Statistical Sciences Extension
Vol. 2, No. 1, Winter 2017, (P 55-66)

گسترهٔ علوم آماری
سال دوم، شماره اول، پاییز و زمستان 1395، (ص 55-66)

کاربردهای نامساوی کرامر - رائو در محاسبهٔ برآوردگرهای کمینه - بیشینه تحت داده‌های سانسور شده و مدل کوزیول - گرین

مهدی شمس[*1]، **غلامرضا حسامیان**[2]

1. استادیار، گروه آمار، دانشگاه کاشان
2. استادیارف گروه آمار، دانشگاه شهرکرد

تاریخ دریافت: 1395/09/21 **تاریخ پذیرش**: 1395/12/24

Applications of Cramer - Rao Inequality in Calculations Minimax Estimators Under Censored Data and Koziol - Green Model

M. Shams[*1], **GH. Hesamian**[2]

1. Assistant Professor, Department of Statistics, Kashan University
2. Assistant Professor, Department of Statistics, Shahrekord University

Received: 2016/12/11 **Accepted:** 2017/03/14

Abstract

In this paper, by using a Cramer - Rao inequality we obtained a lower bound minimax risk based on a random sample of random stopping time, in a condition of the truncated parameter space.

Finally, as applications, we obtained minimax estimators under censored data, Koziol - Green model and exponential distribution family.

Keywords

Minimax Estimator, Censored Data, Koziol - Green Model, Stopping Time.

چکیده

در این مقاله با استفاده از نامساوی کرامر - رائو یک کران پایین برای مخاطرهٔ کمینه بیشینه بر پایهٔ یک نمونه تصادفی با زمان توقف تصادفی، در حالتی که فضای پارامتر بریده باشد، به دست می‌آوریم.

در پایان به‌عنوان کاربردهایی از این مسئله، برآوردگرهای کمینه - بیشینه را تحت داده‌های سانسور شده، مدل کوزیول - گرین و خانواده توزیع‌های نمایی به دست می‌آوریم.

واژگان کلیدی

برآوردگر کمینه - بیشینه، داده‌های سانسور شده، مدل کوزیول - گرین، زمان توقف.

Corresponding Author: mehdi_shams1357@yahoo.com

* **نویسنده مسئول**: مهدی شمس

مقدمه

نامساوی‌های اطلاع، کاربرد فراوان در نظریهٔ برآوردیابی و اتخاذ تصمیم‌های آماری دارند. نامساوی اطلاع برای مخاطره بیز در پیشگوکننده‌ها [11]، خانواده‌های نامنظم [1]، پیشین‌های بریده شده [2]، تحت زیان مقیاسی [4] به کار گرفته شده و از طریق آن رویکردی برای یافتن برآوردگرهای کمینه - بیشینه پیدا می‌شود. روش‌های کارایی مخاطره مجانبی دنباله‌ای [10، 9، 8]، برآوردهای بیز دنباله‌ای با داده‌های سانسور شده [15، 14، 13] و برآوردگرهای کمینه - بیشینه برای پارامتر مقیاس در داده‌های سانسور شده و بریده شده [7 و 6] از مهم‌ترین تحقیقاتی است که در این زمینه پژوهشگران انجام داده‌اند. نامساوی کرامر - رائو یک کران پایین برای مخاطره کمینه بیشینه بر پایهٔ یک نمونه تصادفی با زمان توقف تصادفی، در حالتی که فضای پارامتر بریده باشد، پیدا می‌کند. در این مقاله نامساوی‌های اطلاع برای مخاطره کمینه - بیشینه برآوردگرهای دنباله‌ای تحت زیان مربع خطا برای برآوردگرهای خطی محاسبه می‌شود و در پایان به‌عنوان کاربردهایی از این مسئله، برآوردگرهای کمینه -بیشینه را تحت داده‌های سانسور شده از راست و مدل کوزیول - گرین[1] و خانواده توزیع‌های نمایی به دست می‌آوریم.

