) + \log(b-a),$$

and when $u_1 = 0$ and $u_2 = u$ we have quantile past entropy,

$$PE_q(u) = \log u + \log(b-a).$$

• Power function distribution
Let X have the power function distribution with parameter (α, β) and $Q(p)s$,$\alpha\, p^{\frac{1}{\beta}}$ and $q(p) = \frac{\alpha}{\beta}\, p^{\frac{1}{\beta}-1}$. Then quantile doubly truncated shannon entropy for this distribution is given by,

$$\begin{aligned} H_q(u_1, u_2) &= \log(u_2 - u_1) + \frac{1}{u_2 - u_1}\{\log(\frac{\alpha}{\beta})(u_2 - u_1) + \int_{u_1}^{u_2} (\frac{1}{\beta} - 1) \log(p) dp \\ &= \log(u_2 - u_1) + \log\frac{\alpha}{\beta} - \frac{(\frac{1}{\beta} - 1)(u_1)}{(u_2 - u_1)} \log(u_1) + \frac{(\frac{1}{\beta} - 1)(u_2)}{(u_2 - u_1)} \log(u_2) + 1 - \frac{1}{\beta}. \end{aligned}$$

In special case of $u_1 = u$ and $u_2 = 1$ we have quantile residual entropy,

$$RE_q(u) = \log(\frac{\alpha}{\beta}) + (\frac{\beta-1}{\beta}) + \log(1-u) + (\frac{\beta-1}{\beta})\frac{u \log u}{(1-u)},$$

and when $u_1 = 0$ and $u_2 = u$ we have quantile past entropy,

$$PE_q(u) = \log(\frac{\alpha}{\beta}) + (\frac{\beta-1}{\beta}) + \frac{1}{\beta} \log u.$$

Theorem 2.3. *If X is IIQE (DIQE) and if $\phi(.)$ is nonnegative, increasing and convex (concave) function, then $\phi(X)$ is also IIQE (DIQE) .*

Proof. If $g(y)$ is the pdf of $Y = \phi(X)$, then $g(y) = \frac{f(\phi^{-1}(y))}{\phi'(\phi^{-1}(y))}$;hence ,$g(Q_Y(u)) = \frac{1}{q_Y(u)} = \frac{fQ(u)}{\phi' Q(u)} = \frac{1}{q_x(u)\phi' Q(u)}$, then the interval quantile shannon entropy of Y, $H_q^Y(u_1, u_2)$ is as follow,

$$\begin{aligned} H_q^Y(u_1,u_2) &= \log(u_2-u_1) + \frac{1}{u_2-u_1}\int_{u_1}^{u_2} \log(q_Y(p))dp \\ &= \log(u_2-u_1) + \frac{1}{u_2-u_1}\int_{u_1}^{u_2} \log(q_X(p)\phi'(Q(p)))dp \\ &= \log(u_2-u_1) + \frac{1}{u_2-u_1}\int_{u_1}^{u_2} \log(q_X(p))dp + \frac{1}{u_2-u_1}\int_{u_1}^{u_2} \log(\phi'(Q(p))dp \\ &= H_q^X(u_1,u_2) + E(\log(\phi'(X))|\, \phi^{-1}(u_1) < X < \phi^{-1}(u_2)); (u_1,u_2)\epsilon D \end{aligned}$$

where $H_q^X(u_1, u_2)$ and $H_q^Y(u_1, u_2)$ are the DQE's of X and Y respectively. Now if X is $IIQE$ and $\phi(x)$ is nonnegative ,increasing and convex ,then $\phi(x)$ is also $IIQE$. □

Definition 2. *X said to have less IQE than Y if $H_q^Y(u_1, u_2) \geq H_q^X(u_1, u_2)$ for all $0 \leq u_1 \leq u_2 \leq 1$, and we write $X \leq_{IQE} Y$.*

For example if X and Y are two uniform distribution with parameters $(0, b_1)$ and $(0, b_2)$, and if $b_1 \leq b_2$, then

$$H_q^X(u_1,u_2) = \log(u_2-u_1) + \log(b_1) \leq \log(u_2-u_1) + \log(b_2) = H_q^Y(u_1,u_2),$$

so, $X \leq_{IQE} Y$.

Theorem 2.4. *Let $Z_1 = a_1 X + b_1$ and $Z_2 = a_2 Y + b_2$, where $a_1 > 0, a_2 > 0$ and $b_1 > 0, b_2 > 0$. If $X \leq_{IQE} Y$ and $a_1 \leq a_2$, then $Z_1 \leq_{IQE} Z_2$.*

Proof. Let $q_Z(u)$ be the quantile density function of the variable $Z = aX + b$, where $a > 0$ and $b > 0$. Then, using A.1 we get,

$$H_q^Z(u_1,u_2) = H_q^X(u_1,u_2) + \log(a), \tag{A.3}$$

which proved the theorem. □

It should be noted that all characterization problems of distribution functions discussed in literature are automatically hold for quantile functions under the transformation $x = Q(u)$. Other than these, new characterizations have been found exclusively in the quantile set-up. Nair et al. (2013) have obtained characterization results based on linear hazard quantile function and mean residual quantile function. In the next theorem we obtain a characterization of power distribution.

Theorem 2.5. *The non-negative random variable X have power distribution with parameters α and β if and only if for a real constant C,*

$$H_q(u_1,u_2)=\log(u_2-u_1)+C-\frac{u_1}{u_2-u_1}\log(\frac{\beta(u_2-u_1)}{\alpha}A(u_1))+\frac{u_2}{u_2-u_1}\log(\frac{\beta(u_2-u_1)}{\alpha}A(u_2)).$$

Proof. Using results of Theorem 2.1 and Example 2.2 and some calculations the required result is obtained. □

Remark 2.6. When $C=0$ we have the uniform distribution with parameters $(0,1)$ and when $C=1$ we have the uniform distribution with parameters $(0,\alpha)$.

References

[1] Di Crescenzo, A. and Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distribution. J. Appl. Probab. 39, 434-440.

[2] Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution. Sankhy: The Indian Journal of Statistics, Series A, 48-56.

[3] Ebrahimi, N. and Pellerey, F., (1995). New partial ordering of survival functions based on notion of uncertainty. J. Appl. Probab. 32, 202-211.

[4] Gilchrist, W., (2000). Statistical Modelling with Quantile Functions. Chapman and Hall/CRC, Boca Raton, FL.

[5] Hankin, R.K.S. and Lee, A. (2006). A new family of non-negative distributions. Aust. NZ J. Stat., 48, 67-78.

[6] Misagh, F. and Yari, G.H. (2010). A novel entropy-based measure of uncertainty to lifetime distributions characterizations. In: Proc. ICMS 10. Ref. No. 100196. Sharjah, UAE.

[7] Misagh, F. and Yari, G.H. (2011). On weighted interval entropy. Statistics and Probability Letters, 29, 167-176.

[8] Misagh, F. and Yari, G.H. (2012). Interval entropy and informative distance. Entropy, 14, 480-490.

[9] Nair, N.U. and Sankaran, P.G. (2009). Quantile based reliability analysis. Comm. Statist. Theory Methods, 38, 222-232.

[10] Nair, N.U., Sankaran, P.G. and Vinesh Kumar, B. (2008). Total time on test transforms of order n and its implications in reliability analysis. J. Appl. Probab., 45, 1126-1139.

[11] Nair, N.U., Sankaran, P.G. and Vinesh Kumar, B. (2011). Modelling lifetimes by quantile functions using Parzens score function. Statistics, 1, 1-13.

[12] Nair, N.U., Sankaran, P.G. and N. Balakrishnon, (2013). Quantile-Based Reliability Analysis. Springer, New York, Heidelberg, Dordrecht, London.

[13] Navarro, J. and Ruiz, J.M. (1996). Failure-rate function for doubly-truncated random variables. IEEE Transactions on Reliability 45 (4), 685-690.

[14] Parzen, E. (1979). Nonparametric statistical data modeling. J. Amer. Statist. Assoc., 74, 105-122.

[15] Sankaran, P.G. and Nair, N.U. (2009). Nonparametric estimation of the hazard quantile function. J. Nonparametr. Stat., 21, 757-767.

[16] Sankaran, P.G., Nair, N.U., Sreedevi, E.P., (2010). A quantile based test for comparing cumulative incidence functions of competing risks models. *Statist. Probab. Lett.*, 80, 886-891.

[17] Sunoj, S.M. and Sankaran, P.G., (2012). Quantile based entropy function. Statistics and Probability Letters, 82, 1049-1053.

[18] Sunoj, S. M., Sankaran, P. G., and Nanda, A. K. (2013). Quantile based entropy function in past lifetime. Statistics and Probability Letters, 83(1), 366-372.

[19] Shannon, C.E., (1948). A mathematical theory of communication. Bell Syst. Tech. J., 27, 379-423.

[20] van Staden, P.J. and Loots, M.T., (2009). L-moment estimation for the generalized lambda distribution. In: Third Annual ASEARC Conference, New Castle, Australia.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Fuzzy non-parametric predictive inference for the reliability of k-out-of-m systems

Shokri, S. [1] Sadeghpour-Gildeh, B. [2] Mohtashami Borzadaran, G.R. [3]

and Fathi Vajargah, B. [4]

[1] Department of Statistics, International Campus, Ferdowsi University of Mashhad
[2,3] Department of Statistics, Ferdowsi University of Mashhad
[4] Department of Statistics, Faculty of Mathematical Sciences, University of Guilan

Abstract

The aim of this paper is to present a new method for evaluating the reliability of k-out-of-m systems, titled the fuzzy non-parametric predictive inference. With this end in view, this paper presents fuzzy, lower and upper probabilities for the reliability of k-out-of-m systems. For the sake of reaching this aim, attention has been restricted to the k-out-of-m systems with exchangeable components. Moreover, the problem of evaluating system reliability based on nonparametric predictive inferential (NPI) approach has been considered in this paper, in which defining the parameters of reliability function as crisp values is not possible, and parameters of reliability function are described with the use of a triangular fuzzy number. Formula of a fuzzy reliability function and its -cut set are also presented. Moreover, the fuzzy reliability of structures is defined on the basis of fuzzy numbers, and the fuzzy reliability functions of k-out-of m systems are discussed. Finally, some numerical examples are presented to illustrate how to calculate the fuzzy reliability function and its α-cut set.

Keywords: k-out-of-m systems, Lower and upper probabilities, Nonparametric predictive inference, Fuzzy number.

[1] soheilshokri@yahoo.com,sadeghpour@um.ac.ir, grmohtashami@um.ac.ir, fathi@guilan.ac.ir
[2]
[3]
[4]

1 Introduction

This paper provides a new method for statistical inference on system reliability on the basis of uncertain information resulting from component testing. This Approach is called Fuzzy Nonparametric Predictive Inference (FNPI). We present FNPI for system reliability, in particular FNPI for k-out-of-m systems. Coolen [6] provided an insight into imprecise reliability, discussing a variety of issues and reviewing suggested applications of imprecise probabilities in reliability; see [8, 9, 10] for a detailed overview of imprecise reliability and many references. The nonparametric predictive approach is a statistical approach based on few assumptions about probability distributions together with inferences based on data [7, 8]. Nonparametric predictive inference (NPI) is close in nature to predictive inference for the low structure stochastic case as briefly outlined by Geisser [11]. In recent years, many theoretical aspects and a variety of applications of inferences based on Hills assumption of $A_{(n)}$ for prediction of probabilities for one (or more) future values on the basis of n prior observations have been presented, referred in the literature as nonparametric predictive inference (NPI), see e.g. [1, 2, 5, 7, 8]. Shokri et al. [15] present FNPI for reliability of series systems.The aim of the authors of the current paper is studying the reliability of k-out-of-m systems based on non-parametric predictive inference in a fuzzy environment and proposing a new method titled fuzzy non-parametric predictive inference for the reliability of k-out-of-m systems.

