

توابع زیان متعادل^۱

اکبر اصغرزاده^۲

چکیده

این مقاله، مروری است بر توابع زیان متعادل که هم‌زمان دو محک خوبی برازش^۳ و دقت برآورده را معکس می‌کند. از دیدگاه بیزی برآورد پارامتر میانگین و بردار میانگین با این توابع زیان مورد مطالعه قرار می‌گیرد. مجاز بودن^۴ کلاسی از برآوردهای خطی از میانگین نمونه، در مدل نرمال مورد بررسی قرار می‌گیرد.

واژه‌های کلیدی: توابع زیان متعادل، برآوردهای بیزی، مجاز بودن یک برآوردهای.

بیزی مورد بحث و بررسی قرار گرفته است. روذریگز^۵ و زلنر^۶ مسئله برآوردهای میانگین زمان لازم برای شکست را با تابع زیان متعادل مطالعه کردند. چانگ و کیم^۷ [۱] برآورد هم‌زمان میانگین نرمال چندمتغیره را وقتی که تابع زیان متعادل است، مطالعه کرده و برآوردهای خطی از نوع جیمز و استاین^۸ [۱۱] را برای این تابع زیان بدست آورند. برای مطالعه بیشتر روی مسئله برآوردهای تابع زیان متعادل می‌توان به دی و همکاران^۹ [۲]. سنجری و اصغرزاده^{۱۰} [۶ و ۷] و گروبر^{۱۱} [۲] مراجعه کرد.

در بخش ۲ این مقاله، تابع زیان متعادل تعریف و انواع مختلف آن بیان می‌شود. در بخش ۳، برآوردهای میانگین جامعه را از دیدگاه بیزی مورد بررسی قرار داده و مجاز بودن کلاسی از برآوردهای خطی میانگین نمونه، مورد مطالعه قرار می‌گیرد.

۲. تابع زیان متعادل

فرض کنید که $\mathbf{X} = (X_1, X_2, \dots, X_n)$ یک بردار از مشاهدات باشد

$$\mathbf{X} = \theta \mathbf{1} + \mathbf{u} \quad (1)$$

که θ میانگین مشترک^{۱۲} X_i ها، $\mathbf{1}$ بردار $1 \times n$ تابی با عناصر ۱ و یک بردار $n \times 1$ تابی خطأ باشد. افرون بر این، فرض کنید که برای پارامتر θ یک چگالی پسین $\pi(\theta | \Delta)$ که $(\mathbf{X}, \pi) = \Delta$ و π یک چگالی پیشین برای θ است، وجود دارد. فرض های مختلفی را می‌توان درباره توزیع \mathbf{X} و π در نظر گرفت.

Rodrigues -۸
Chung and Kim -۹
James and Stein -۱۰
Dey et.al. -۱۱
Gruber -۱۲

۱. مقدمه

در تئوری تصمیم‌آماری، توابع زیان معمولاً تأکید روی دقت برآورده دارند، در حالیکه خوبی برازش نیز یک محک مهمی است. از ایسو، زلنر^۸ [۸] در سال ۱۹۹۴ توابع زیان متعادل را معرفی نمود که هم‌زمان دو محک خوبی برازش و دقت برآورده را معکس می‌کنند. در گذشته، توابع زیان منعکس کننده یکی از این دو محک اما نه هر دو، برای آسالیز میانگین‌ها، رگرسیون و دیگر مسائل برآورده بکار رفته‌اند. برای مثال، برآوردهای کمترین مربیات^۹، خوبی برازش را منعکس می‌کنند، در حالیکه، استفاده از تابع درجه ۲، صرفاً دقت برآورده را نشان می‌دهد. همانطور که شناخته شده است، تأکید تها روحی محک دقت یک برآورده‌گر، به عنوان مثال میانگین مربع خطأ، اغلب منجر به برآوردهای اریب می‌شود. از طرف دیگر، استفاده از محک خوبی برازش، همانند مجموع مربعات باقیمانده‌ها^{۱۰} در یک مسئله رگرسیون، منجر به برآوردهای می‌شود که خوبی برازش را می‌دهد و نااریب است. با این وجود، چنین برآوردهای ممکن است به دقت یک برآورده‌گر اریب با مقدار اریبی کم نباشد. از این‌رو، معرفی توابع زیانی که هم‌زمان دو محک خوبی برازش (یا فقسدان اریبی) و دقت برآورده را لحاظ کنند، ضروری به نظر می‌رسد. توابع زیان متعادل، این ضرورت را برآورده می‌سازند.