ابتدا مفاهیمی از نظریهٔ تصمیم مانند تصمیم بیز و تصمیم کمینه - بیشینه مرور می‌شوند.

فرض می‌شود $\mathbf{X}=(X_1,\dots,X_n)$ متغیرهای تصادفی مستقل و هم توزیع در فضای اندازه‌پذیر[2] (X,S) با اندازهٔ احتمال $P_{\boldsymbol\theta}$ باشند که در آن $\boldsymbol\theta\in\Theta\subset R^m$ و $m>1$. همچنین فرض می‌شود خانواده توزیع $\{P_{\boldsymbol\theta},\boldsymbol\theta\in\Theta\}$ نسبت به اندازهٔ σ - متناهی μ مطلقاً پیوسته باشند و $f(.,\boldsymbol\theta)=dP_{\boldsymbol\theta}/d\mu$. به‌علاوه فرض می‌شود $q(\boldsymbol\theta)$ یک چگالی احتمال پیشین[3] برای $\boldsymbol\theta$ با تکیه‌گاه $S_q\subset\Theta$ باشد. قاعده تصمیم (غیر تصادفی)[4] $\boldsymbol\theta^*$ تابعی اندازه‌پذیر از فضای نمونه‌ای X به فضای کارها[5] A است به‌طوری‌که به یک مشاهده نمونه تصادفی عملی را نسبت می‌دهد. فضای کارها معین می‌کند که مسئلهٔ تصمیم مورد مطالعه، یک مسئلهٔ برآورد نقطه‌ای و یا فاصله اطمینان و یا آزمون فرضیه است. برای مثال، در مسئلهٔ برآورد نقطه‌ای کارها حدس نقاط و در مسئلهٔ آزمون فرضیه دو کار پذیرش فرضیه صفر یا رد آن هستند. همچنین در مسئلهٔ برآورد نقطه‌ای تابع تصمیم همان برآوردگر است و در مسئله آزمون فرضیه تابع تصمیم همان تابع آزمون است. مجموعهٔ همهٔ توابع تصمیم ممکن فضای تصمیم[6] نامیده می‌شود. در نظریهٔ تصمیم هدف انتخاب بهینه یک تابع تصمیم است. تابع نامنفی $L(a,\theta)$ از فضای حاصل‌ضرب دکارتی فضای پارامتر در فضای کارها به اعداد حقیقی مثبت را تابع زیان[7] می‌نامند که نشان‌دهنده زیان ناشی از اخذ تصمیم a است، هنگامی که در حالت واقعی $\boldsymbol\theta$ باشد [13]. انتخاب تابع زیان بستگی به مسئلهٔ مورد نظر دارد و یکی از متداول‌ترین آنها تابع زیان توان دوم خطا (SEL) است. تابع زیان ملاکی برای ارزیابی یک تصمیم است. رویکرد کلاسیک برای اخذ چنین تصمیمی، مینیمم‌سازی متوسط زیان است که این متوسط زیان را تابع مخاطره[8] تصمیم d^* می‌نامند و به‌صورت

$$R\left(d^*,q\right)=E_{X^n|q}\left(L\left(d^*\left(\mathbf{X}\right),q\right)\right)$$

تعریف می‌شود [13]. در روش‌های کلاسیک پارامتر θ یک مقدار ثابت و نامعلوم در نظر گرفته می‌شود و یک نمونه تصادفی X از جمعیتی که دارای توزیع $f_n(\mathbf{x},\boldsymbol\theta)$ است، جمع‌آوری می‌شود تا بر اساس آن در مورد t تصمیم‌گیری شود. در روش بیزی Θ کمیتی در نظر گرفته می‌شود که خود یک متغیر تصادفی، با مقدار مشاهده شده θ است و تغییرات آن به وسیلهٔ یک توزیع پیشین[9] $q(\boldsymbol\theta)$ مشخص می‌شود. این توزیع پیشین بر اساس اعتقادات و تجربیات قبلی آزمایشگر یا اطلاعاتی در مورد توزیع پیشین (مانند گشتاورها) و یا ساختار مشاهدات و... تعیین می‌شود. پس از مشاهده نمونه، اطلاعات در مورد Θ (توزیع پیشین) تصحیح می‌گردد.