2 Concepts on Nonparametric Predictive Approach

Hill [13] introduced the assumption $A_{(n)}$ for predictions concerning future observations. This assumption was suggested particularly for states in which there is no strong prior knowledge about the probability distribution for a random variable of interest. Hill [13] explained $A_{(n)}$ in detail. Inferences based on $A_{(n)}$ are nonparametric and predictive, and can be considered proper if there is hardly any knowledge about the random variable of interest other than the n observations, or if one does not want to use such knowledge; for example, in the case of studying the influences of additional assumptions underlying other statistical techniques [7]. $A_{(n)}$ is not enough for deriving exact probabilities for many events of interest but presents optimal bounds for probabilities for all events of interest involving X_{n+1}. These bounds are lower and upper probabilities in the theory of interval probability[16]. Augustin and Coolen [1] showed that the interval probabilities obtained, based only on the $A_{(n)}$ assumption, have powerful consistency characteristics in the theory of interval probability [6]. Coolen [5] used $A_{(n)}$ for NPI in the case of Bernoulli data, presenting lower and upper probabilities for the number of successes in m future trials based on the number of successes in n observed trials. This was possible by considering similar representations of Bernoulli data as was used by Bayes [3] under the title of balls on a table.

Theorem 2.1. *Assume $n+m$ number of Bernoulli's exchangeable experiments whose result can be success or failure. Assume:*
$Y_{n+1}^{n+m} \rightarrow$ *The random variable of number of successes of m Bernoulli's future ($n+1 \rightarrow +m$) experiments.*

$Y_1^n \rightarrow$ *The random variable of the number of successes in n Bernoulli's previous (1 to n) experiments.*
For the sake of simplicity, we define $\binom{s+r_0}{s} = 0$*; therefore, the upper and lower probabilities of non-parametric predictive inferences are*

$$\bar{P}\left(Y_{n+1}^{n+m} \in R_t | Y_1^n = s\right) = \binom{n+m}{m}^{-1} \sum_{j=1}^{t} \left[\binom{s+r_j}{s} - \binom{s+r_{j-1}}{s}\right] \binom{n-s+m-r_j}{n-s}, \tag{A.1}$$

and

$$\underline{P}\left(Y_{n+1}^{n+m} \in R_t | Y_1^n = s\right) = 1 - \bar{P}\left(Y_{n+1}^{n+m} \in R_t^c | Y_1^n = s\right). \tag{A.2}$$

where $R_t = \{r_1, \ldots, r_t\}$ *with* $\circ \leq r_1 < r_2 < \ldots < r_t \leq m$, $1 \leq t \leq m+1$ *and* $R_t^c = \{\circ, 1, \ldots, m\} \setminus R_t$.

Proof. See [5]. □

3 Non-parametric predictive inference for a *k* -out-of-*m* system

The class of k-out-of-m systems, also called 'voting systems', was introduced by Birnbaum [4]. These are systems that consist of m exchangeable components (often the confusing term identical components is used) such that the system functions if and only if at least k of its components function. Since the value of m is usually larger than the value of k, redundancy is generally built into a k-out-of-m system [10].

Theorem 3.1. *considering a k-out-of-m system, the event* $Y_{n+1}^{n+m} \geq k$ *is of interest as this corresponds to successful functioning of a k-out-of-m system, following n tests of components that are exchangeable with the m components in the system considered. Given data consisting of s successes from n components tested, the NPI lower and upper probabilities for the event denoting that the k-out-of-m system functions successfully are also denoted by* $\overline{P}\left(S\left(m:k\right)|(n,s)\right)$ *and* $\underline{P}\left(S\left(m:k\right)|(n,s)\right)$*, respectively. For* $k \in \{1, 2, \cdots, m\}$ *and* $0 < s < n$

$$\begin{aligned}
&\overline{P}\left(S\left(m:k\right)|(n,s)\right) \\
&= \overline{P}\left(Y_{n+1}^{n+m} \geq k \middle| Y_1^n = s\right) \\
&= \binom{n+m}{m}^{-1} \left[\binom{s+k}{s}\binom{n-s+m-k}{n-s} + \sum_{l=k+1}^{m} \binom{s+l-1}{s-1} \binom{n-s+m-l}{n-s}\right], \\
&\underline{P}\left(S\left(m:k\right)|(n,s)\right) = \underline{P}\left(Y_{n+1}^{n+m} \geq k \middle| Y_1^n = s\right) \\
&= 1 - \overline{P}\left(Y_{n+1}^{n+m} \leq k-1 \middle| Y_1^n = s\right) \\
&= 1 - \binom{n+m}{m}^{-1} \left[\sum_{l=0}^{k-1} \binom{s+l-1}{s-1} \binom{n-s+m-l}{n-s}\right].
\end{aligned}$$

Proof. See [10]. □

4 Fuzzy Non-parametric Predictive Inference for the reliability of k -out-of-m systems

In this Section, we consider the problem of the evaluation of system reliability based on the nonparametric predictive inferential (NPI) approach, in which the defining the parameters of reliability function in definite quantities is not possible and parameters of reliability function are described using a triangular fuzzy number.The theory of sets and fuzzy logic was first proposed by Zadeh [17]. This theory has found wide applications in many fields such as computer, system analysis, electronic and recently in social sciences, economics and industry. Fuzzy logic is a theory for uncertain conditions. This theory can form many of concepts, variables and systems which are imprecise and vague in a mathematical form and provide the way for reasoning, control and decision-making in uncertain conditions [14, 17].

4.1 Fuzzy number of success in tested components(s)

The number of success in tested components can be defined by linguistic variables. One of the circumstances that can be assumed is when the number of functioning items is defined as linguistic variables. Fuzzy numbers can be used for showing functioning items. Assume that the number of functioning items can be shown as the following triangular fuzzy number:

$$\tilde{s} = \mathrm{TFN}(s_1, s_2, s_3), \quad s(\alpha) = (s_1 + (s_2 - s_1)\alpha\ , s_3 + (s_2 - s_3)\alpha).$$

Therefore, the fuzzy upper non-parametric predictive probability is as

$$\begin{aligned}&\widetilde{\overline{P}}\left(S(m:k)|(n,s)\right) = \widetilde{\overline{P}}\left(Y_{n+1}^{n+m} \geq k \middle| Y_1^n = s\right)\\ &= \binom{n+m}{m}^{-1}\left[\binom{\tilde{s}+k}{\tilde{s}}\binom{n-\tilde{s}+m-k}{n-\tilde{s}} + \sum_{l=k+1}^{m}\binom{\tilde{s}+l-1}{\tilde{s}-1}\binom{n-\tilde{s}+m-l}{n-\tilde{s}}\right].\end{aligned}$$

As a result

$$\widetilde{\overline{P}}(\alpha) = \left\{\binom{n+m}{m}^{-1}\left[\binom{s+k}{s}\binom{n-s+m-k}{n-s} + \sum_{l=k+1}^{m}\binom{s+l-1}{s-1}\binom{n-s+m-l}{n-s}\right]\middle| \ s \in s(\alpha)\right\}, 0 \leq \alpha \leq 1.$$

Too fuzzy lower non-parametric predictive probability is as

$$\underline{P}_{\ l}(\alpha) = min\left\{1 - \binom{n+m}{m}^{-1}\left[\sum_{l=0}^{k-1}\binom{s+l-1}{s-1}\ \binom{n-s+m-l}{n-s}\right]\middle| s \in s(\alpha)\right\},$$

$$\underline{P}_{\ r}(\alpha) = \ max\left\{1 - \binom{n+m}{m}^{-1}\left[\sum_{l=0}^{k-1}\binom{s+l-1}{s-1}\ \binom{n-s+m-l}{n-s}\right]\middle| s \in s(\alpha)\right\}.$$

4.2 Fuzzy numbers of tested components(n)

One of the other instances that can be assumed is when the number of tested components is defined as linguistic variables. Fuzzy numbers can be used for the depiction of the number of sample elements. Assume that n numbers of tested components are defined as the following triangular numbers:

$$\tilde{n} = \mathrm{TFN}(n_1,\, n_2,\, n_3), \qquad n\,(\alpha) = (n_1 + (n_2 - n_1)\,\alpha\,, n_3 + (n_2 - n_3)\,\alpha).$$

So fuzzy upper non-parametric predictive probability is as

$$\begin{aligned}&\widetilde{\overline{P}}\,(S\,(m:k)|(n,s)) = \widetilde{\overline{P}}\left(Y_{n+1}^{n+m} \geq k \middle| Y_1^n = s\right)\\ &= \binom{\tilde{n}+m}{m}^{-1}\left[\binom{\tilde{s}+k}{\tilde{s}}\ \binom{\tilde{n}-\tilde{s}+m-k}{\tilde{n}-\tilde{s}} + \sum_{l=k+1}^{m}\binom{\tilde{s}+l-1}{\tilde{s}-1}\ \binom{\tilde{n}-\tilde{s}+m-l}{\tilde{n}-\tilde{s}}\right].\end{aligned}$$

As a result,

$$\begin{aligned}\widetilde{\overline{P}}\,(\alpha) = \Bigg\{&\binom{n+m}{m}^{-1}\left[\binom{s+k}{s}\ \binom{n-s+m-k}{n-s}\right.\\ &\left.+\sum_{l=k+1}^{m}\binom{s+l-1}{s-1}\ \binom{n-s+m-l}{n-s}\right]\Bigg|\, s\in s\,(\alpha)\,,\ n\in n(\alpha)\Bigg\}, 0\leq\alpha\leq 1.\end{aligned}$$

Also fuzzy lower nonparametric predictive probability is given as follows,

$$\underline{\tilde{P}}\,(\alpha) = [\underline{P}\ _l\,(\alpha)\,,\, \underline{P}\ _r\,(\alpha)]\,,$$

$$\begin{aligned}\underline{P}\ _l\,(\alpha) = &\min\left\{1-\binom{n+m}{m}^{-1}\left[\sum_{l=0}^{k-1}\binom{s+l-1}{s-1}\ \binom{n-s+m-l}{n-s}\right]\right.\\ &\left.\Bigg|\, s\in s\,(\alpha)\,,\ n\in n(\alpha)\right\},\\ \underline{P}\ _r\,(\alpha) = &\max\left\{1-\binom{n+m}{m}^{-1}\left[\sum_{l=0}^{k-1}\binom{s+l-1}{s-1}\ \binom{n-s+m-l}{n-s}\right]\right.\\ &\left.\Bigg|\, s\in s\,(\alpha)\,,\ n\in n(\alpha)\right\}.\end{aligned}$$

4.3 example

Consider a 2-out-of-6 system with exchangeable components (so $k = 2$, $m = 6$), and the only information available is the result of a test of "Approximately 5" components, also exchangeable with the 6 to be used in the system. Assume that the numbers of successes in the tests are expressed as "Approximately 2". Triangular fuzzy numbers are more suitable to convert this definition into

a fuzzy number. The number of components to be converted to a triangular fuzzy number as $\widetilde{n} = \text{TFN } (4, 5, 6)$, and The number of successes in the tests to be converted to a triangular fuzzy number is as $\widetilde{s} = \text{TFN } (1, 2, 3)$. The FNPI lower and upper probabilities for successful functioning of the system are

$$\begin{aligned} \tilde{n} &= \text{TFN}(4,\ 5,\ 6), \\ \tilde{s} &= \text{TFN}(1,\ 2,\ 3), \\ n\,(\alpha) &= (4+\alpha\ , 6-\alpha), \\ s\,(\alpha) &= (1+\alpha\ , 3-\alpha). \end{aligned}$$

$$\overline{\widetilde{P}}\,(\alpha) = \left\{ \binom{n+6}{6}^{-1} \left[\binom{s+2}{s} \binom{n-s+6-2}{n-s} + \sum_{l=k+1}^{m} \binom{s+l-1}{s-1} \binom{n-s+6-l}{n-s} \right] \middle| s \in s\,(\alpha)\,, \right.$$
$$\left. n \in n\,(\alpha) \right\}, 0 \leq \alpha \leq 1.$$

Table 1 and 2 show α-cuts related to $\underline{\widetilde{P}}$ fuzzy lower non-parametric predictive probability and $\overline{\widetilde{P}}$ fuzzy upper non-parametric predictive probability. Figures 1 and 2 show the corresponding diagrams of membership function.

Table 1: α-cuts related to $\underline{\tilde{P}}$ fuzzy lower non-parametric predictive probability.

