توابع زیان متعادل[1]

اکبر اصغرزاده[2]

چکیده

این مقاله، مروری است بر توابع زیان متعادل که همزمان دو محک خوبی برازش[3] و دقت برآورد را منعکس می‌کنند . از دیدگاه بیزی بـرآورد پـارامتر میـانگین و بردار میانگین با این توابع زیان مورد مطالعه قرار می‌گیرد. مجاز بودن[4] کلاسی از برآوردگرهای خطی از میانگین نمونه، در مدل نرمال مورد بررسی قرار می‌گیرد.

واژه‌های کلیدی: توابع زیان متعادل، برآورد بیز، مجاز بودن یک برآوردگر.

۱. مقدمه

در تئوری تصمیم آماری، توابع زیان معمـولاً تأکیـد روی دقـت بـرآورد دارند، در حالیکه خوبی برازش نیز یـک محـک مهمـی اسـت. از اینـرو، زلنر[5] [۸] در سال ۱۹۹۴ توابع زیان متعادل را معرفی نمود که همزمان دو محک خوبی برازش و دقت برآورد را منعکس می‌کننـد. در گذشـته، توابع زیان منعکس‌کننده یکی از این دو محک اما نه هر دو، برای آنـالیز میانگین‌ها، رگرسیون و دیگر مسائل برآورد بکار رفتـه‌انـد. بـرای مثـال، برآورد کمترین مربعات[6]، خوبی برازش را منعکس می‌کنـد، در حالیکـه، استفاده از تابع درجه ۲، صرفاً دقت برآورد را نشان می‌دهد. همانطور که شناخته شده است، تأکید تنها روی محک دقت یک برآوردگر، به عنوان مثال میانگین مربع خطا، اغلب منجر به برآوردگرهای اریب می‌شـود. از طرف دیگر، استفاده از محک خـوبی بـرازش، هماننـد مجمـوع مربعـات باقیمانده‌ها[7] در یک مسأله رگرسیون، منجر به برآوردگری می‌شـود کـه خوبی برازش را می‌دهد و نااریب است. با این وجود، چنـین برآوردگـری ممکن است به دقت یک برآوردگر اریب با مقـدار اریبـی کـم نباشـد. از اینرو، معرفی توابع زیانی که همزمان دو محک خوبی برازش (یا فقـدان اریبی) و دقت برآورد را لحاظ کنند، ضروری به نظر می‌رسد. توابع زیان متعادل، این ضرورت را برآورده می‌سازند.

در مقاله [۸]، برآورد میانگین و بردار ضرایب رگرسیونی از دیدگاه بیزی مورد بحث و بررسی قرار گرفته است. رودریگز[8] و زلنر[۵] مـسأله برآورد میانگین زمان لازم برای شکست را با تابع زیان متعـادل مطالعـه کردند. چانگ و کیم[9] [۱] برآورد همزمان میانگین نرمال چنـدمتغیره را وقتی که تابع زیان متعادل است، مطالعه کـرده و برآوردگرهـای از نـوع جیمز و استاین[10] را برای این تابع زیان بدسـت آوردنـد. بـرای مطالعـه بیشتر روی مـسأله بـرآورد بـا تـابع زیـان متعـادل مـی‌تـوان بـه دی و همکاران[11] [۲]، سـنجری و اصـغرزاده[۶ و ۷] و گروبـر[12][۳] مراجعـه کرد.

در بخش ۲ این مقاله، تابع زیان متعادل تعریف و انواع مختلف آن بیان می‌شود. در بخش ۳، برآورد میانگین جامعـه را از دیـدگاه بیـزی مـورد بررسی قرار داده و مجاز بودن کلاسی از برآوردگرهـای خطـی میـانگین 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 بطوریکه

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

که θ میانگین مشترک X_i ها، $\mathbf{1}$ بردار $n\times 1$ تایی با عناصر ۱ و $\mathbf{u}$ یک بردار $n\times 1$ تایی خطا باشد. افزون بر این، فرض کنید که برای پارامتر θ یک چگالی پسین $\pi(\theta\mid\Delta)$ که $\Delta=(\mathbf{X},\pi_0)$ و π_0 یک چگالی پیشین برای θ است، وجود دارد. فرض های مختلفی را می‌توان درباره توزیع $\mathbf{X}$ و π_0 در نظر گرفت.

۱- Balanced loss functions

۲- گروه آمار، دانشگاه مازندران

۳- Goodness if fit

۴- Admissibility

۵- Zellner

۶- least squares estimation

۷- Sum of squared residuals

۸- Rodrigues

۹- Chung and Kim

۱۰- James and Stein

۱۱- Dey et.al.