در مقاله [۸] برآوردهای میانگین و بردار ضرایب رگرسیونی از دیدگاه

Balanced loss functions -۱

۲ گروه امل، دانشگاه مازندران

Goodness of fit -۲

Admissibility -۴

Zellner -۵

least squares estimation -۶

Sum of squared residuals -۷

۳. برآورد میانگین

در این بخش ابتدا برآورد بیزی پارامتر میانگین و بردار میانگین را با تابع زیان متعادل مورد مطالعه قرار می‌دهیم. سپس مجاز بودن کلاسی از برآوردهای خطی به فرم $a\bar{X} + b$ در حالتیکه توزیع جامعه نرمال است مورد بررسی قرار می‌گیرد.

۳.۱. برآورد بیزی میانگین

فرض کنید X_1, X_2, \dots, X_n یک نمونه تصادفی از جامعه‌ای با میانگین θ باشد (بر اساس مدل ۱). هدف پیدا کردن برآوردهای بیز پارامتر θ تحت تابع زیان L_1 می‌باشد. ریسک پسین برآوردهای $\hat{\theta}$ می‌شود.

$$E[L_1(\hat{\theta}, \theta) | \mathbf{X}] = \frac{\omega}{n} \sum_{i=1}^n (X_i - \hat{\theta})^2 + (1-\omega) E[(\hat{\theta} - \theta)^2 | \mathbf{X}]$$

که $(X_1, X_2, \dots, X_n) = \mathbf{X}$. برای یافتن برآوردهای بیز، کافی است $\hat{\theta}$ را طوری پیدا کنیم که ریسک پسین کمترین مقدار را داشته باشد. با حل معادله

$$\frac{d}{d\hat{\theta}} E[L_1(\hat{\theta}, \theta) | \mathbf{X}] = 0$$

برآوردهای بیز $\hat{\theta}$ به صورت زیر به دست می‌آید.

$$\hat{\theta}_B = \omega \bar{X} + (1-\omega) E(\theta | \mathbf{X}) \quad (3)$$

که در آن $E(\theta | \mathbf{X})$ میانگین پسین می‌باشد. همانطوریکه مشاهده می‌شود، $\hat{\theta}_B$ ترکیبی از برآوردهای غیر بیز \bar{X} و برآوردهای بیز $E(\theta | \mathbf{X})$ می‌باشد.

در حالتی که θ یک بردار میانگین و عناصر نمونه برداری باشند، به کمک مشتق گیری برداری برآوردهای بیز، مشابه با حالت قبل به صورت زیر به دست می‌آید [۷].

$$\hat{\theta} = \omega \bar{X} + (1-\omega) E(\theta | \mathbf{X})$$

که در آن \hat{X} بردار میانگین نمونه و $E(\theta | \mathbf{X})$ بردار میانگین پسین است.

۳.۲. چند مثال

مثال ۱: فرض کنید X_1, X_2, \dots, X_n یک نمونه تصادفی از توزیع نرمال $N(\theta, \sigma^2)$ باشد بطوریکه توزیع پیشین θ نیز نرمال با میانگین μ و واریانس σ^2 است. در آن صورت به سادگی می‌توان نشان داد که توزیع پسین θ نرمال با میانگین و واریانس

برای پارامتر θ یک چگالی پسین $\pi(\theta | \Delta)$ که $\Delta = (\mathbf{X}, \pi)$ است، وجود دارد. فرض های مختلفی را می‌توان درباره توزیع \mathbf{X} و π در نظر گرفت. تابع زیان متعادل برای برآوردهای $\hat{\theta}$ به صورت زیر معروفی می‌شود.

$$L_1(\hat{\theta}, \theta) = \frac{\omega}{n} (\mathbf{X} - \hat{\theta}\mathbf{I})(\mathbf{X} - \hat{\theta}\mathbf{I})' + (1-\omega)(\hat{\theta} - \theta)^2$$

$$= \frac{\omega}{n} \sum_{i=1}^n (X_i - \hat{\theta})^2 + (1-\omega)(\hat{\theta} - \theta)^2 \quad (2)$$

که در آن $1 \leq \omega \leq 0$ و $\hat{\theta}$ یک برآوردهای θ است. اولین جمله در رابطه (۲)، خوبی برازش را نشان می‌دهد، در حالیکه دومین جمله دقت برآوردهای را نشان می‌دهد. در رابطه (۲)، بجای تابع زیان درجه دوم $(\hat{\theta} - \theta)^2$ ، می‌توان از تابع زیان دیگری همچون $| \hat{\theta} - \theta |$ استفاده نمود. دقت کنید که برای $\omega = 0$ ، این تابع زیان به یک تابع زیان درجه دوم تبدیل می‌شود، در حالیکه برای $\omega = 1$ ، این تابع صرفاً خوبی برازش را نشان می‌دهد.

