نابرابری کرامر ـ رائو در آمار بیز یا نابرابری ون ـ تریز

مینا توحیدی* و جواد بهبودیان†

چکیده

در این مقاله نخست نابرابری کرامر ـ رائو را در آمار بیز مورد تجزیه و تحلیل قرار داده و یک کران پایین برای مخاطره بیز برآوردکننده‌های پارامتر مجهول در یک توزیع آماری ارائه می‌دهیم. شرایط برقراری نابرابری اخیر را با نابرابری کرامر ـ رائو مقایسه می‌کنیم و در پایان مثالهایی می‌آوریم و نقش برآوردکننده بیز را در نابرابری نامبرده یادآور می‌شویم.

۱ پیشگفتار

اکثر کتابهای آماری، نابرابری کرامر ـ رائو را معرفی کرده نشان داده‌اند که این نابرابری در نظریهٔ آمار از اهمیت ویژه‌ای برخوردار است. شما نیز با این نابرابری آشنایی دارید اما برای یادآوری یک بار دیگر به آن اشاره می‌کنیم.

فرض کنید $X_1,\ldots,X_n$ نمونه‌ای تصادفی از متغیر تصادفی X، با چگالی $f(x|\theta)$ و $\theta \in A$ باشد و فضای پارامتری A بازه‌ای روی اعداد حقیقی تشکیل دهد. اگر شرایط زیر را داشته باشیم:

الف ـ برای تمام مقادیر $\theta \in A$، $\frac{\partial f(x|\theta)}{\partial\theta}$ موجود باشد؛

ب ـ $E_\theta[\frac{\partial}{\partial\theta}\ln f(X|\theta)] = 0$؛

ج ـ $I_X(\theta) = E_\theta[\frac{\partial}{\partial\theta}\ln f(X|\theta)]^2$ مقداری متناهی باشد؛

د ـ برای $\hat\theta(X_1,\ldots,X_n)$ به عنوان یک برآوردکننده پارامتر θ، تساوی زیر برقرار باشد:

$$\frac{\partial}{\partial\theta}\int\cdots\int\hat\theta(x_1,\ldots,x_n)\prod_{i=1}^{n}f(x_i|\theta)dx_1\ldots dx_n =$$

$$\int\cdots\int\hat\theta(x_1,\ldots,x_n)[\frac{\partial}{\partial\theta}\prod_{i=1}^{n}f(x_i|\theta)]dx_1\ldots dx_n$$

آنگاه نابرابری زیر (نابرابری کرامر ـ رائو) برقرار است:

$$\mathrm{Var}_\theta(\hat\theta) \geq \frac{[1+b'(\theta)]^2}{nI_X(\theta)}$$

$b(\theta) = E_\theta(\hat\theta) - \theta$ اریبی برآوردکنندهٔ $\hat\theta$ برای پارامتر θ است.

در آمار کلاسیک، واریانس برآوردکننده‌ها مورد مقایسه قرار می‌گیرند ولی در آمار بیز، برآوردکننده‌ها به وسیله مخاطره ـ بیز آنها با هم مقایسه می‌شوند. ما در این مقاله یک کران پایین برای مخاطره بیز برآوردکننده‌ها پیدا می‌کنیم. همچنین باید توجه داشت که بررسی شرایط لازم برای برقرار نابرابری کرامر ـ رائو، در اکثر موارد دشوار است و اگر هم این شرایط برقرار شوند، تنها برآوردکننده‌هایی مورد مقایسه قرار می‌گیرند که در شرط (د) صدق کنند. اما نابرابریهایی که در این مقاله ارائه خواهد شد، دارای چنین مشکلاتی نیست.

قضیهٔ اصلی این مقاله بر اساس مرجع [۱] است که با روشی ساده ثابت شده و در مثال به عنوان کاربرد آن ارائه گردیده است.

۲ نابرابری کرامر ـ رائو در آمار بیز و اثبات آن

در این بخش، با تعریف یک تابع توزیع احتمال روی فضای پارامتری A، به نام توزیع پیشین متغیر تصادفی Θ یک کران پایین برای مخاطره بیز برآوردکنندهٔ $\hat\theta$ ارائه می‌دهیم.