α	$\underline{P}_{\ l}(\alpha)$	$\underline{P}_{\ r}(\alpha)$	α	$\underline{P}_{\ l}(\alpha)$	$\underline{P}_{\ r}(\alpha)$
0	0.3333	0.7273	0.55	0.4799	0.6531
0.05	0.3480	0.7213	0.60	0.4915	0.6453
0.10	0.3625	0.7152	0.65	0.5029	0.6374
0.15	0.3766	0.7090	0.70	0.5141	0.6292
0.20	0.3905	0.7026	0.75	0.5250	0.6208
0.25	0.4041	0.6960	0.80	0.5356	0.6122
0.30	0.4174	0.6893	0.85	0.5460	0.6034
0.35	0.4305	0.6824	0.90	0.5562	0.5944
0.40	0.4432	0.6754	0.95	0.5661	0.5852
0.45	0.4557	0.6681	1	0.5758	0.5758
0.50	0.4679	0.6607			

5 Conclusions

Despite of the usefulness of reliability of k-out-of-m systems based on non-parametric predictive approach, it has a main difficulty in defining its parameters as crisp values. For these cases, the fuzzy set theory is the most suitable tool to analyze reliability of k-out-of-m systems based on nonparametric predictive approach. The obtained results show that the fuzzy definitions of

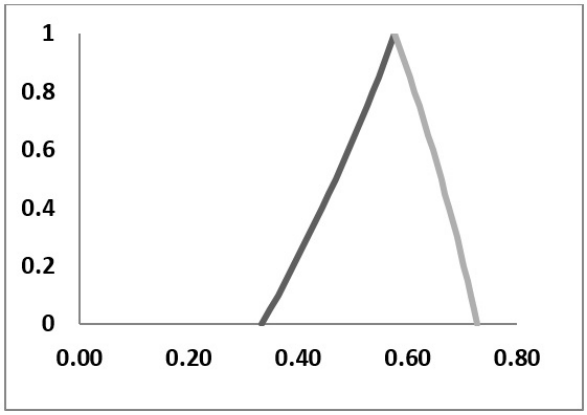


Figure 1: the diagram of membership function of lower non-parametric predictive probability

Table 2: α-cuts related to$\widetilde{\overline{P}}$ fuzzy upper non-parametric predictive probability.

α	$\overline{P}_{\ l}(\alpha)$	$\overline{P}_{\ r}(\alpha)$	α	$\overline{P}_{\ l}(\alpha)$	$\overline{P}_{\ r}(\alpha)$
0	0.6667	0.8788	0.55	0.7517	0.8425
0.05	0.6756	0.8759	0.60	0.7581	0.8386
0.10	0.6843	0.8730	0.65	0.7643	0.8346
0.15	0.6927	0.8700	0.70	0.7703	0.8304
0.20	0.7009	0.8669	0.75	0.7762	0.8262
0.25	0.7088	0.8637	0.80	0.7819	0.8218
0.30	0.7165	0.8604	0.85	0.7874	0.8173
0.35	0.7240	0.8570	0.90	0.7928	0.8127
0.40	0.7312	0.8535	0.95	0.7980	0.8079
0.45	0.7382	0.8499	1	0.8030	0.8030
0.50	0.7450	0.8463			

parameters provide more flexibility and more usability. In this article the non- parametric predictive probability has been analyzed for reliability of k-out-of-m systems with fuzzy parameters. We calculated the fuzzy reliability function and its α-cut set.

References

[1] Augustin, T. and Coolen, F. P. A. (2004), Nonparametric predictive inference and interval probability, *Journal of Statistical Planning and Inference*, **124**, 251-272.

[2] Augustin, T. Coolen, F. P. A., Cooman, G. D. and Troffaes, M.C.M. (2014), *Introduction to Imprecise Probabilities* , Wiley, Series in Probability and Statistics.

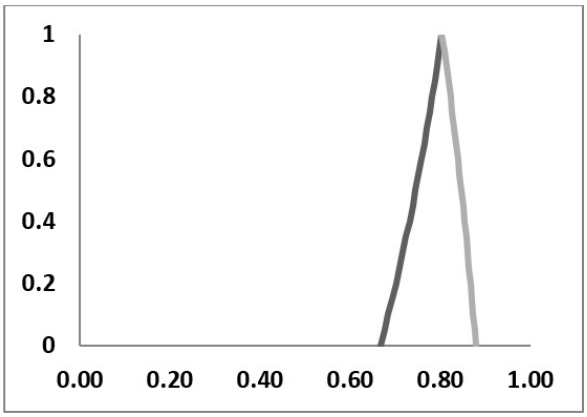


Figure 2: the diagram of membership function of upper non-parametric predictive probability

[3] Bayes,T. (1763), An essay towards solving a problem in the doctrine of chances, *Philosophical Transactions of the Royal Society London*, **53**, 370-418.

[4] Birnbaum, Z. W. Esary, J. D. and Saunders, S. (1961), Multi-component systems and structures and their reliability, *Technometrics*, **3**, 55-77.

[5] Coolen, F. P. A. (1998), Low structure imprecise predictive inference for Bayes' problem, *Statistics & Probability Letters*, **36**, 349-357.

[6] Coolen, F. P. A. (2004), On the use of imprecise probabilities in reliability, *Quality and Reliability Engineering International*, **20**, 193-202.

[7] Coolen, F. P. A. (2011) , Nonparametric predictive inference, *In International encyclopedia of statistical science*, ed. M. Lovric , Springer, Berlin, 968-970.

[8] Coolen, F. P. A. (2015), *Nonparametric Predictive Inference*, Wiley Series in Probability and Statistics, Wiley.

[9] Coolen, F. P. A. , Coolen-Schrijner P. and Yan, K. J. (2002), Nonparametric predictive inference in reliability, *Reliability Engineering and System Safety*, **78**, 185-193.

[10] Coolen-Schrijner, P. Coolen, F. P. A. and MacPhee, I. M. (2008), Nonparametric predictive inference for systems reliability with redundancy allocation, *Journal of Risk and Reliability*, **222**, 463-476.

[11] Geisser, S. (1993), *Predictive Inference: An Introduction*, Chapman & Hall, London.

[12] Hill, B.M. (1968), Posterior distribution of percentiles: Bayes' theorem for sampling from a population, *Journal of the American Statistical Association*, **63** , 677-691.

[13] Hill, B.M. (1988), De Finetti's theorem, induction, and A(n) or Bayesian nonparametric predictive inference (with discussion), *Bayesian Statistics* ,**3** ,211-241.

[14] Kahraman, C. and Kaya, I. (2010), *Production Engineering and Management under Fuzziness*, Springer, Berlin, 484–513.

[15] Shokri, S. Sadeghpour-Gildeh, B. Mohtashami Borzadaran, Gh. R. Fathi Vajargah, B. (2016), Fuzzy Nonparametric Predictive Inference for the Reliability of Series Systems, *Asian Journal of Fuzzy and Applied Mathematics*, **4**, 86-96.

[16] Weichselberger, K. (1995), Axiomatic foundations of the theory of interval-probability, *In: Mammitzsch, Proceedings of the Second GauU Symposion, Section B. De Gruyter, Berlin*, 47-64.

[17] Zadeh, L.A. (1965), Fuzzy sets, *Information and Control*, **8**, 338-353.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Cumulative residual inaccuracy in upper record values

Tahmasebi, S. [1] Daneshi, S. [2] and Salehi, M. [3]

[1,3] Department of Statistics, Persian Gulf University
[2] Department of Statistics, Faculty of Mathematics, Shahrood University of Technology

Abstract

In this paper, we propose a measure of cumulative residual inaccuracy between distribution of the nth upper record value and parent distribution F. We discuss some reliability properties of the proposed measure.

Keywords: Cumulative residual inaccuracy; Upper record values; Measure of inaccuracy.

1 Introduction

Suppose that X and Y are two non-negative random variables with reliability functions $\bar{F}(x)$, $\bar{G}(x)$, respectively. [3] defined the cumulative residual inaccuracy based on $\bar{F}(x)$, $\bar{G}(x)$ as

$$I(\bar{F},\bar{G}) = -\int_0^{+\infty} \bar{F}(x)\log\bar{G}(x)dx. \tag{A.1}$$

Now, let $\{X_m, m \geq 1\}$ be a sequences of independent and identically distributed random variables with cumulative distribution function (cdf) F and density function f. An observation X_j will be called an upper record value if its value exceeds all previous observations. Thus, X_j is an upper

[1]tahmasebi@pgu.ac.ir
[2]stahmasby@yahoo.com
[3]sdaneshi445@gmail.com

record value if $X_j > X_i$ for every $i < j$. Let $\{X_n = R_n\}$ denote the nth upper record value arising from the $\{X_m, m \geq 1\}$. Then the density function and survival function of R_n, which are denoted by f_{R_n} and $\bar{F}_{R_n}$, respectively, are given by

$$f_{R_n}(x) = \frac{[-\log \bar{F}(x)]^{n-1}}{(n-1)!} f(x), \quad x > 0, \quad n \geq 1,$$

$$\bar{F}_{R_n}(x) = \sum_{j=0}^{n-1} \frac{[-\log \bar{F}(x)]^j}{j!} \bar{F}(x).$$

Record values apply in problems such as industrial stress testing, meteorological analysis, hydrology, sporting and economics. In reliability theory, records values are used to study, for example, technical systems which are subject to shocks, e.g. peaks of voltages. For more details about records and their applications, one may refer to [1]. 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 $\bar{F}_{R_n}$ and $\bar{F}$ and study its characterization results.

2 Main results

We propose the cumulative residual measure of inaccuracy between $\bar{F}_{R_n}$ and $\bar{F}$ as follows:

$$\begin{aligned} I(\bar{F}_{R_n}, \bar{F}) &= -\int_0^{+\infty} \bar{F}_{R_n}(x) \log\left(\bar{F}(x)\right) dx \\ &= -\int_0^{+\infty} \sum_{j=0}^{n-1} \frac{[-\log \bar{F}(x)]^j}{j!} \bar{F}(x) \log\left(\bar{F}(x)\right) dx \\ &= \sum_{j=0}^{n-1} \int_0^{+\infty} \frac{[-\log \bar{F}(x)]^{j+1}}{j!} \bar{F}(x) dx \\ &= \sum_{j=0}^{n-1} (j+1) E_{R_{j+2}} \left(\frac{1}{\lambda_F(X)} \right), \end{aligned} \tag{A.1}$$

where $\lambda_F(X)$ is the hazard rate function of F and R_{j+2} is a random variable with reliability $\bar{F}_{R_{j+2}}$. In the following example, we calculate $I(\bar{F}_{R_n}, \bar{F})$ for some specific lifetime distributions which are widely used in reliability theory and life testing.

Example 2.1. (a) If X is uniformly distributed in $[0, \theta]$, then it is easy to see that $I(\bar{F}_{R_n}, \bar{F}) = \theta \sum_{j=0}^{n-1} \frac{j!(j+1)^2}{2^{j+2}}$, for all integers $n \geq 1$.
(b) If X has a Weibull distribution with survival function $\bar{F}(x) = e^{-\alpha x^\beta}$, $x \geq 0$, $\alpha, \beta > 0$, then for all integers $n \geq 1$, we obtain $I(\bar{F}_{R_n}, \bar{F}) = \frac{1}{\beta \alpha^{\frac{1}{\alpha}}} \sum_{j=0}^{n-1} \frac{(\frac{1}{\beta}+j)!}{j!}$.
(c) If X has a Pareto distribution with survival function $\bar{F}(x) = \left(\frac{\lambda}{x+\lambda}\right)^\gamma$, $x \geq 0$, $\gamma > 1, \lambda > 0$, then $I(\bar{F}_{R_n}, \bar{F}) = \frac{\lambda}{\gamma - 1} \sum_{j=0}^{n-1} (j+1) \left(\frac{\gamma}{\gamma-1}\right)^{j+1}$, for all integers $n \geq 1$.

(d) Let X be an exponential distribution with mean $\frac{1}{\lambda}$, then $I(\bar{F}_{R_n}, \bar{F}) = \frac{n(n+1)}{2\lambda}$.

(e) Let X be a nonnegative random variable which has an Exponential-Inverse Gaussian distribution with survival function $\bar{F}(x) = e^{\frac{\alpha}{\beta}[1-\sqrt{1+2\beta x}]}$, $x \geq 0$, $\alpha, \beta > 0$, then for all integers $n \geq 1$, we obtain $I(\bar{F}_{R_n}, \bar{F}) = \frac{1}{\alpha^2} \sum_{j=0}^{n-1} j!(j+1)^2[\alpha + (j+2)\beta]$.