تابع زیان متعادل در (۲) را می‌توان به صورت زیر تعمیم و گسترش داد.

$$L_\alpha(\hat{\theta}, \theta) = \omega q(\theta) \frac{\sum_{i=1}^n (X_i - \hat{\theta})^2}{n} + (1-\omega) q(\theta)(\hat{\theta} - \theta)^2$$

که در آن $1 \leq \omega \leq 0$ و $q(\theta)$ تابعی مثبت از θ است که تابع وزن α نامیده می‌شود. در حالتی خاص که $q(\theta) = 1$ باشد، این تابع زیان به تابع زیان L_1 در (۲) تبدیل می‌شود. برای حالت بسرباری، فرض کنید $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ که $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ یک نمونه تصادفی $\mathbf{X}_\alpha = (X_{\alpha 1}, X_{\alpha 2}, \dots, X_{\alpha n})'$ باشد. این تابع زیان به از جامعه‌ای با بردار میانگین $(\theta_1, \theta_2, \dots, \theta_p)'$ باشد. برای برآوردهای بردار میانگین $\hat{\theta}$ ، تابع زیان متعادل به صورت زیر در نظر گرفته می‌شود.

$$L_\alpha(\hat{\theta}, \theta) = \frac{\omega}{N} \sum_{i=1}^N (\mathbf{X}_i - \hat{\theta})' Q (\mathbf{X}_i - \hat{\theta})$$

$$+ (1-\omega)(\hat{\theta} - \theta)' Q (\hat{\theta} - \theta)$$

که $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_p)'$ یک برآوردهای θ و Q یک ماتریس معین مثبت از بعد $p \times p$ است. دقت کنید در حالت خاص که $\omega = 0$ و $Q = I$ باشد، تابع زیان درجه دوم $, \|\hat{\theta} - \theta\|^2$ حاصل خواهد شد.

کنیم در آن صورت توزیع پسین، نرمال چندمتغیره با بردار میانگین $\bar{\mathbf{X}}$ و ماتریس کواریانس \mathbf{I} خواهد بود.
 $\frac{1}{N+m} \bar{\mathbf{X}}$ بنابراین برآوردگر بیز می‌شود:

$$E(\theta | \mathbf{X}) = \begin{bmatrix} \frac{n\bar{X} + \mu}{\sigma^2 + \tau^2} \\ \vdots \\ \frac{n\bar{X} + \mu}{\sigma^2 + \tau^2} \end{bmatrix}$$

$$E(\theta | \mathbf{X}) = \omega \bar{\mathbf{X}} + (1 - \omega) \hat{\theta}_B = \bar{\mathbf{X}} \frac{\omega + Nm}{1 + Nm}$$

۳.۳ مجاز بودن برآوردگر $a\bar{X} + b$

با توجه به مثالهای فوق، برآوردهای بیز به فرم خطی $a\bar{X} + b$ می‌باشد. در این بخش مجاز بودن این کلاس از برآوردهای خطی تحت مدل (۱) مورد مطالعه قرار می‌گیرد.تابع ریسک برآوردگر $a\bar{X} + b$ تحت تابع زیان متعادل L به صورت زیر می‌باشد،

$$R(\theta, a\bar{X} + b) = \frac{\omega}{n} \sum_{i=1}^n E(X_i - a\bar{X} - b)^2 + (1 - \omega) E(a\bar{X} + b - \theta)^2$$

به سادگی می‌توان نشان داد که این تابع ریسک را می‌توان به صورت زیر نوشت:

$$R(\theta, a\bar{X} + b) = [(a - 1)\theta + b]^2$$

$$+ \frac{\sigma^2}{n} [(a - \omega)^2 + \omega(n - \omega)]$$

به کمک تابع فوق می‌توان ریسک برآوردگر بیز را پیدا نموده و همچنین ناحیه‌ای که در آن $a\bar{X} + b$ مجاز است، مشخص کرد.