---

1. Koziol- Green Model(KGM)
2. Right Censored data
3. Prior density
4. (Nonrandomaized) Decision Rule
5. Action Space
6. Decision Space
7. Loss Function
8. Risk Function
9. Prior Distribution

توزیع پیشین تصحیح شده را توزیع پسین[1] $g_{X|\theta}(\theta)$ می‌نامند. حال در رویکرد بیز بر اساس این توزیع پسین تصمیم‌گیری می‌شود. برای مثال فرض کنید قطعات تولیدی یک کارخانه دارای توزیع نمایی با میانگین θ ساعت باشند. به علت فرسودگی دستگاه‌ها این میانگین نیز در سال تغییر می‌کند و تغییرات آن طبق یک تابع توزیع احتمال $q(\theta)$ می‌باشد. مخاطره بیز[2] تابع تصمیم d^* با توجه به توزیع پیشین $q(\theta)$ به صورت امید ریاضی $R(d^*,\theta)$ نسبت به $q(\theta)$ تعریف می‌شود. اصل مخاطره بیز به این معنی است که اطلاعات تابع مخاطره در یک عدد (مخاطره بیز) خلاصه می‌شود و دو تصمیم به وسیلهٔ مقایسه این مخاطره بیز با یکدیگر مقایسه می‌شوند. در واقع یک تصمیم بیز[3]، مرتبط با توزیع پیشین $q(\theta)$ و تابع زیان L، تصمیمی است که مخاطره بیز را مینیمم کند. یک قاعده تصمیم d^* کمینه - بیشینه[4] است اگر تابع مخاطره یعنی

$$\sup_{\theta\in\Theta} R(d^*,\theta) = \inf_d \sup_{\theta\in\Theta} R(d,\theta)$$

را مینیمم کند. به عبارت دیگر تصمیم کمینه - بیشینه آن تصمیمی است که در بدترین حالت تابع مخاطره $R(\theta,d)$ نسبت به δ، بهترین حالت را انتخاب می‌کند و در حقیقت یک رویکرد محافظه‌کارانه برای انتخاب بهترین تصمیم است. برای کسب اطلاعات بیشتر در مورد تصمیم‌های بیز و کمینه - بیشینه به [13] مراجعه کنید.

کران پایین برای مقادیر کمینه - بیشینه

در ابتدا داده‌های سانسور شده از راست، مدل کوزیول - گرین و زمان توقف را معرفی می‌کنیم و سپس با استفاده نامساوی کرامر - رائو یک کران پایین برای مخاطره کمینه - بیشینه بر پایهٔ یک نمونه تصادفی با زمان توقف تصادفی، در حالتی که فضای پارامتر بریده باشد، پیدا می‌کنیم. در پایان کاربردهایی از این مفهوم را بیان می‌کنیم.

تعریف 1: [5] (برش تصادفی - داده‌های سانسور شده از راست): اگر $X_1,\dots,X_n$ و $Y_1,\dots,Y_n$ متغیرهای تصادفی مستقل و هم‌توزیع به ترتیب با تابع توزیع F و G باشند (Y_i زمان برش مربوط به X_i است)، X_i و Y_i قابل مشاهده نیستند و در این حالت معمولاً فقط (W_n,D_n) و... و (W_1,D_1) را مشاهده می‌کنیم که در آن برای هر i، W_i و D_i به صورت زیر تعریف می‌شوند:

$$W_i = \min(X_i, T_i)$$

$$D_i = I(X_i \le T_i) = \begin{cases} 1 & X_i \le T_i \\ 0 & X_i > T_i \end{cases}$$