Proposition 2.2. *Let $a, b > 0$. For $n = 1, 2, \ldots$ it holds that*

$$I(\bar{F}_{aR_n+b}, \bar{F}_{aX+b}) = aI(\bar{F}_{R_n}, \bar{F}).$$

Proof. From (A.1) and noting that $\bar{F}_{aX+b}(x) = \bar{F}(\frac{x-b}{a})$, we have

$$I(\bar{F}_{aR_n+b}, \bar{F}_{aX+b}) = -\int_0^{+\infty} \sum_{j=0}^{n-1} \frac{[-\log \bar{F}_{aX+b}(x)]^{j+1}}{j!} \bar{F}_{aX+b}(x)dx = aI(\bar{F}_{R_n}, \bar{F}).$$

□

Proposition 2.3. *Let X be an absolutely continuous nonnegative random variable with $I(\bar{F}_{R_n}, \bar{F}) < \infty$, for $n \geq 1$. Then, we have*

$$I(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1} \frac{1}{j!} E\left[h_{j+1}(X)\right], \tag{A.2}$$

where

$$h_{j+1}(x) = \int_0^x \left[-\log \bar{F}(z)\right]^{j+1} dz, \quad x \geq 0.$$

Proof. From (A.1) and using Fubini's theorem, we obtain

$$\begin{aligned} I(\bar{F}_{R_n}, \bar{F}) &= \sum_{j=0}^{n-1} \int_0^{\infty} \frac{[-\log \bar{F}(z)]^{j+1}}{j!} \bar{F}(z)dz \\ &= \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^{\infty} \left[\int_z^{\infty} f(x)dx\right] [-\log \bar{F}(z)]^{j+1} dz \\ &= \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^{\infty} f(x) \left[\int_0^x [-\log \bar{F}(z)]^{j+1} dz\right] dx = \sum_{j=0}^{n-1} \frac{1}{j!} E\left[h_{j+1}(X)\right]. \end{aligned}$$

□

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

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

$$I(\bar{F}_{R_n}, \bar{F}) \geq \sum_{j=0}^{n-1} \frac{[\mathcal{E}(X)]^{j+1}}{j!}, \tag{A.3}$$

where

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

is the cumulative residual entropy (see [5]).

Proof. Since $\bar{F}(x) \geq (\bar{F}(x))^n$, for $n \geq 1$, we have

$$I(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^{\infty} \bar{F}(x)[-\log \bar{F}(x)]^{j+1} dx \geq \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^{\infty} [-\bar{F}(x) \log \bar{F}(x)]^{j+1} dx.$$

By noting that $g(x) = x^n, n \geq 1$, is a convex function, Jensen's inequality gives

$$I(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^{\infty} \bar{F}(x)[-\log \bar{F}(x)]^{j+1} dx \geq \sum_{j=0}^{n-1} \frac{1}{j!} \left(- \int_0^{\infty} \bar{F}(x) \log \bar{F}(x) dx \right)^{j+1},$$

which proof follows by recalling (A.3). □

Corollary 2.5. *Let X be a nonnegative absolutely continuous random variable. Then,*

$$I(\bar{F}_{R_n}, \bar{F}) \geq \sum_{j=0}^{n-1} \frac{C^{j+1}}{j!} \left[e^{(j+1)H(X)} \right],$$

where $C = \exp\left(\int_0^1 \log(x|\log x|) dx \right) = 0.2065$, and $H(X) = -\int_0^{+\infty} f(x) \log f(x) dx$ is the Shannon entropy of X.

Proof. The proof follows by recalling (A.3) and Proposition 4.2 of [2]. □

Proposition 2.6. *Let X be an absolutely continuous nonnegative random variable with finite mean $\mu = E(X)$. Then, we have*

$$I(\bar{F}_{R_n}, \bar{F}) \geq \sum_{j=0}^{n-1} \frac{1}{j!} \left[h_{j+1}(\mu) \right],$$

where $h_{j+1}(\mu)$ is defined in Proposition 2.3.

Proof. By noting that $h_{j+1}(.)$ is a convex function, applying Jensens inequality in (A.2) the proof is completed. □

Proposition 2.7. *Let X be a nonnegative random variable with absolutely continuous cumulative distribution function $F(x)$. Then for $n = 1, 2, ...$ we have*

$$I(\bar{F}_{R_n}, \bar{F}) \leq \sum_{j=0}^{n-1} \frac{1}{j!} \int_0^{\infty} [-\log \bar{F}(x)]^{j+1} dx.$$

Proof. By using (A.1) proof is easy. □

In the following, we obtain some results of $I(\bar{F}_{R_n}, \bar{F})$ and its connection with notions of reliability theory. For that we present the following definitions:

1. The random variable X is said to be smaller than Y according to stochastically ordering (denoted by $X \leq^{st} Y$) if $P(X \geq x) \leq P(Y \geq x)$ for all x. It is known that $X \leq^{st} Y \Leftrightarrow E(\phi(X)) \leq E(\phi(Y))$ for all increasing functions ϕ.

2. We say that X is smaller than Y in the hazard rate order, denoted by $X \leq^{hr} Y$, if $\frac{\bar{G}(x)}{\bar{F}(x)}$ is increasing with respect to x.

3. A non-negative random variable X is said to have increasing (decreasing) failure rate IFR (DFR) if $\lambda_F(x) = \frac{f(x)}{\bar{F}(x)}$ is increasing(decreasing) in x.

4. A non-negative random variable X with cdf F is said to have increasing(decreasing) failure rate average IFRA(DFRA) if $\frac{\lambda_F(x)}{x}$ is increasing (decreasing) function in $x > 0$. Note that IFR and DFR classes of distributions are included to IFRA and DFRA classes of distributions, respectively.

Proposition 2.8. *If X has the exponential distribution with mean $\mu = \frac{1}{\theta}$, then as the hazard rate is constant, we obtain the following property $I(\bar{F}_{R_n}, \bar{F}) = \frac{n(n+1)}{2}\mu$ which is an increasing function of n.*

Proposition 2.9. *Let X and Y be two nonnegative random variables with reliability functions $\bar{F}(x)$, $\bar{G}(x)$, respectively. If $X \leq^{hr} Y$ and X is DFR, then*

$$I(\bar{F}_{R_n}, \bar{F}) \leq I(\bar{G}_{R_n}, \bar{G}),$$

for $n = 1, 2, \ldots$.

Proof. It is well known that $X \leq_{hr} Y$ implies $X \leq_{st} Y$ (see [6]). Hence, we have

$$\bar{F}_{R_{j+2}} \leq \bar{G}_{\tilde{R}_{j+2}},$$

where $\bar{G}_{\tilde{R}_{j+2}}$ is the survival function of $\tilde{R}_{j+2}$. That is , $R_{j+2} \leq_{st} \tilde{R}_{j+2}$ holds. This is equivalent (see [6]) to have

$$E(\phi(R_{j+2})) \leq E(\phi(\tilde{R}_{j+2})),$$

for all increasing functions ϕ such that these expectations exist. Thus, if we assume that X is DFR and $\lambda_F(x)$ is its failure rate, then $\frac{1}{\lambda_F(x)}$ is increasing and we have

$$I(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1}(j+1)E_{R_{j+2}}\left(\frac{1}{\lambda_F(X)}\right) \leq \sum_{j=0}^{n-1}(j+1)E_{\tilde{R}_{j+2}}\left(\frac{1}{\lambda_F(X)}\right).$$

On the other hand, $X \leq_{hr} Y$ implies that the respective failure rate functions satisfy $\lambda_F(x) \geq \lambda_G(y)$. Hence, we have

$$\sum_{j=0}^{n-1}(j+1)E_{\tilde{R}_{j+2}}\left(\frac{1}{\lambda_F(X)}\right) \leq \sum_{j=0}^{n-1}(j+1)E_{\tilde{R}_{j+2}}\left(\frac{1}{\lambda_G(Y)}\right) = I(\bar{G}_{R_n}, \bar{G}).$$

Therefore, using both expressions we obtain $I(\bar{F}_{R_n}, \bar{F}) \leq I(\bar{G}_{R_n}, \bar{G})$. □

Proposition 2.10. *If X is IFRA (DFRA), then for $n = 1, 2, \ldots,$ we have*

$$I(\bar{F}_{R_n}, \bar{F}) \leq (\geq) \sum_{j=0}^{n-1} \frac{1}{j!} E\left[X\left(-\log \bar{F}(X)\right)^j\right]. \tag{A.4}$$

Proof. From (A.1), we have

$$I(\bar{F}_{R_n}, \bar{F}) = \sum_{j=0}^{n-1} \int_0^{+\infty} \frac{[-\log \bar{F}(x)]^j}{j!} [-\log \bar{F}(x)] \bar{F}(x) dx. \tag{A.5}$$

Now, since X is IFRA (DFRA), $\frac{-\log \bar{F}(x)}{x}$ is increasing (decreasing) with respect to $x > 0$, which implies that

$$-\bar{F}(x) \log \bar{F}(x) \leq (\geq) x f(x), \quad x > 0. \tag{A.6}$$

Hence, the proof is completed by noting (A.4) and (A.5). □

Proposition 2.11. *Let X and Y be two nonnegative random variables with survival function $\bar{F}(x)$ and $\bar{G}(x)$, respectively. If $X \leq^{hr} Y$, then for $n = 1, 2, \ldots,$ it holds that*

$$\frac{I(\bar{F}_{R_n}, \bar{F})}{E(X)} \leq \frac{I(\bar{G}_{R_n}, \bar{G})}{E(Y)}.$$

Proof. By noting that the function $h_{j+1}(x) = \int_0^x [-\log \bar{F}(z)]^{j+1} dz$ is an increasing convex function, under the assumption $X \leq^{hr} Y$, it follows by [6],

$$\sum_{j=0}^{n-1} \frac{1}{j!} \left[\frac{E\left[h_{j+1}(X)\right]}{E(X)}\right] \leq \sum_{j=0}^{n-1} \frac{1}{j!} \left[\frac{E\left[h_{j+1}(Y)\right]}{E(Y)}\right].$$

Hence, the proof is completed by recalling (A.2). □

Assume that X_θ^* denotes a nonnegative absolutely continuous random variable with the survival function $\bar{H}_\theta(x) = [\bar{F}(x)]^\theta, x \geq 0$. This model is known as a proportional hazards rate model. We now obtain the cumulative residual measure of inaccuracy between $\bar{H}_{R_n}$ and $\bar{H}$ as follows:

$$\begin{aligned} I(\bar{H}_{R_n}, \bar{H}) &= -\int_0^{+\infty} \bar{H}_{R_n}(x) \log\left(\bar{H}(x)\right) dx \\ &= \sum_{j=0}^{n-1} \theta^{j+1} \int_0^{+\infty} \frac{[-\log \bar{F}(x)]^{j+1}}{j!} [\bar{F}(x)]^\theta dx. \end{aligned} \tag{A.7}$$

Proposition 2.12. *If $\theta \geq (\leq) 1$, then for any $n \geq 1$, we have*

$$I(\bar{H}_{R_n}, \bar{H}) \leq (\geq) \sum_{j=0}^{n-1} (j+1) \theta^{j+1} \mathcal{E}_{j+1}(X),$$

where $\mathcal{E}_{j+1}(X)$ *is the generalized cumulative residual entropy of* X*, defined by [4] as*

$$\mathcal{E}_{j+1}(X) = \int_0^{+\infty} \frac{\bar{F}(x)[-\log \bar{F}(x)]^{j+1}}{(j+1)!}.$$

Proof. Suppose that $\theta \geq (\leq) 1$, then it is clear $[\bar{F}(x)]^{\theta} \leq (\geq)\bar{F}(x)$, and hence (A.7) yields

$$I(\bar{H}_{R_n}, \bar{H}) \leq (\geq) \sum_{j=0}^{n-1} (j+1)\theta^{j+1}\mathcal{E}_{j+1}(X).$$

□

References

[1] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N.(1992), *A first course in order statistics*, John Wiley, New York.

[2] Di Crescenzo, A. and Longobardi, M.(2009), On cumulative entropies.*Journal of Statistical Planning and Inference*, **139**, 4072-4087.