قضیه ۱. فرض کنید X_1, X_2, \dots, X_n یک نمونه تصادفی از جامعه‌ای (نه لزوماً نرمال) با میانگین θ و واریانس σ^2 باشد در آن صورت $a\bar{X} + b$ برای θ تحت تابع زیان L غیر مجاز است هرگاه یکی از شرایط زیر برقرار باشد:

الف) $a > 1$

ب) $a < \omega$

ج) $a = 1, b \neq 0$

اثبات: کافی است در هر حالت برآوردهای معرفی شود که از لحاظ ریسک بر $a\bar{X} + b$ برتری داشته باشد.
 (الف) اگر $a > 1$ ، آنگاه

$$(a - \omega)^2 > (1 - \omega)^2$$

$$\text{Var}(\theta | \mathbf{X}) = \frac{1}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}$$

می‌باشد. از اینرو، برآوردگر بیز در (۳) به صورت زیر در می‌آید:

$$\hat{\theta}_B = \frac{n\tau^2 + \sigma^2\omega}{n\tau^2 + \sigma^2} \bar{X} + (1 - \omega) \frac{\mu\tau^2}{n\tau^2 + \sigma^2} \quad (5)$$

که به فرم خطی $a\bar{X} + b$ می‌باشد.

مثال ۲: فرض کنید X_1, X_2, \dots, X_n یک نمونه تصادفی از توزیع برنولی با تابع چگالی احتمال

$$f(x | p) = p^x (1 - p)^{1-x}, \quad x = 0, 1, \quad 0 < p < 1.$$

باشد. اگر توزیع پیشین p ، توزیع بتا $Beta(\alpha, \beta)$ با چگالی

$$\pi(p) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1} (1-p)^{\beta-1}, \quad 0 < p < 1, \alpha, \beta > 0$$

فرض شود. در آن صورت توزیع پسین p ، توزیع $Beta(\alpha + \sum_{i=1}^n x_i, n - \sum_{i=1}^n x_i + \beta)$ خواهد بود. از اینرو برآوردگر بیز می‌شود:

$$\begin{aligned} \hat{p}_B &= \omega \bar{X} + (1 - \omega) E(p | \mathbf{X}) \\ &= \omega \bar{X} + (1 - \omega) \frac{\alpha + \sum_{i=1}^n x_i}{n + \alpha + \beta} \\ &= \frac{n + \alpha\omega + \beta\omega}{n + \alpha + \beta} \bar{X} + (1 - \omega) \frac{\alpha}{n + \alpha + \beta} \end{aligned}$$

مثال ۳: فرض کنید که X_1, X_2, \dots, X_n یک نمونه تصادفی از یک توزیع نرمال چندمتغیره با بردار میانگین $\theta_p = (\theta_1, \dots, \theta_p)$ و ماتریس کواریانس I^{σ^2} باشد. اگر توزیع پیشین θ را نیز یک توزیع نرمال چندمتغیره با بردار میانگین θ و ماتریس کواریانس mI فرض

قضیه ۲: اگر X_1, X_2, \dots, X_n یک نمونه تصادفی از (θ, σ^2) باشد در آن صورت $a\bar{X} + b$ برای θ مجاز است هرگاه $a = 1, b = 0$ یا $\omega \leq a < 1, b \in \mathbb{R}$ داشته باشیم.

اثبات: از مثال ۱، برآوردگر بیز θ به صورت زیر می‌باشد:

$$\hat{\theta}_B = \frac{n\tau^* + \sigma^* \omega}{n\tau^* + \sigma^*} \bar{X} + (1 - \omega) \frac{\mu\tau^*}{n\tau^* + \sigma^*}$$

همانطوریکه مشاهده می‌شود ضریب \bar{X} یعنی $\frac{n\tau^* + \sigma^* \omega}{n\tau^* + \sigma^*}$ اکیداً ω و ۱ است. از طرفی چون تابع زیان L ، اکیداً محدب است، لذا برآوردگر بیز فوق، یکتا بوده و بنابراین مجاز نیز خواهد بود. این ثابت می‌کند که $a\bar{X} + b$ برای θ مجاز است. به کمک روش حدی بیز می‌توان مجاز بودن $a\bar{X} + b$ را اثبات کرد که اثبات این در مرجع [۶] آمده است.