*) خانم مینا توحیدی، دانشجوی کارشناسی‌ارشد آمار، دانشگاه شیراز
†) دکتر جواد بهبودیان، استاد آمار، دانشگاه شیراز

قضیه: فرض کنید $X_1, \ldots, X_n$ یک نمونه تصادفی از X، با چگالی $f(x|\theta)$ و $\theta \in A$ بوده و $\pi(\theta)$ یک تابع چگالی احتمال روی فضای پارامتری $A = (\underline{\theta}, \bar{\theta})$ باشد، به طوری که π در نقاط ابتدا و انتهای بازهٔ A همگرا به صفر گردد. با تعریف دو تابع زیر:

$$I_X(\theta) = E_\theta\{\frac{\partial}{\partial\theta} \ln f(X|\theta)\}^2,$$
$$J(\pi) = E\{\frac{\partial}{\partial\theta} \ln \pi(\theta)\}^2$$

و فرض $E_\theta[\frac{\partial}{\partial\theta} \ln f(X|\theta)] = 0$ داریم:

$$\text{مخاطره بیز}\ \hat{\theta} = r(\hat{\theta})$$
$$= E[\hat{\theta}(X_1, \ldots, X_n) - \Theta]^2 \geq \frac{1}{nE[T_X(\Theta)] + J(\pi)}$$

امید ریاضی در این نابرابری نسبت به توزیع توأم (X, Θ) است. $I_X(\theta)$ اطلاع فیشر برای θ در X و $J(\pi)$ اطلاع فیشر برای θ در چگالی پیشین π نامیده می‌شود.

اثبات: قضیه را برای $n = 1$ اثبات می‌کنیم. روش اثبات برای $n > 1$ مانند حالت $n = 1$ است. نخست تساوی زیر را در نظر می‌گیریم:

$$\iint [\hat{\theta}(x) - \theta][\frac{\partial}{\partial\theta} f(x|\theta)\pi(\theta)] d\theta dx =$$
$$\int \hat{\theta}(x) \left\{ \int [\frac{\partial}{\partial\theta} f(x|\theta)\pi(\theta)] d\theta \right\} dx$$
$$- \iint \theta [\frac{\partial}{\partial\theta} f(x|\theta)\pi(\theta)] d\theta dx \quad (1)$$

حال توجه کنید که:

$$\int [\frac{\partial}{\partial\theta} f(x|\theta)\pi(\theta)] d\theta = f(x|\theta)\pi(\theta)|_{\underline{\theta}}^{\bar{\theta}} = 0 \qquad (2)$$

از طرفی، با استفاده از انتگرال جزء به جزء خواهیم داشت:

$$\int \theta [\frac{\partial}{\partial\theta} f(x|\theta)\pi(\theta)] d\theta = \theta f(x|\theta)\pi(\theta)|_{\underline{\theta}}^{\bar{\theta}} - \int f(x|\theta)\pi(\theta) d\theta$$
$$= - \int f(x|\theta)\pi(\theta) d\theta \qquad (3)$$

با به‌کارگیری (۲) و (۳) در رابطه (۱)، داریم:

$$\iint [\hat{\theta}(x) - \theta][\frac{\partial}{\partial\theta} f(x|\theta)\pi(\theta)] d\theta dx =$$
$$\iint f(x|\theta)\pi(\theta) d\theta dx = 1$$

حال با استفاده از نابرابری کوشی - شوارتس برحسب انتگرال دوگانه با دو تابع زیر: $G(x,\theta) = [\frac{\partial}{\partial\theta} \ln f(x|\theta)\pi(\theta)]\sqrt{f(x|\theta)\pi(\theta)}$ و $F(x,\theta) = [\hat{\theta}(x) - \theta]\sqrt{f(x|\theta)\pi(\theta)}$ رابطه زیر به‌دست می‌آید:

$$1 = \iint F(x,\theta) G(x,\theta) d\theta dx$$
$$= [\iint F(x,\theta) G(x,\theta) d\theta dx]^2$$
$$\leq \iint [\hat{\theta}(x) - \theta]^2 f(x|\theta)\pi(\theta) d\theta dx$$
$$\times \iint [\frac{\partial}{\partial\theta} \ln f(x|\theta)\pi(\theta)]^2 f(x|\theta)\pi(\theta) d\theta dx$$

با توجه به اینکه $E_\theta[\frac{\partial}{\partial\theta} \ln f(x|\theta)] = 0$، داریم

$$\iint [\frac{\partial}{\partial\theta} \ln f(x|\theta)] [\frac{\partial}{\partial\theta} \ln \pi(\theta)] f(x|\theta)\pi(\theta) d\theta dx$$
$$= \int [\int [\frac{\partial}{\partial\theta} \ln f(x|\theta)] f(x|\theta) dx] [\frac{\partial}{\partial\theta} \ln \pi(\theta)] \pi(\theta) d\theta = 0$$