[3] Kumar, V. and Taneja, H.C.(2015), Dynamic cumulative residual and past inaccuracy measures. *J. Stat. Theory Appl*, **14**(4), 399-412.

[4] Psarrakos, G. and Navarro, J.(2013), Generalized cumulative residual entropy and record values. *Metrika*, **76**, 623-640.

[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] Shaked, M. and Shanthikumar, J.G.(2007), *Stochastic orders*, Springer, New York.

[7] Thapliyal, R. and Taneja, H.C.(2015), On residual inaccuracy of order statistics. *Statist. Probab. Lett*, **97**, 125-131.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Endogeneity problem in recurrent event data analysis

Tavakoli Dinani, Z. [1] and Rikhtehgaran, R. [2]

[1,2] Department of Mathematical Sciences, Isfahan University of Technology

Abstract

In this paper, we analyze recurrent event data in the framework of mixed-effects models to control the between and within subject variabilities among observations. We assume that longitudinal variables are informative for the analysis of recurrent data and thus are considered as covariates in the structure of underlying mixed-effects model. Since longitudinal variables are stochastic, they may be correlated with the random effects. This correlation causes biased estimates of regression coefficients. To solve this problem, we propose jointly modelling of longitudinal and recurrent event data in the framework of shared-random effects models. Bayes estimates of model parameters are achieved by the use of Gibbs sampling algorithm. A simulation study is conducted to show the performance of the proposed model.

Keywords: Endogeneity, Gibbs sampler, Longitudinal data, Mixed-effects models, Recurrent event data.

1 Introduction

In many research fields, subjects may experience the outcome of interest more than once over a period of observations. These types of outcomes are termed recurrent events. This sort of data is used in many applied fields such as medical studies and reliability. In clinical applications, recurrent events are often specific medical conditions, such as hospitalizations due to a particular disease,

[1] z.tavakoli@math.iut.ac.ir
[2] r_rikhtehgaran@cc.iut.ac.ir

cancer metastasis, cardiovascular events or epileptic seizures. In industrial applications, recurrent event data are generally related to the analysis of repairable systems.

Usually, through recording recurrent event data, measuring other variables than the time of events may also be important. Repeated measurements for these types of variables are termed longitudinal data. Recurrent event data and longitudinal data are often analyzed separately. Nevertheless, when recurrent variables are dependent on longitudinal variables, ignoring the dependence between them may lead to bias estimates or inefficient inferences [1]. In this article, we assume that longitudinal variables are informative for the analysis of time to event data. Thus, we enter longitudinal responses as covariates for the recurrent model. On the other hand, to control the between and within subject variabilities among time to event observations, we assume a mixed-effects modeling structure. In these models, random effects can capture a part of variability among recurrent event observations that are inherited by the variety of subjects.

However, in the analysis of mixed-effects models, it is usually assumed that random effects are independent of covariates. Violation of this assumption is critical and may cause biases for estimates of some regression coefficients [3]. The assumption of independence among random effects and time-invariant covariates are usually fulfilled, in practice. However, when there are time-varying or longitudinal covariates, this assumption may be violated. The reason of this violation is that longitudinal covariates are usually stochastic, thus they are probable to have non-zero correlation with both random effects and errors. These types of covariates are termed endogenous and the problem of having biased estimates due to the existence of endogeonus covariates is called endogeneity.

Thus, when longitudinal responses are considered as covariates in the structure of random-effects models for the analysis of recurrent event data, the endogeneity problem usually happens.

To solve this problem, we propose jointly modelling of the recurrent event data and longitudinal reponses in the framework of a mixed-effects model. In this approch, we consider a joint distribution for the random effects of the main model used for the analysis of recurrent data and the subsequent model of the longitudinal responses. Thus, through this approach, the dependence among random effects and longitudinal covariates has taken to account. By conducting a simulation study we show that the resulting model can considerably reduce the bias and improve the performance of the model.

The remainder of this paper is organized as follows. In Section 2, we introduce the proposed joint modeling structure. In Section 3, we drive all full conditional distributiones required for implementing the Gibbs sampler, to achive the Bayes estimates of parameters. The last section presents a simulation study.

2 Model specification

Let T_{ij} denotes the time to the j-th event, $j = 1, \cdots, J_i$, for the i-th subject, $i = 1, \cdots, n$. A common model for analyzing recurrent event data is given by [4],

$$T_{ij}^{*} = \mathbf{z}_{ij}'\boldsymbol{\delta} + b_i + \eta_{ij}, \tag{A.1}$$

where $T_{ij}^* = \ln T_{ij}$, $\mathbf{z}_{ij}$ is a p-dimensional vector of covariates at the j-th event for the i-th subject and $\boldsymbol{\delta}$ includes regression coefficients. Variables b_i's are random intercepts which are considered to capture the between subject variability and η_{ij}'s are residual terms. It is usually assumed that the random intercepts and residuals are independent and $b_i \overset{\text{iid}}{\sim} N(0, \sigma_b^2)$ and $\eta_{it} \overset{\text{iid}}{\sim} N(0, \sigma_\eta^2)$.

Usually in the analysis of recurrent data, measuring some longitudinal variables may also be informative. Putting the longitudinal covarites, y_{ij}'s, into the model can provides, in many cases, much information about the time of events. Thus, we consider the following model,

$$T_{ij}^* = z_{ij}'\boldsymbol{\delta} + \gamma y_{ij} + b_i + \eta_{ij}, \tag{A.2}$$

where γ denotes the effect of longitudinal covariates on the logarithm of the time to event variable, T_{ij}^*. Since the y_{ij} is a random variable, thus it may depend on the random effects and error terms. This dependency known as the endogeneity problem causes bias to the estimation of regression coefficients [5]. Indeed, there may be some features of subjects which are not entered to the model as time-invariant covarites and thus their effects are present in random intercepts, b_i's. In these situations, it is more probable that these omitted time-invariant covariates be correlated with the longitudinal covariate, y_{ij}. Thus, the longitudinal covariate, y_{ij}, is an endogeneous variable. To solve the endogeneity problem, we propose joint modelling of T_{ij}^* and y_{ij} in the framework of a shared random-effects model. Thus, in a vector form representation, we assume that,

$$\mathbf{T}_i^* | \boldsymbol{\delta}, b_i, \sigma_\eta^2 \overset{ind}{\sim} N_{J_i}(\mathbf{Z}_i\boldsymbol{\delta} + \gamma \mathbf{y}_i + b_i \mathbf{1}_{J_i}, \sigma_\eta^2 \mathbf{I}_{J_i}), \tag{A.3}$$

$$\mathbf{Y}_i | \boldsymbol{\beta}, \alpha_i, \sigma_\varepsilon^2 \overset{ind}{\sim} N_{J_i}(\mathbf{X}_i\boldsymbol{\beta} + \alpha_i \mathbf{1}_{J_i}, \sigma_\varepsilon^2 \mathbf{I}_{J_i}), \tag{A.4}$$

$$(\alpha_i, b_i)' \sim N_2(\mathbf{0}, \boldsymbol{\Xi}), \tag{A.5}$$

where $\mathbf{T}_i^* = (T_{i,1}^*, \cdots, T_{i,J_i}^*)'$, $\mathbf{Y}_i = (Y_{i,1}, \cdots, Y_{i,J_i})'$, $\mathbf{Z}_i = (\mathbf{z}_{i,1}', \cdots, \mathbf{z}_{i,J_i}')'$, $\mathbf{X}_i = (\mathbf{x}_{i,1}', \cdots, \mathbf{x}_{i,J_i}')'$, and $\mathbf{x}_{i,j}$ includes covariates values at the time of j-th event for the i-th subject. The vector $\boldsymbol{\beta}$ includes regression coefficients. Indeed, in this hierarchical representation of the model, the correlation between random effects, b_i, and the longitudinal covaraite, y_{ij}, is considered through the framework of the shared random-effects model. In this model, the dependence between random effects, α_i and b_i, reduces the bias created in these models. We illustrate this reduction in biases through conducting a simulation study in section 4.

3 Bayesian Estimation

In this section, we obtain parameter estimates in a Bayesian framework by the use of Markov Chain Monte carlo (MCMC) methods such as the Gibbs sampler [6]. To perform this sampler, the complete conditional posterior distributions are required to be driven. To do this, we use the following conditionally conjugate priors to obtain the parameter estimates: Inverse gamma priors, $IG(\tau_1, \tau_2)$ for σ_ε^2 and $IG(\eta_1, \eta_2)$ for σ_η^2, the $N(\boldsymbol{\lambda}, \boldsymbol{\Lambda})$ for $\boldsymbol{\theta}$, the $N(\boldsymbol{\varsigma}, \boldsymbol{S})$ for $\boldsymbol{\delta}$, the $N(\Upsilon_1, \Upsilon_2)$ for γ, the $N_2(\mathbf{0}', \boldsymbol{\Xi})$ for $\boldsymbol{\kappa}_i = (\alpha_i, b_i)'$ and $IW(\boldsymbol{\upsilon}, \varphi)$ for $\boldsymbol{\Xi}$. Then, by assuming the hierarchical representation of the model in Equations (2.3) to (2.5), the complete conditional posterior distributions, after doing some algebra, are obtained as follows

$$\boldsymbol{\beta} | \boldsymbol{\kappa}_i, \boldsymbol{\delta}, \gamma, \sigma_\varepsilon^2, \sigma_\eta^2, \boldsymbol{\Xi}, \mathbf{Y}, \mathbf{T}^* \sim \mathbf{N_p}(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta), \tag{A.1}$$

where, $\boldsymbol{\mu}_\beta = \boldsymbol{\Sigma}_\beta(\boldsymbol{\lambda}'\boldsymbol{\Lambda}^{-1} + \frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^{n}(\mathbf{y_i} - \alpha_\mathbf{i})'\mathbf{X_i})$ and $\boldsymbol{\Sigma}_\beta = (\boldsymbol{\Lambda}^{-1} + \frac{1}{\sigma_\varepsilon^2}\sum_{i=1}^{n}\mathbf{X_i'X_i})^{-\mathbf{1}}$.

For σ_ε^2 we have

$$\sigma_\varepsilon^2|\boldsymbol{\beta}, \boldsymbol{\kappa}_i, \boldsymbol{\Xi}, \boldsymbol{\delta}, \sigma_\eta^2, \mathbf{Y}, \mathbf{T}^* \sim \mathbf{IG}(\tau_\mathbf{1}^*, \tau_\mathbf{2}^*), \tag{A.2}$$

where $\tau_1^* = \tau_1 + \frac{1}{2}\sum_{i=1}^{n} J_i$ and $\tau_2^* = \tau_2 + \frac{1}{2}\sum_{i=1}^{n}(\mathbf{y_i} - \mathbf{X_i}\boldsymbol{\beta} - \alpha_\mathbf{i})'(\mathbf{y_i} - \mathbf{X_i}\boldsymbol{\beta} - \alpha_\mathbf{i})$.

For the vector parameter $\boldsymbol{\delta}$, we derive

$$\boldsymbol{\delta}|\boldsymbol{\beta}, \boldsymbol{\kappa}_i, \boldsymbol{\Xi}, \gamma, \sigma_\varepsilon^2, \sigma_\eta^2, \mathbf{Y}, \mathbf{T}^* \sim \mathbf{N_q}(\boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta), \tag{A.3}$$

where $\boldsymbol{\mu}_\delta = \boldsymbol{\Sigma}_\delta(\boldsymbol{\varsigma}'\boldsymbol{S}^{-1} + \frac{1}{\sigma_\eta^2}\sum_{i=1}^{n}(\mathbf{T_i^*} - \gamma\mathbf{y_i} - \mathbf{b_i})'\mathbf{Z_i})$ and $\boldsymbol{\Sigma}_\delta = (\boldsymbol{S}^{-1} + \frac{1}{\sigma_\eta^2}\sum_{i=1}^{n}\mathbf{Z_i'Z_i})^{-\mathbf{1}}$.