به سادگی می‌توان نشان داد که تابع چگالی و اطلاع فیشر تحت مدل فوق به صورت زیر است

$$h_\theta(w,\delta) = \left[f_\theta(w)\,\overline{G}_\theta(w)\right]^{\delta} \times \left[g_\theta(w)\,\overline{F}_\theta(w)\right]^{1-\delta}, \quad w>0,\ \delta=0,1,$$

$$I_{(W,D)}(\theta) = \int_R \left(\frac{\partial}{\partial\theta}\log(\overline{G}_\theta f_\theta)\right)^2 \overline{G}_\theta f_\theta\,dx + \int_R \left(\frac{\partial}{\partial\theta}\log(\overline{F}_\theta g_\theta)\right)^2 \overline{F}_\theta g_\theta\,dx$$

که در آن $\overline{F}=1-F$ و $\overline{G}=1-G$.

تعریف 2: (مدل کوزیول – گرین) (KGM) [12]: گوییم زوج (W,D) از مدل کوزیول - گرین پیروی می‌کند، هرگاه عدد $\beta>0$ موجود باشد به طوری که

$$1-G_{\theta,\beta}(x) = (1-F_\theta(x))^{\beta}.$$

فرض کنید $X_1,X_2,\dots$ دنباله‌ای از متغیرهای تصادفی مستقل و هم‌توزیع با خانواده توزیع احتمال $\{P_\theta,\theta\in\Theta\}$ باشند که در آن Θ یک فاصله باز از اعداد حقیقی است. مسئله مورد علاقه، برآورد کمینه -

1. Posterior Distribution
2. Bayes Risk
3. Bayes Decision
4. Minimax Decision

بیشینه پارامتر θ^s ($s \neq 0$ عدد حقیقی معلوم) تحت تابع زیان وزنی زیر است:

$$L(T_N,\theta) = w(\theta)(T_N - \theta^s)^2 \qquad (1)$$

به طوری که T_N، یک برآوردگر θ^s تحت نمونه‌ای به حجم تصادفی N است. فرض کنید $c(\theta)$ متوسط هزینه یک مشاهده باشد. معمولاً فرض می‌شود که هزینه مشاهده هر X_i، تابعی از X_i مانند $x(X_i)$ باشد، از این رو متوسط هزینه مشاهده تمام نمونه $X_1,\dots,X_N$ بنا بر لم والد[1] [13] به صورت زیر است:

$$E_\theta\left(\sum_{i=1}^{N}(X_i)\right) = E_\theta(N)E_\theta(x(X_1))$$

که در آن N زمان توقف[2] است.

از این رو مخاطره کل T_N در تخمین θ^s برابر است با:

$$R(T_N,\theta) = E_\theta[L(T_N,\theta)] + C(\theta)E_\theta(N) \qquad (2)$$

تعریف 3: متغیر تصادفی N را زمان توقف نسبت به دنباله $\{X_i\}$ها گوییم، اگر رخداد پیشامد "$N = n$" ($\{W : N(W) = n\}$) به وسیله؛ مشاهدهٔ مقادیر تحقق فرآیند تا بار nام $(X_1,\dots,X_n)$ تعیین شود.

برای مثال فرض کنید به یک سکه اریب در هر دقیقه یک تلنگر زده می‌شود، اگر متغیر تصادفی N_1 را پیشامد مشاهدهٔ خط در لحظه‌ای که به سکه ضربه زده می‌شود، تعریف می‌کنیم در این صورت N_1 زمان توقف آزمایش است. حال اگر $T_n = T(X_1,\dots,X_n)$، برآوردگر θ^s که $s \neq 0$ باشد، با در اختیار داشتن دنبالهٔ آماره‌های $\{T_n\}$ و زمان توقف N ساختاری از برآوردگرهای دنباله‌ای[3] T_N حاصل می‌شود. قضیهٔ زیر کران پائین مقادیر کمینه - بیشینه T_N معرفی می‌کند.