اینک می‌نویسیم:

$$\iint [\frac{\partial}{\partial\theta} \ln f(x|\theta)\pi(\theta)]^2 f(x|\theta)\pi(\theta) d\theta dx$$
$$= \iint [\frac{\partial}{\partial\theta} \ln f(x|\theta) + \frac{\partial}{\partial\theta} \ln \pi(\theta)]^2 f(x|\theta)\pi(\theta) d\theta dx$$
$$= \iint [\frac{\partial}{\partial\theta} \ln f(x|\theta)]^2 f(x|\theta)\pi(\theta) d\theta dx$$
$$+ \iint [\frac{\partial}{\partial\theta} \ln \pi(\theta)]^2 f(x|\theta)\pi(\theta) d\theta dx$$
$$= \int \left\{ \int [\frac{\partial}{\partial\theta} \ln f(x|\theta)]^2 f(x|\theta) dx \right\} \pi(\theta) d\theta +$$
$$\int [\int f(x|\theta) dx] [\frac{\partial}{\partial\theta} \ln \pi(\theta)]^2 \pi(\theta) d\theta$$
$$= E[I_X(\Theta)] + J(\pi)$$

در نتیجه

$$r(\hat{\theta}) = E[\hat{\theta}(X) - \Theta]^2 \geq \frac{1}{E[I_X(\Theta)] + J(\pi)}$$

لازم است بیفزاییم که:

۱) نابرابری بالا را، نابرابری ون - تریز (Van-Trees) هم می‌نامند. توجه کنید که در این نابرابری $\hat{\theta}$ هر برآوردکننده‌ای برای θ است و نیازی به برقراری شرط (د) در مورد $\hat{\theta}$ نیست. البته طبق شرایط قضیه $\hat{\theta}$ نمی‌تواند برآوردکنندهٔ بیز نااریب باشد [۲ صفحه ۲۴۴].

۲) اگر $\hat{\psi}(X)$ برآوردکننده‌ای برای تابع $\psi(\theta)$ باشد، نابرابری زیر به‌دست می‌آید:

$$E[\hat{\psi}(X) - \psi(\Theta)]^2 \geq \frac{E[\psi'(\Theta)]^2}{E[I_X(\Theta)] + J(\pi)}$$

۳) حالت تساوی در نابرابری ون - تریز هنگامی برقرار می‌شود که در نابرابری کشی - شوارتس بالا تساوی برقرار شود، یعنی داشته باشیم

$$C(\hat{\theta}(x) - \theta) = \frac{\partial}{\partial\theta} \ln f(x|\theta)\pi(\theta)$$

که در آن C عددی ثابت است. این برابری بدین معناست که چگالی پسین به صورت زیر است:

$$\pi(\theta|x) \propto f(x|\theta)\pi(\theta) = e^{-\frac{۱}{۲}C(\hat{\theta}(x)-\theta)^{۲}}$$

به سخنی دیگر چگالی پسین، به شکل چگالی توزیع نرمال با میانگین $\hat{\theta}(x)$ است. پس به‌طور کلی می‌توان گفت تنها هنگامی نابرابری ون-تریز تبدیل به تساوی می‌شود که چگالی پسین نرمال باشد و داشته باشیم $E[\Theta|x] = \hat{\theta}(x)$، یعنی $\hat{\theta}(x)$ برآوردکنندهٔ بیز با تابع زیان درجه دو باشد.

۳ دو مثال برای توزیع نرمال

مثال ۱. فرض کنید $\underline{X} = (X_{۱},\dots,X_n)$ نمونه‌ای تصادفی از $N(\theta,\sigma^{۲})$ با $\sigma^{۲}$ معلوم باشد. برای برآورد پارامتر θ وقتی که با تابع زیان درجه دو کار می‌کنیم، در آمار بیز معمولاً فرض می‌شود که پارامتر θ خود یک متغیر تصادفی با توزیع $N(\mu,\tau^{۲})$ است. واضح است که در این صورت، داریم:

$$I_x(\theta) = \frac{۱}{\sigma^{۲}}, \qquad J(\tau) = \frac{۱}{\tau^{۲}}$$

در نتیجه کران پایین نابرابری ون - تریز برابر است با

$$\frac{۱}{nE[I_x(\Theta)] + J(\pi)} = \frac{۱}{\frac{n}{\sigma^{۲}} + \frac{۱}{\tau^{۲}}}$$

همچنین می‌دانیم که [۲]، یک برآوردکنندهٔ بیز برای پارامتر θ عبارت است از:

$$\delta_B(\underline{x}) = \left(\frac{n/\sigma^{۲}}{n/\sigma^{۲} + ۱/\tau^{۲}}\right)\bar{x} + \left(\frac{۱/\tau^{۲}}{n/\sigma^{۲} + ۱/\tau^{۲}}\right)$$

با مخاطره بیز

$$r(\delta_B) = \frac{۱}{\frac{n}{\sigma^{۲}} + \frac{۱}{\tau^{۲}}}$$

ملاحظه می‌شود که در این مثال مخاطره بیز برآوردکنندهٔ بیز با کران پایین نابرابری ون-تریز برابر می‌شود، زیرا چگالی پسین، نرمال است و $\delta_B(\underline{x}) = E(\Theta|\underline{x})$.