For the vector parameter γ we have

$$\gamma|\boldsymbol{\beta}, \boldsymbol{\kappa}_i, \boldsymbol{\delta}, \boldsymbol{\Xi}, \sigma_\varepsilon^2, \sigma_\eta^2, \mathbf{Y}, \mathbf{T}^* \sim \mathbf{N}(\mu_\gamma, \boldsymbol{\Sigma}_\gamma), \tag{A.4}$$

where $\mu_\gamma = \Sigma_\gamma(\frac{\Upsilon_1}{\Upsilon_2} + \frac{1}{\sigma_\eta^2}\sum_{i=1}^{n}(\mathbf{T_i^*} - \mathbf{b_i} - \mathbf{Z_i}\boldsymbol{\delta})'\mathbf{y_i})$ and $\Sigma_\gamma = (\frac{1}{\Upsilon_2} + \frac{1}{\sigma_\eta^2}\sum_{i=1}^{n}\mathbf{y_i'y_i})^{-\mathbf{1}}$.

Further, for the residual variance, σ_η^2, we show that

$$\sigma_\eta^2|\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{\kappa}_i, \boldsymbol{\Xi}, \gamma, \sigma_\varepsilon^2, \mathbf{Y}, \mathbf{T}^* \sim \mathbf{IG}(\eta_\mathbf{1}^*, \eta_\mathbf{2}^*), \tag{A.5}$$

where $\eta_1^* = \eta_1 + \frac{1}{2}\sum_{i=1}^{n} J_i$ and $\eta_2^* = \eta_2 + \frac{1}{2}\sum_{i=1}^{n}(\mathbf{T_i^*} - \mathbf{Z_i}\boldsymbol{\delta} - \gamma\mathbf{y_i} - \mathbf{b_i})'(\mathbf{T_i^*} - \mathbf{Z_i}\boldsymbol{\delta} - \gamma\mathbf{y_i} - \mathbf{b_i})$.

For each random intercept, we derive

$$\boldsymbol{\kappa}_i|\boldsymbol{\beta}, \boldsymbol{\delta}, \gamma, \sigma_\varepsilon^2, \sigma_\eta^2, \boldsymbol{\Xi}, \mathbf{Y}, \mathbf{T}^* \sim \mathbf{N}(\mathbf{K_1^*}, \mathbf{K_2^*}), \tag{A.6}$$

where $\boldsymbol{K}_1^* = \boldsymbol{K}_2^*(\mathbf{Y_i^*} - \mathbf{X_i^*}\boldsymbol{\beta}^*)'\boldsymbol{\Sigma}^{-1}\mathbf{W}^*$, $\boldsymbol{K}_2^* = (\boldsymbol{\Xi}^{-1} + \mathbf{W}^{*\prime}\boldsymbol{\Sigma}^{-1}\mathbf{W}^*)^{-\mathbf{1}}$, $\mathbf{Y_i^*} = (\mathbf{Y_i}, \mathbf{T_i^*})'$, $\boldsymbol{\beta}^* = (\boldsymbol{\beta}, \boldsymbol{\delta}^*)'$, $\boldsymbol{\delta}^* = (\boldsymbol{\delta}, \gamma)'$,

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_\varepsilon^2 I_{J_i} & 0I_{J_i} \\ 0I_{J_i} & \sigma_\eta^2 I_{J_i} \end{bmatrix}$$

and the matrices $\mathbf{X_i^*}$ and $\mathbf{W}^*$ are defined according to $\boldsymbol{\beta}^*$ and $\boldsymbol{\kappa}_i$, respectively.

Moreover, for $\boldsymbol{\Xi}$ we have

$$\boldsymbol{\Xi}|\boldsymbol{\beta}, \boldsymbol{\delta}, \gamma, \sigma_\varepsilon^2, \sigma_\eta^2, \boldsymbol{\kappa}_i, \mathbf{Y}, \mathbf{T}^* \sim \mathbf{IW}(\boldsymbol{\upsilon}^*, \varphi^*), \tag{A.7}$$

where $\boldsymbol{\upsilon}^* = \boldsymbol{\upsilon} + n$ and $\varphi^* = \varphi + \sum_{i=1}^{n}\boldsymbol{\kappa}_i\boldsymbol{\kappa}_i'$.

Now by sampling from the above distributions, iteratively and set appropriate burn-in until converges to stationary distributions and averaging the simulated values, the Bayes estimates are obtained.

4 Simulation Study

In this section, we conduct a simulation study to evaluate performance of the proposed model. The recurrent event data generating process is organized as follows

$$T_{ij}^* = \delta_0 + \delta_1 z_{ij} + \gamma_1 y_{i,j} + b_i + \eta_{ij}, \tag{A.1}$$

for $i = 1, ..., 200$ and $j = 1, ..., 20$, where $\ln T_{ij} = T^*_{ij}$ and

$$Y_{ij} = \beta_0 + \alpha_i + \varepsilon_{ij}, \tag{A.2}$$

where $\varepsilon_{ij} \overset{\text{iid}}{\sim} N(0, \sigma^2_\varepsilon)$ and $\eta_{ij} \overset{\text{iid}}{\sim} N(0, \sigma^2_\eta)$. We set intercept parameters, β_0 and δ_0 equal to 0 and regression coefficient δ_1 equals to 0.5. The parameter γ_1 is assumed to be 1 in order to prevent from large values for T^*_{ij}. We also assume that $z_{ij} \overset{\text{iid}}{\sim} N(5, \sigma^2_z)$ and $\boldsymbol{\kappa}_i = (\alpha_i, b_i)' \overset{\text{iid}}{\sim} N_2(\mathbf{0}', \boldsymbol{\Xi})$ so that

$$\boldsymbol{\Xi} = \begin{bmatrix} \sigma^2_\alpha & \rho\sigma_\alpha\sigma_b \\ \rho\sigma_\alpha\sigma_b & \sigma^2_b \end{bmatrix}.$$

where σ^2_α and σ^2_b are both set to 1 and $\rho = 0.8$. Also, it is assumed that the parameter σ^2_ε is equal to 1 , and σ^2_η and σ^2_z are equal to 0.01.

We fit the following models:

- Model M1: We consider the proposed joint modeling structure specified by Equations (2.3) to (2.5) in the framework of a shared- random effects model,

- Model M2: We consider the model specified by Equations (2.3) and (2.4) and assume that $\alpha_i \overset{\text{iid}}{\sim} N(0, \sigma^2_\alpha)$, $b_i \overset{\text{iid}}{\sim} N(0, \sigma^2_b)$ and α_i and b_i are independent.

Table 1: Bayesian estimation results of Models M1 and M2 for the simulated data set.

Model	β_0	σ^2_ε	δ_0	δ_1	γ	σ^2_η
M1	1.187	1.068	1.618	0.609	0.8151	103.8
	(0.659)	(0.037)	(0.0967)	(0.57)	(0.247)	(3.654)
M2	-1.739	1.068	-0.148	0.741	1.826	104.6
	(0.739)	(0.038)	(0.065)	(0.554)	(0.049)	(3.682)

* Bayesian standard deviations are given in parentheses.

In order to achieve the Bayes estimates of parameters, we use conditionally conjugates priors with large variances. Indeed we assume that the hyper-parameters for the inverse-Gamma prior of σ^2_ε and σ^2_η are equal to 0.01. The inverse-Wishart prior for $\boldsymbol{\Xi}$ has 2 degrees of freedom with a diagonal scale matrix with values equal to 0.01. Independent normal priors $N(0, \sigma^2_\beta)$, $N(0, \sigma^2_\delta)$ and $N(0, \sigma^2_\gamma)$ with σ^{-2}_β, σ^{-2}_δ and σ^{-2}_γ all equall to 0.01. Then, we use the OpenBUGs software [7] to run the Gibbs sampler algorithm. We use 100000 samples generated after 1000 burn-in. Results, after convergence is achieved, are reported in Table 1. It is seen that the regression coefficients in the model of the recurrent event data, i.e. δ_1 and γ, are highly biased in model M2. While the amount of the biases is drastically reduced in the proposed model, Model M1.

Also, the Deviance information criterion (DIC) is computed for each model [2]. Values of this criterion for Models M1 and M2 are 19010 and 19030, respectively. As is seen, the proposed model, Model M1, is better fitted to the simulated data set, in terms of DIC.

References

[1] Bhattacharjee, A. and Nath, D.C. (2014), Joint longitudinal and survival data modelling: an application in anti-diabets drug therapeutic effect, *Statistics In Transition*, **15**, 437-452.

[2] Lunn, D., Spiegelhalter, D. and Thomas, A. and Best, N. (2009), The BUGS project: Evolution, critique and future directions, *Statistics in Medicine*, **28** (25), 3049-3067.

[3] Palta, M. and Yao, T.J. (1991), Analysis of longitudinal data with unmeasured confounders, *Biometrics*, **47**, 13551369.

[4] Rathbun, S.L. and Shiffman, S. (2016), Mixed effects models for recurrent events data with partially observed time-varying covariates: Ecological Momentary Assessment of Smoking, *Biometrics*, **72**, 46-55.

[5] Rikhtehgaran, R., Kazemi, I. and Verbeke, G. (2016), A comparative study on estimation methods to deal with the endogeneity in linear randomintercept models with an extension, *Journal of Statistical Computation and Simulation*, **87**, 171-186.

[6] Robert, C. and Casella, G. (2004), Monte Carlo Statistical Methods, 2nd ed. Springer-Verlag, New York.

[7] Spiegelhalter, D.J., Thomas, A., Best, N.G. and Lunn, D. (2003), WinBUGS Version 1.4 Users Manual, *Biostatistics*, **51**, 151168.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

Some new results for weighted distributions

Velayati-Yazdsefat, P. [1] and Salehi, E. [2]

[1,2] Department of Basic Sciences, Birjand University of Technology

Abstract

In this paper, we derive some new results on preservation properties of the star shaped, moment-generating-function and Laplace transform orders under weighted distributions. Then we apply the obtained results to establish our results about preservation of the new better (worse) than used in star shaped; NBUS (NWUS), new better (worse) than used in the moment-generating-function; NBU_{mgf} (NWU_{mgf}) and new better (worse) than used in Laplace transform, NBUL (NWUL) classes for weighted distributions.

Keywords: Residual lifetime, Stochastic order, Aging class, Weighted distribution, Reliability.

1 Introduction

One of the most important and widely used the topics in industry and medicine is the weighted distributions. Weighted distributions are useful to model data in situations where the distribution of the observed data does not coincide with the original distribution of the data (see [7]). Sometimes, there is not exist a suitable sampling for recording observations and using classical sampling, because some reasons such as the invisibility of some events, partial destruction observations, sampling with unequal chances for the observations, etc. In these cases, observations are recorded according to some weight functions i.e, we use the weighted distributions instead of the parent distributions. So

[1]parham.amari@gmail.com
[2]salehi@birjandut.ac.ir

that the study of the weighted distributions has always been considering by many researchers. In the recent decades, several researchers have also worked to investigate the properties of weighted distributions in the context of stochastic orderings and aging classes. Among others, we refer to [2, 6, 5]. Recently, Izadkhah et al. in [3, 4] investigated the preservation of various stochastic orders and aging classes for weighted distributions under different scenarios.

This article is organized as follows. Some concepts and necessary preliminaries are provided in Section 2. New results on preservation of the star shaped, moment-generating-function and Laplace transform orders and NBUS (NWUS), NBU$_{mgf}$ (NWU$_{mgf}$) and NBUL (NWUL) classes under weighted distributions are obtained in Section 3.

2 Preliminaries

In this section, we present the required definitions and concepts to obtain the maim results. Let F and $\bar{F}$ denote the distribution function and reliability function of a non-negative random variable X, respectively. The residual lifetime of X at age $t \geq 0$, denoted by $X_t = \{X - t | X > t\}$, is the remaining life of X given survival at age t. The reliability function of X_t, denoted by $\bar{F}_t$, is

$$\bar{F}_t(x) = \frac{\bar{F}(x+t)}{\bar{F}(t)}.$$

In following, we present some stochastic orders considered in this note. [8] is comprehensive reference on these concepts and their properties.