۱۲- Gruber

برای پارامتر θ یک چگالی پسین $\pi(\theta \mid \Delta)$ که $\Delta = (\mathbf{X}, \pi_0)$ و π_0 یک چگالی پیشین برای θ است، وجود دارد. فرض های مختلفی را می‌توان درباره توزیع $\mathbf{X}$ و π_0 در نظر گرفت.
تابع زیان متعادل برای برآورد پارامتر θ به صورت زیر معرفی می‌شود.

$$\begin{aligned} L_1(\hat{\theta},\theta) &= \frac{\omega}{n}(\mathbf{X}-\hat{\theta}\mathbf{1})'(\mathbf{X}-\hat{\theta}\mathbf{1}) + (1-\omega)(\hat{\theta}-\theta)^2 \\ &= \frac{\omega}{n}\sum_{i=1}^{n}(X_i-\hat{\theta})^2 + (1-\omega)(\hat{\theta}-\theta)^2 \qquad (۲) \end{aligned}$$

که در آن $0 \le \omega \le 1$ و $\hat{\theta}$ یک برآوردگر θ است. اولین جمله در رابطه (۲)، خوبی برازش را نشان می‌دهد، در حالیکه دومین جمله دقت برآوردگر را نشان می‌دهد. در رابطه (۲)، بجای تابع زیان درجه دوم $(\hat{\theta}-\theta)^2$، می‌توان از توابع 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.

$$L_2(\hat{\theta},\theta) = \omega q(\theta)\frac{\sum_{i=1}^{n}(X_i-\hat{\theta})^2}{n} + (1-\omega)q(\theta)(\hat{\theta}-\theta)^2$$

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

$$\begin{aligned} L_3(\hat{\theta},\theta) &= \frac{\omega}{N}\sum_{i=1}^{N}(\mathbf{X}_i-\hat{\boldsymbol{\theta}})'Q(\mathbf{X}_i-\hat{\boldsymbol{\theta}}) \\ &\quad + (1-\omega)(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})'\mathbf{Q}(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}) \end{aligned}$$

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

۳. برآورد میانگین

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

۳. ۱. برآورد بیزی میانگین

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

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

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

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

برآوردگر بیز θ به صورت زیر به دست می‌آید.

$$\hat{\theta}_B = \omega\overline{X} + (1-\omega)E(\theta \mid \mathbf{X}) \qquad (۳)$$

که در آن $E(\theta \mid \mathbf{X})$ میانگین پسین می‌باشد. همانطوریکه مشاهده می‌شود، $\hat{\theta}_B$ ترکیبی از برآوردگر غیر بیز $\overline{X}$ و برآوردگر بیز $E(\theta \mid \mathbf{X})$ می‌باشد.
در حالتی که $\boldsymbol{\theta}$ یک بردار میانگین و عناصر نمونه برداری باشند، به کمک مشتق‌گیری برداری برآوردگر بیز، مشابه با حالت قبل به صورت زیر به دست می‌آید[۷].

$$\hat{\boldsymbol{\theta}} = \omega\overline{\mathbf{X}} + (1-\omega)E(\theta \mid \mathbf{X})$$

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

۳. ۲. چند مثال

مثال۱: فرض کنید $X_1, X_2, \ldots, X_n$ یک نمونه تصادفی از توزیع نرمال $N(\theta,\sigma^2)$ باشد بطوریکه توزیع پیشین θ نیز نرمال با میانگین μ و واریانس τ^2 است. در آن صورت به سادگی می‌توان نشان داد که توزیع پسین θ ، نرمال با میانگین و واریانس

۱- Weight function

$$E(\theta \mid \mathbf{X}) = \frac{\left[\frac{n\overline{X}}{\sigma^{2}} + \frac{\mu}{\tau^{2}}\right]}{\left[\frac{n}{\sigma^{2}} + \frac{1}{\tau^{2}}\right]}$$

(۴)

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

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

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

که به فرم خطی $a\overline{X} + b$ می‌باشد.

مثال۲: فرض کنید $X_1, X_2, \ldots, X_n$ یک نمونه تصادفی از توزیع برنولی با تابع چگالی احتمال

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

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:

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

مثال۳: فرض کنید که $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n$ یک نمونه تصادفی از یک توزیع نرمال چندمتغیره با بردار میانگین $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)'$ و ماتریس کواریانس $\sigma^{2}\mathbf{I}$ باشد. اگر توزیع پیشین $\boldsymbol{\theta}$ را نیز یک توزیع نرمال چندمتغیره با بردار میانگین $\mathbf{0}$ و ماتریس کواریانس mI فرض کنیم در آن صورت توزیع پسین، نرمال چندمتغیره با بردار میانگین $\overline{\mathbf{X}}\frac{N}{N + 1/m}$ و ماتریس کواریانس $\mathbf{I}\frac{1}{N + 1/m}$ خواهد بود. بنابراین برآوردگر بیز می‌شود:

$$E(\theta \mid \mathbf{X}) = \omega\overline{\mathbf{X}} + (1-\omega)\hat{\boldsymbol{\theta}}_B = \overline{\mathbf{X}}\,\frac{\omega + Nm}{1 + Nm}$$