مثال ۲. فرض کنید $\underline{X} = (X_{۱},\dots,X_n)$ یک نمونه تصادفی از توزیع $N(۰,\sigma^{۲})$ باشد. برای برآورد $\sigma^{۲}$، وقتی با تابع زیان درجه دو کار می‌کنیم[۲] با فرض اینکه $\tau = \frac{۱}{۲\sigma^{۲}}$ دارای توزیع گاما با پارامترهای g و $\frac{۱}{\alpha}$ باشد، داریم:

$$E(\tau) = \frac{g}{\alpha}, \qquad E(\tau^{۲}) = \frac{g(g+۱)}{\alpha^{۲}}$$

$$E(\frac{۱}{\tau}) = \frac{\alpha}{g-۱}, \qquad E(\frac{۱}{\tau^{۲}}) = \frac{\alpha^{۲}}{(g-۱)(g-۲)}, \qquad g > ۲$$

در نتیجه

$$I_x(\sigma^{۲}) = \frac{۱}{۲\sigma^{۴}}, \qquad E[I_x(\sigma^{۲})] = ۲E(\tau^{۲}) = \frac{۲g(g+۱)}{\alpha^{۲}}$$

$$J(x) = E[-\frac{g-۱}{\sigma^{۲}} + \frac{\alpha}{۲\sigma^{۴}}]^{۲} = ۴E[-\tau(g-۱) + \alpha\tau^{۲}]^{۲}$$

پس از محاسبات لازم، کران پایین نابرابری ون - تریز برابر خواهد بود با

$$\frac{\alpha^{۲}}{۲g(g+۱)(n+۲g+۲۲)}$$

از طرفی برآوردکنندهٔ بیز برای $\sigma^{۲} = \frac{۱}{۲\tau}$ برابر است با

$$\delta_B(\underline{x}) = \frac{\alpha + Y}{n + ۲g - ۲}$$

با فرض $Y = \sum_{i=۱}^{n} X_i^{۲}$. ضمناً تابع مخاطره δ_B برابر است با

$$R(\sigma^{۲},\delta_B) = \frac{۲n\sigma^{۴}}{(n+۲g-۲)^{۲}} + \frac{(\alpha-(۲g-۲)\sigma^{۲})^{۲}}{(n+۲g-۲)^{۲}}$$

در نتیجه

$$\begin{aligned}
r(\delta_B) &= \frac{n}{۲(n+۲g-۲)^{۲}}E(\frac{۱}{\tau^{۲}}) + \frac{(g-۱)^{۲}}{(n+۲g-۲)^{۲}}\operatorname{var}(\frac{۱}{\tau}) \\
&= \frac{n}{۲(n+۲g-۲)^{۲}}\left(\frac{\alpha^{۲}}{(g-۱)(g-۲)}\right) \\
&\quad + \frac{(g-۱)^{۲}}{(n+۲g-۲)^{۲}}\left(\frac{\alpha^{۲}}{(g-۱)^{۲}(g-۲)}\right) \\
&= \frac{\alpha^{۲}}{۲(g-۱)(g-۲)(n+۲g-۲)}
\end{aligned}$$

ملاحظه می‌شود که در این مثال نابرابری ون - تریز یک نابرابری اکید است، زیرا چگالی پسین نرمال نیست.

در پایان دو پرسش را مطرح می‌کنیم:

۱) آیا می‌توان مثالی مناسب پیدا کرد که شرط (د) را نداشته باشد، یعنی نابرابری کرامر - رائو برایش صادق نباشد، امّا نابرابری ون - تریز برای آن برقرار گردد؟ ظاهراً پیداکردن چنین مثالی آسان نیست.

۲) آیا در صورتی که $\pi(\theta)$ چگالی ناسره (مانند $\pi(\theta) = ۱$) باشد، نابرابری ون - تریز چه موقع می‌تواند برقرار گردد؟ ملاحظه می‌شود که در مثال ۱ با فرض $\tau^{۲} = \infty$ (پیشین یکنواخت ناسره) نابرابری ون - تریز برقرار است، ولی در مثال ۲ با فرض $\alpha = ۱$ و $g = ۰$ (پیشین ناسره e^{-x}/x) نابرابری برقرار نمی‌شود.

قدردانی: از پیشنهادهای داور محترم که برای بهبود مقاله سودمند بوده‌اند، سپاسگزاری می‌شود.

مراجع

[1] Gill, R. D. & Levit, B. Y. (1995) Applications of the Van-Trees inequality. Bernoulli 1, 59-79.

[2] Lehmann, E. L. (1983) Theory of Point Estimation. John Wiley & sons, New York.

[3] Van Trees, H. L. (1968) Detection, Estimation and Modulation theory, Part 1. New York, Wiley.