Definition 1. *Let X and Y be two non-negative random variables with distribution functions F and G, and survival functions $\bar{F}$ and $\bar{G}$, respectively. X is said to be smaller than Y in the*

(i) usual stochastic order, denoted by $X \leq_{st} Y$, if for all x, $\bar{F}(x) \leq \bar{G}(x)$;

(ii) increasing concave order, denoted by $X \leq_{icv} Y$, if for all $t \geq 0$,

$$\int_0^t \bar{F}(x)dx \leq \int_0^t \bar{G}(x)dx;$$

(iii) star shaped order, denoted by $X \leq_{ss} Y$, if for all $t \geq 0$,

$$\int_t^\infty xf(x)dx \leq \int_t^\infty xg(x)dx;$$

(iv) moment-generating-function order, denoted by $X \leq_{mgf} Y$, if for all $s \geq 0$,

$$\int_0^\infty e^{sx}\bar{F}(x)dx \leq \int_0^\infty e^{sx}\bar{G}(x)dx;$$

(v) the Laplace transform order, denoted by $X \leq_{lt} Y$, if for all $s \geq 0$,

$$\int_0^\infty e^{-sx}\bar{F}(x)dx \leq \int_0^\infty e^{-sx}\bar{G}(x)dx;$$

Below, some of the important aging classes are presented (see [1]).

Definition 2. *Let X be a non-negative random variable with distribution function F and survival function $\bar{F}$. X or F is said to be*

(i) new better (worse) than used in second order, denoted by NBU(2) (NWU(2)), if for all $x, t \geq 0$,

$$\int_0^x \bar{F}(u+t)du \leq (\geq)\bar{F}(t)\int_0^x \bar{F}(u)du,$$

or equivalently, $X_t \leq_{icv} (\geq_{icv}) X$.

(ii) NBUS (NWUS) if for all $x, t \geq 0$,

$$\frac{1}{x}\int_x^\infty [\bar{F}(u+t) - \bar{F}(u)\bar{F}(t)]du \leq (\geq)[\bar{F}(x+t) - \bar{F}(x)\bar{F}(t)],$$

or equivalently, $X_t \leq_{ss} (\geq_{ss}) X$.

(iii) NBU_{mgf} (NWU_{mgf}), if for all $t, s \geq 0$,

$$\int_0^\infty e^{sx}\bar{F}(x+t)dx \leq (\geq)\bar{F}(t)\int_0^\infty e^{sx}\bar{F}(x)dx,$$

or equivalently, $X_t \leq_{mgf} (\geq_{mgf}) X$.

(iv) NBUL (NWUL), if for all $t, s \geq 0$,

$$\int_0^\infty e^{-sx}\bar{F}(x+t)dx \leq (\geq)\bar{F}(t)\int_0^\infty e^{-sx}\bar{F}(x)dx,$$

or equivalently, $X_t \leq_{lt} (\geq_{lt}) X$.

In the following, we introduce the weighted distribution. Let X and Y be two random variables with absolutely continuous distribution functions F and G, probability density functions f and g, and survival functions $\bar{F}$ and $\bar{G}$, respectively. Also, assume that $u_X = \sup\{x : F(x) < 1\}$ and $u_Y = \sup\{x : G(x) < 1\}$ are the respective upper bounds of X and Y, and $l_X = \inf\{x : F(x) > 0\}$ and $l_Y = \inf\{x : G(x) > 0\}$ are their corresponding lower bounds. Then, the weighted version of X (Y) with a non-negative weight function $w_1(x)$ ($w_2(x)$) for which $0 < E(w_1(X)) < \infty$ ($0 < E(w_2(Y)) < \infty$), is denoted by X^w (Y^w) which has the probability density function

$$f^{w_1}(x) = \frac{w_1(x)f(x)}{\eta_1} \quad \left(g^{w_2}(x) = \frac{w_2(x)g(x)}{\eta_2}\right),$$

where $\eta_1 = E(w_1(X))$ and $\eta_2 = E(w_2(Y))$. The distribution functions of X^w and Y^w are equal to

$$F^{w_1}(x) = \frac{1}{\eta_1}A_1(x)F(x), \qquad G^{w_2}(x) = \frac{1}{\eta_2}A_2(x)G(x),$$

and their corresponding survival functions equal to

$$\bar{F}^{w_1}(x) = \frac{1}{\eta_1}B_1(x)\bar{F}(x), \qquad \bar{G}^{w_2}(x) = \frac{1}{\eta_2}B_2(x)\bar{G}(x),$$

where $A_1(x) = E(w_1(X)|X \leq x)$, $A_2(x) = E(w_2(Y)|Y \leq x)$, $B_1(x) = E(w_1(X)|X > x)$ and $B_2(x) = E(w_2(Y)|Y > x)$. For more details see Jain et al. [5].

3 Main results

In the section, we obtain some new results on preservation of aging properties and stochastic orders under weighted distributions. First, we express two theorems about preservation of various stochastic orders under the weighted distributions.

Theorem 3.1. *Let w_i be increasing for some $i = 1, 2$, and let $\frac{w_2(x)}{\eta_2} \geq \frac{w_1(x)}{\eta_1}$ for all $x \geq 0$. If $X \leq_{ss} Y$, then $X^{w_1} \leq_{ss} Y^{w_2}$.*

Proof. Suppose that w_i is increasing for some $i = 1, 2$; and let $\frac{w_2(x)}{\eta_2} \geq \frac{w_1(x)}{\eta_1}$ for all $x \geq 0$. Then, for all $t \geq 0$, we have

$$\begin{aligned}\int_t^\infty (xg^{w_2}(x) - xf^{w_1}(x))\, dx &= \int_t^\infty x\left(\frac{w_2(x)g(x)}{\eta_2} - \frac{w_1(x)f(x)}{\eta_1}\right) dx \\ &\geq \int_t^\infty \frac{w_i(x)}{\eta_i}(xg(x) - xf(x))dx \\ &= \int_{-\infty}^\infty \Delta_i(x)(xg(x) - xf(x))dx, \qquad (A.1)\end{aligned}$$

where $\Delta_i(x) = \frac{1}{\eta_i}[w_i(x)I_{[t,\infty)}(x)]$, $i = 1, 2$. Since $\Delta_i(x)$ is non-negative and increasing function, and then by using Lemma 7.1(a) in Barlow and Proschan [1], the expression in (A.1) is non-negative. Hence, the proof is complete. □

Corollary 3.2. *If w_i is decreasing for some $i = 1, 2$, such that $\frac{w_2(x)}{\eta_2} \leq \frac{w_1(x)}{\eta_1}$ for all $x \geq 0$. Then*

$$X^{w_1} \leq_{ss} Y^{w_2} \Rightarrow X \leq_{ss} Y.$$

Theorem 3.3. *Let B_i be increasing for some $i = 1, 2$; and let $\frac{B_1(x)}{\eta_1} \leq \frac{B_2(x)}{\eta_2}$ for all $x \geq 0$.*

(i) If $X \leq_{mgf} Y$, then $X^{w_1} \leq_{mgf} Y^{w_2}$.

(ii) If $X \leq_{lt} Y$, then $X^{w_1} \leq_{lt} Y^{w_2}$.

Proof. We only prove case (i) because the other case is similar. $X \leq_{mgf} Y$ holds if and only if for all $s \geq 0$, $\int_0^\infty e^{sx}(\bar{G}(x) - \bar{F}(x))dx \geq 0$. Then

$$\begin{aligned}\int_0^\infty e^{sx}\left[\bar{G}^{w_2}(x) - \bar{F}^{w_1}(x)\right] dx &= \int_0^\infty e^{sx}\left[\frac{B_2(x)}{\eta_2}\bar{G}(x) - \frac{B_1(x)}{\eta_1}\bar{F}(x)\right] dx \\ &\geq \int_0^\infty \frac{B_i(x)}{\eta_i}\left[e^{sx}(\bar{G}(x) - \bar{F}(x))\right] dx \\ &= \int_{-\infty}^\infty \Delta_i(x)\left[e^{sx}(\bar{G}(x) - \bar{F}(x))\right] dx,\end{aligned}$$

where $\Delta_i(x) = \frac{1}{\eta_i}B_i(x)I_{[0,\infty)}(x)$. Similar to Theorem 3.1 and again using Lemma 7.1 (a) in Barlow and Proschan [1], the proof is complete. □

Corollary 3.4. *Suppose that B_i is decreasing for some $i = 1, 2$; and $\frac{B_1(x)}{\eta_1} \geq \frac{B_2(x)}{\eta_2}$ for all $x \geq 0$. Then*

(i) $X^{w_1} \leq_{mgf} Y^{w_2} \Rightarrow X \leq_{mgf} Y$.

(ii) $X^{w_1} \leq_{lt} Y^{w_2} \Rightarrow X \leq_{lt} Y$.

Now, we present some results on preservation of aging classes under weighted distributions.

Theorem 3.5. *Let A be decreasing, $\frac{1}{B(t)}E(w(X_t + t)|X \leq x) \geq \frac{A(x)}{\eta}$, for all $x, t \geq 0$. If X is NBU(2), then X^w is also NBU(2).*

Proof. X is NBU(2) if and only if $X_t \leq_{icv} X$ for all $t \geq 0$. Take $A_2(x) = A(x)$, $A_1(x) = E\left(w(X_t + t)|X \leq x\right)$. From Lemma 2.1(ii) in Izadkhah et al. [4], we have $E(w(X_t + t)) = B(t)$. Using Theorem 7 in Izadkhah et al. [3], $X_t \leq_{icv} X$ is equivalent to $(X_t)^{w_1} \leq_{icv} X^{w_2}$ for all $t \geq 0$. On the other hand, from Lemma 2.2 in Izadkhah et al. [4], $(X_t)^{w_1} \overset{\text{st}}{=} (X^w - t|X^w > t)$ for all $t \geq 0$. Thus, it follows that $(X^w - t|X^w > t) \leq_{icv} X^w$, for all $t \geq 0$, which means that X^w is NBU(2). □

Corollary 3.6. *If A is increasing and $\frac{1}{B(t)}E(w(X_t + t)|X \leq x) \leq \frac{1}{\eta}A(x)$, for all $x, t \geq 0$. Then $X^w \in$NBU(2) implies $X \in$NBU(2).*

Theorem 3.7. *Let w be decreasing (increasing) and let $\frac{w(x+t)}{B(t)} \leq (\geq)\frac{w(x)}{\eta}$ for all $x, t \geq 0$. Then X is NBUS if and only if X^w is NBUS.*

Proof. From Definition 2 (ii), $X_t \leq_{ss} X$ for all $t \geq 0$. By using Lemma 2.1 (ii) in Izadkhah et al. [4], we have,

$$E(w(X_t + t)) = B(t).$$

Taking $w_2(x) = w(x)$ and $w_1(x) = w(x+t)$, and from Theorem 3.1, we have $(X_t)^{w_1} \leq_{ss} X^{w_2}$ for all $t > 0$. Using Lemma 2.2 in Izadkhah et al. [4], $(X_t)^{w_1} \overset{\text{st}}{=} (X^w - t|X^w > t)$ for all $t > 0$. Therefore, for all $t > 0$,

$$(X^w - t|X^w > t) \leq_{ss} X^w.$$

and it means that X^w is NBUS. □

Theorem 3.8. *Let B be increasing (decreasing) and let $\frac{B(x+t)}{B(t)} \geq (\leq)\frac{B(x)}{\eta}$ for all $x \geq 0$. Then*

(i) X is NBU_{mgf} if and only if X^w is NBU_{mgf}.

(i) X is NBUL if and only if X^w is NBUL.

Proof. From Lemma 2.1 (ii) in Izadkhah et al. [4], $E(w(X_t + t)) = B(t)$. Taking $B_2(x) = B(x)$ and $B_1(x) = B(x + t)$. Since, X is NBU_{mgf}, then $X_t \leq_{mgf} X$ for all $s \geq 0$ and from Theorem 3.3, $(X_t)^{w_1} \leq_{mgf} X^{w_2}$ for all $s \geq 0$. Also, using Lemma 2.2 in Izadkhah et al. [4], we have for all $t \geq 0$, $(X_t)^{w_1} \overset{\text{st}}{=} (X^w - t|X^w > t)$. Therefore, it follows that $(X^w - t|X^w > t) \leq_{mgf} X^w$, for all $t \geq 0$, and hence X^w is NBU_{mgf}.

The proof of the other case can be obtained from Definition 2 (iii), and using Theorem 3.3 and Lemma 2.2 in Izadkhah et al. [4] in the same way as in the proof of case (i). □

References

[1] Barlow, R. E. and Proschan, F. (1975), *Statistical Theory of Reliability and Life Testing: Probability Models*. Holt, Rinhart, and Winston.