نتیجه ۱: با توجه به دو قضیه فوق ناحیه مجاز بودن $a\bar{X} + b$ به صورت زیر می‌باشد:

$$\{(a, b) : \omega \leq a < 1, \forall b\} \cup \{1, 0\}$$

در حالتیکه $\omega = 0$ باشد نوار عمودی زیر

$$\{(a, b) : 0 \leq a < 1, \forall b\} \cup \{1, 0\}$$

حاصل خواهد شد که ناحیه مجاز بودن متناظر با تابع زیان درجه دوم است (مرجع [۴] صفحات ۳۲۵-۳۲۲). بنابراین نتیجه حاصل با نتایجی که از قبیل برای تابع زیان درجه دوم وجود داشته است مطابقت دارد. یادآور می‌شویم که در حالتی خاص که $\omega = 0$ باشد، تابع زیان متعادل به تابع زیان درجه دوم کاهش پیدا می‌کند.

و بنابراین

$$\begin{aligned} R(\theta, a\bar{X} + b) &= [(a-1)\theta + b]^r + \frac{\sigma^r}{n} [(a-\omega)^r + \omega(n-\omega)] \\ &\geq \frac{\sigma^r}{n} [(a-\omega)^r + \omega(n-\omega)] \\ &> \frac{\sigma^r}{n} [(1-\omega)^r + \omega(n-\omega)] \\ &= R(\theta, \bar{X}) \end{aligned}$$

لذا در این حالت $a\bar{X} + b$ نمی‌تواند مجاز باشد زیرا \bar{X} بر آن برتری دارد.

ب) اگر $\omega < a$ باشد، آنگاه $(1-\omega)^r > (a-\omega)^r$ بنابراین داریم:

$$\begin{aligned} R(\theta, a\bar{X} + b) &= (a-1)^r \left[\theta + \frac{b}{a-1} \right] + \frac{\sigma^r}{n} [(a-\omega)^r + \omega(n-\omega)] \\ &> (\omega-1)^r \left[\theta + \frac{b}{a-1} \right] + \frac{\sigma^r}{n} [(a-\omega)^r + \omega(n-\omega)] \\ &\geq (\omega-1)^r \left[\theta + \frac{b}{a-1} \right] + \frac{\sigma^r}{n} \omega(n-\omega) \\ &= \left[(\omega-1)\theta + \frac{b(\omega-1)}{a-1} \right] + \frac{\sigma^r}{n} \omega(n-\omega) \\ &= R(\theta, \omega\bar{X} + \frac{b(\omega-1)}{a-1}) \end{aligned}$$

ج) اگر $a = 1, b \neq 0$ باشد، آنگاه

$$\begin{aligned} R(\theta, \bar{X} + b) &= \frac{\sigma^r}{n} [(1-\omega)^r + \omega(n-\omega)] + b^r \\ &> \frac{\sigma^r}{n} [(1-\omega)^r + \omega(n-\omega)] \\ &= R(\theta, \bar{X}) \end{aligned}$$

لذا در این حالت $b + \bar{X}$ نمی‌تواند مجاز باشد زیرا \bar{X} بر آن برتری دارد.

مراجع

- [1] Chung, Y. and Kim, C., 1997, Simultaneous Estimation of the Multivariate Normal Mean under **Balanced Loss Function**, *Commun. Stat.- Theory and Methods*, Vol.26, pp. 1599-1611.
- [2] Dey, D.K., Ghosh, M. and Strawderman, W., 1999, On Estimation with Balanced Loss Functions, *Statistics & Probability Letters*, V. 45, No. 2, pp. 97-101.
- [3] Gruber, M.H.G., 2004, The Efficiency of Shrinkage Estimators with respect to Zellner's Balanced Loss Function, *Commun. Stat.- Theory and Methods*, Vol. 33, No. 2, pp. 235-249.
- [4] Lehmann, E.L. and Casella, G., 1998, *Theory of Point Estimation*, Springer Verlag, New York.

- [5] Rodrigues, J. and Zellner, A., 1995, Weighted Balanced Loss Function for the Exponential Mean Time to Failure, *Commun. Stat. -Theory and Methods*, V. 23, pp. 3609-3616.
- [6] Sanjari Farsipour, N. and Asgharzadeh, A., 2004, Estimation of a Normal Mean Relative Balanced Loss Functions, *Statistical Paper*, Vol. 45, pp. 279-286.
- [7] Sanjari Farsipour, N. and Asgharzadeh, A., 2003, Bayesian Multivariate Normal Analysis under Balanced Loss Function, *Pak. J. Stat.* , V. 19, No. 2, pp. 231-240.
- [8] Zellner, A., 1994, Bayesian and Non-Bayesian Estimation Using Balanced Loss Functions, *Statistical Decision Theory and Related Topics V*, (J.O. Berger and S.S. Gupta Eds.), New York: Springer-Verlag, pp. 377-390.