۳. مجاز بودن برآوردگر $a\overline{X} + b$

با توجه به مثالهای فوق، برآوردگرهای بیز به فرم خطی $a\overline{X} + b$ می‌باشد. در این بخش مجاز بودن این کلاس از برآوردهای خطی تحت مدل (۱) مورد مطالعه قرار می‌گیرد. تابع ریسک برآوردگر $a\overline{X} + b$ تحت تابع زیان متعادل L_1 به صورت زیر می‌باشد،

$$\begin{aligned}
R(\theta, a\overline{X} + b) &= \frac{\omega}{n}\sum_{i=1}^{n} E(X_i - a\overline{X} - b)^{2} \\
&\quad + (1-\omega)E(a\overline{X} + b - \theta)^{2}
\end{aligned}$$

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

$$\begin{aligned}
R(\theta, a\overline{X} + b) &= [(a-1)\theta + b]^{2} \\
&\quad + \frac{\sigma^{2}}{n}\left[(a-\omega)^{2} + \omega(n-\omega)\right]
\end{aligned}$$

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

قضیه ۱. فرض کنید $X_1, X_2, \ldots, X_n$ یک نمونه تصادفی n تایی از جامعه‌ای (نه لزوماً نرمال) با میانگین θ و واریانس σ^{2} باشد در آن صورت $a\overline{X} + b$ برای θ تحت تابع زیان L_1 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 برآوردگری معرفی شود که از لحاظ ریسک بر $a\overline{X} + b$ برتری داشته باشد.

الف) اگر $a > 1$، آنگاه

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

و بنابراین

$$
\begin{aligned}
R(\theta,a\bar{X}+b) &= [(a-1)\theta+b]^2+\frac{\sigma^2}{n}\left[(a-\omega)^2+\omega(n-\omega)\right]\\
&\ge \frac{\sigma^2}{n}\left[(a-\omega)^2+\omega(n-\omega)\right]\\
&> \frac{\sigma^2}{n}\left[(1-\omega)^2+\omega(n-\omega)\right]\\
&= R(\theta,\bar{X})
\end{aligned}
$$

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

ب) اگر $a<\omega$ باشد، آنگاه $(a-\omega)^2>(1-\omega)^2$ بنابراین داریم:

$$
\begin{aligned}
R(\theta,a\bar{X}+b) &= (a-1)^2\left[\theta+\frac{b}{a-1}\right]^2+\frac{\sigma^2}{n}\left[(a-\omega)^2+\omega(n-\omega)\right]\\
&> (\omega-1)^2\left[\theta+\frac{b}{a-1}\right]^2+\frac{\sigma^2}{n}\left[(a-\omega)^2+\omega(n-\omega)\right]\\
&\ge (\omega-1)^2\left[\theta+\frac{b}{a-1}\right]^2+\frac{\sigma^2}{n}\omega(n-\omega)\\
&= \left[(\omega-1)\theta+\frac{b(\omega-1)}{a-1}\right]^2+\frac{\sigma^2}{n}\omega(n-\omega)\\
&= R(\theta,\omega\bar{X}+\frac{b(\omega-1)}{a-1})
\end{aligned}
$$

ج) اگر $a=1, b\neq 0$ باشد، آنگاه

$$
\begin{aligned}
R(\theta,\bar{X}+b) &= \frac{\sigma^2}{n}\left[(1-\omega)^2+\omega(n-\omega)\right]+b^2\\
&> \frac{\sigma^2}{n}\left[(1-\omega)^2+\omega(n-\omega)\right]\\
&= R(\theta,\bar{X})
\end{aligned}
$$

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

قضیه۲: اگر $X_1,X_2,\dots,X_n$ یک نمونه تصادفی از $N(\theta,\sigma^2)$ باشد در آن صورت $a\bar{X}+b$ برای θ مجاز است هرگاه داشته با شیم $\omega\le a<1, b\in\Re$ یا $a=1, b=0$.

اثبات: از مثال۱، برآوردگر بیز θ به صورت زیر می‌باشد:

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

همانطوریکه مشاهده می‌شود ضریب $\bar{X}$ یعنی $\frac{n\tau^2+\sigma^2\omega}{n\tau^2+\sigma^2}$ اکیداً بین ω و ۱ است. از طرفی چون تابع زیان L_1، اکیداً محدب است، لذا برآوردگر بیز فوق، یکتا بوده و بنابراین مجاز نیز خواهد بود. این ثابت می‌کند که $a\bar{X}+b$، برای $\omega<a<1, b\in\Re$ مجاز است. به کمک روش حدی بیز می‌توان مجاز بودن $\omega\bar{X}+b, \bar{X}$ را اثبات کرد که اثبات این در مرجع [۶] آمده است.

نتیجه۱. با توجه به دو قضیه فوق ناحیه مجاز بودن $a\bar{X}+b$ به صورت زیر می‌باشد:

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

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

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

حاصل خواهد شد که ناحیه مجاز بودن متناظر با تابع زیان درجه دوم است(مرجع[۴] صفحات ۳۲۵-۳۲۳) . بنابراین نتیجه حاصل با نتایجی که از قبل 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.

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.