[2] Bartoszewicz, J. and Skolimowska, M. H. (2006), Preservation of classes of life distributions and stochastic order under weighting. *Statistics and Probability Letters*,**76**, 587-596.

[3] Izadkhah, S. Amini, M. and Mohtashami Borazadaran, G. R. (2015), On preservation under univariate weighted distributions. *Applictions of mathematics*, **60**, 6809-6818.

[4] Izadkhah, S., Rezaei, A. H. Amini, M. and Mohtashami Borzadaran, G. R. (2013), A General Approach for Preservation of Some Aging Classes Under Weighting. *Communications in Statistics- Theory and Methods*, **42**, 1899-1909.

[5] Jain, K., Singh, H. and Bagai, I. (1989), Relations for reliability measures of weighted distributions. *Communications in Statistics-Theory and Methods*, **18**, 4393-4412.

[6] Misra, N., Gupta, N. and Dhariyal, I. D. (2008), Preservation of some aging properties and stochastic orders by weighted distributions. *Communications in Statistics- Theory and Methods*, **37**, 627-644.

[7] Rao, C. R. (1965), On discrete distributions arising out of methods of ascertainment. *Sankhy Ser. A*, **27**, 311-324.

[8] Shaked, M. and Shanthikumar, J. G. (2007), *Stochastic orders*, Springer, New York.

Third Seminar on
Reliability Theory and its Applications
May 16-17, 2017

An algorithm to assess t-signature

Zarezadeh, S. [1] and Siavashi, M. [2]

[1] Department of Statistics, Shiraz University
[2] Department of Computer Science, Engineering, and IT, School of Electrical and Computer Engineering, Shiraz University

Abstract

Due to the importance of *signature* vector in studying the reliability of networks, some researchers explored the problem of signature estimation. The signature is used when at most one link may fail at each time instant. Recently, the concept of t-signature has been defined to get the reliability of network for the case where the failure of more than one link is possible at each time instant. The t-signature is a probability vector and depends only on the network structure. In this paper, we propose an algorithm to compute the t-signature. The performance of the proposed algorithm is evaluated for some networks.

Keywords: Network reliability, BFS algorithm, Signature.

1 Introduction

In recent years, the reliability of networks has been studied by many researchers. A network is defined as a collection of nodes (vertices), and links (edges). Rail stations, telecommunication centers, and computers are examples of nodes, and rail ways, communication channels, and the cables between the computers are examples of links. Some nodes of network are considered as *terminals set* and the states of network are usually defined based on the connection between the

[1] s.zarezadeh@shirazu.ac.ir
[2] m.siavashi@cse.shirazu.ac.ir

terminals. In the sequel, We assume that the network has two states: *up*, and *down* and the network nodes are absolutely reliable. In other words, the links are subject to failure and the network state may change in the process of links failure.

There are different models to get the reliability of network. One of these models, which has been considerably explored, represents the reliability of network as a mixture of the reliability of ordered links lifetimes; see, e.g., [3] and [2]. Let T be the lifetime of a network having n links with independent and identically distributed (i.i.d.) lifetimes $T_1, \ldots, T_n$. Under the assumption that there are not ties between the occurrence times of links failures, $P(T_i = T_j) = 0$, $i \neq j$, the network reliability is written as

$$P(T > t) = \sum_{i=1}^{n} s_i P(T_{i:n} > t), \quad t > 0,$$

in which $T_{i:n}$ is the ith ordered links lifetimes and $s_i = P(T = T_{i:n})$, $i = 1, \ldots, n$. The probability vector $\mathbf{s} = (s_1, \ldots, s_n)$ is called *signature* which does not depend on the random mechanism of links failures and is only determined based on the network structure. Similar mixture representation is hold for the reliability of network when we know the links of network are destroyed based on a counting process; see, [2] and [5].

The signature has also combinatorially definition as follows:
Consider a network with n links where $\pi = (e_{i_1}, e_{i_2}, \ldots, e_{i_n})$ denotes a permutation of the ordinal number of network links. Let all links be in up state and moving from left to right of permutation, turn the state of each link from up to down. By the assumption that all permutations are equally probable, the signature is defined as $\mathbf{s} = (s_1, \ldots, s_n)$ where

$$s_i = \frac{n_i}{n!}, \qquad i = 1, \ldots, n,$$

where n_i is the number of permutations in which the failure of the ith link causes to change the network state to down state.

As mentioned, the notion of signature is applicable in studying the reliability of networks that more than one link can not be fail at each time instant; see, e.g. [1], [4]. Zarezadeh et al. [6] studied the reliability of two-state networks under the assumption that their links are subject to shocks such that the failure of more than one link is possible at the occurrence time of each shock. Let the network be subject to shocks. Each shock may lead to link failures and the network finally fails by one of these shocks. Assume that $N(t)$ denotes the number of links that fail up to time t, and T is the network lifetime. Under the assumption that the process of occurrence of the shocks is independent of the number of failed links, it was shown that

$$P(T > t) = \sum_{i=1}^{n} s_i^{\tau} P(N(t) \leq i - 1). \tag{A.1}$$

where $\mathbf{s}^{\tau} = (s_1^{\tau}, s_2^{\tau}, \ldots, s_n^{\tau})$ is called *t-signature* vector. The t-signature has the following combinatorial definition:
If it is possible to fail more that one link at instant time, then the way of the order of links failure is different from ordinal permutation applied in definition of signature. All ways of links failure are

obtained in two stages: first we obtain all partitions of $\{1,\ldots,n\}$ and then all permutation of each partition are considered. Therefore, the number of ways of order of links failures, denoted by n^*, has been obtained as

$$n^* = \sum_{j=1}^{n} \sum_{k=0}^{j} \binom{j}{k} (-1)^k (j-k)^n; \tag{A.2}$$

see Lemma 1 of [6].

Let the discrete random variable M denote the minimum number of links that their failures cause the network failure in each way of links failure order. Clearly, M takes values on $\{1,2,\ldots,n\}$. Suppose n_i is the number of ways of the order of link failures in which $M=i$. Assuming that all the number of ways of the order of link failures are equally likely, the t-signature vector associated to the network is defined as $\mathbf{s}^\tau = (s_1^\tau,\ldots,s_n^\tau)$ where

$$s_i^\tau = \frac{n_i}{n^*}, \qquad i=1,\ldots,n.$$

It is notable that t-signature, similar to the concept of signature, depends only on the structure of the network and does not depend on the random mechanism of the link failures.

In this paper, we propose an algorithm to compute the t-signature vector. For some networks, the performance of the algorithm is examined.

2 The proposed algorithm

In this section, we give an algorithm to obtain t-signature of a network. Let us first introduce the following notations.

Notations:

results:	An array to save the result of algorithm
all_orders:	All ways of order of links failures
order:	An element of *all_orders*. Note that *order* is a vector whose elements may be subsets of $\{1,\ldots,n\}$.
source_node	The source node (terminal)
destination_node	The destination node (terminal)
has_route:	A function which surveys the connection between two terminals
removed:	An array to save the removed links
failure:	The element of *order* which fails.
all_permutations:	A function which computes all possible permutation
perm:	An element of all_permutations
M	The minimum number of links that their failures cause the network failure in each order.

The adjacency matrices usually contain few nonzero elements. These matrices are called sparse and it is better to store only the nonzero elements instead of storing all elements of the

matrix. Besides memory efficiency, this makes the calculations go much faster. We can indicate the structure of network by such method.

Now, we give the algorithm.

Algorithm:

```
results := [ ]
for each order in all_orders:
  removed := [ ]
  M := 0
  for each failure in order:
    if not graph.has_route(source_node,destination_node,removed+failure):
      all_Ms := [ ]
      for each permutation in all_permutations(failure):
        m := 1
        without_edges := removed[:]
        for each edge in permutation:
          without_edges := without_edges + [edge]
          if graph.has_route(source_node, destination_node, without_edges):
            m := m + 1
          else
            break
        all_Ms := all_Ms + [m]
      M = M + minimum(all_Ms)
      results := results + [(order, M)]
      break
    M := M + length(failure)
```

removed := removed + failure

As seen in Table 1, the number of n^* is very large even for $n = 9$. Then we need to estimate t-signature for large value of n^*. To this, we select the samples using the method of probability proportional to size sampling.

Table 1: The amount of n^* for different values of n.

n	$n!$	n^*
2	2	3
3	6	13
4	24	75
5	120	541
6	720	4,683
7	5,040	47,293
8	40,320	545,835
9	362,880	7,087,261
10	3,628,800	102,247,563
11	39,916,800	1,622,632,573
12	479,001,600	28,091,567,595

3 Experimental results

In this section, we examine the algorithm for some networks. The computer program is developed in Python v2.7.10. To run this program we use an intel core i5-4200U processor 1.6 GHz and 8 GB RAM under Windows 8 64-bit.

Example 3.1. Consider the bridge network with graph as shown in Figure 1. The network is defined to be in up state if and only if nodes a and d are connected. From Table 1, $n^* = 541$ and hence we can obtain the exact value of t-signature.

Using the proposed algorithm, t-signature of the network is obtained as

$$\mathbf{s}^\tau = (0, \frac{154}{541}, \frac{309}{541}, \frac{78}{541}, 0).$$

Example 3.2. Consider the network with graph depicted in Figure 2. This network has 5 nodes and 8 links in which the network is in up state if and only if there is a path between nodes a and d. As seen in Table 1, $n^* = 545,835$. To explore the accuracy of estimation, we get both exact t-signature and estimated signature.

Table 2 represents the exact value of t-signature and estimated t-signature with sample sizes $n = 30000$ and $n = 450000$, respectively. Figure 3 shows the accuracy of estimation even for $n = 30000$.

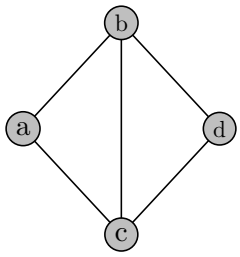


Figure 1: The bridge network.

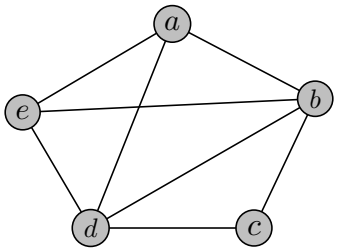


Figure 2: Network with 5 nodes, 8 links, and terminals set $\{a, d\}$

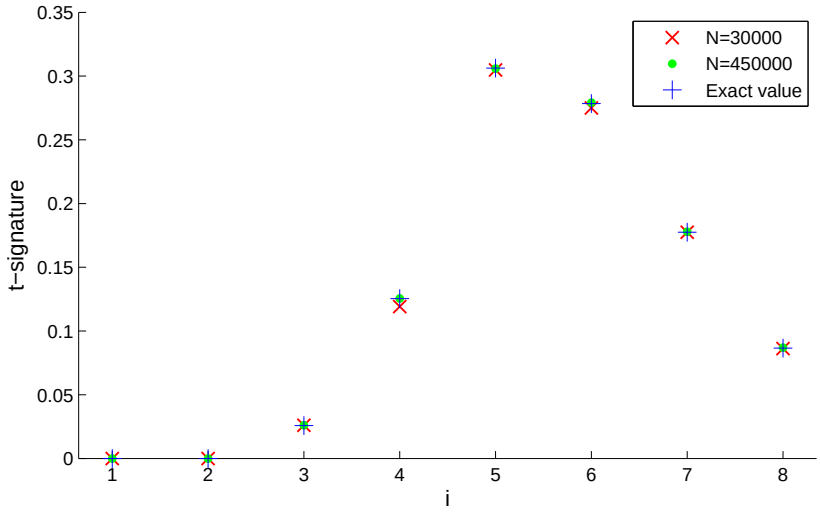


Figure 3: The plot of t-signature for Example 3.2.

References

[1] Gertsbakh, I. and Shpungin, Y. (2009), *Models of Network Reliability. Analysis, Combinatorics, and Monte Carlo*, CRC Press.

مجموعه مقالات سومین سمینار تخصصی

نظریه قابلیت اعتماد و کاربردهای آن

۲۶ و ۲۷ اردیبهشت ۱۳۹۶

گروه آمار دانشگاه فردوسی مشهد