1. Wald lemma
2. Stopping Time
3. Sequential Estimator

قضیه 1: تحت شرایط برقراری نامساوی کرامر - رائو می‌دانیم که:

$$E_\theta(T_N - \theta^s)^2 \geq b^2(\theta) + \frac{(s\theta^{s-1}+b(\theta))^2}{E_\theta(N)I(\theta)}$$

که در آن

$$I(\theta) = Var_\theta\left(\frac{\partial}{\partial\theta}\log P_\theta(X_1)\right),$$
$$b(\theta) = E_\theta(T_N) - \theta^s$$

در این صورت اگر $w(0)$ و $C(0)$ به ترتیب نشان دهندهٔ تابع وزن و تابع هزینه باشند:

الف) اگر $0 < C(\theta) \leq s^{-2}\theta^{2s+2}I(\theta)w(\theta)$ به ازای یک $\delta > 0$ و برای هر $\theta \in \Theta$ که $\theta < \delta$ ($\theta > k$ به ازای بعضی از مقادیر $k > 0$) و

$$\lim_{\theta\to 0}\theta^{-2}\frac{C(\theta)}{I(\theta)},\ \lim_{\theta\to 0}W(\theta)\theta^{2s}$$
$$\left(\lim_{\theta\to\infty}\theta^{-2}\frac{C(\theta)}{I(\theta)},\ \lim_{\theta\to\infty}w(\theta)\theta^{2s}\right)$$

موجود و متناهی باشند آن گاه:

$$\limsup_{\theta\to 0,\{\infty\}}\left\{E_\theta[T_N - \theta^s]^2 w(\theta) + C(\theta)E_\theta(N)\right\}$$
$$= \limsup_{\theta\to 0,\{\infty\}} R(T_N,\theta)$$
$$\geq \lim_{\theta\to 0,\{\infty\}} |s|\theta^{s-1}\sqrt{\frac{C(\theta)w(\theta)}{I(\theta)}}\left(2 - \frac{|s|}{\theta^{s+1}}\sqrt{\frac{C(\theta)}{I(\theta)w(\theta)}}\right)$$

ب) اگر $C(\theta) \leq s^{-2}\theta^{2s+2}I(\theta)w(\theta)$ به ازای یک $\delta > 0$، برای هر $\theta \in \Theta$ که $\theta < \delta$ ($\theta > k$، به ازای بعضی از مقادیر $k > 0$) و $\lim_{\theta\to 0}w(\theta)\theta^{2s}$ ($\lim_{\theta\to\infty}w(\theta)\theta^{2s}$) موجود و متناهی باشد، آن گاه:

$$\limsup_{\theta\to 0,\{\infty\}}\left\{E_\theta[T_N-\theta^s]^2 w(\theta)+C(\theta)E_\theta(N)\right\}$$
$$=\limsup_{\theta\to 0,\{\infty\}} R\left(T_N,\theta\right)$$
$$\ge \lim_{\theta\to 0,\{\infty\}} w(\theta)\theta^{2s}$$

کران‌های پایین مقادیر کمینه - بیشینه در قضیه فوق را به اختصار ICRB [1] می‌نامیم. قضیهٔ فوق نشان می‌دهد که کران پایین مخاطره کمینه - بیشینه هر برآوردگر که در نامساوی الف و ب صدق کند، بستگی به برآوردگر و به زمان توقف ندارد، از این رو قضیه معرفی می‌شود.

قضیه 2: اگر T_N برآوردگری باشد که سوپریمم مخاطرهٔ آن برابر ICRB باشد، آن گاه T_N دنباله‌ای از برآوردگرهای کمینه - بیشینه θ^s است.

اثبات: از این که $ICRB$ به T_N بستگی ندارد، داریم:

$$\inf_{T'}\sup_{\Theta} R\left(T',\theta\right)=ICRB$$
$$=\sup_{\Theta} R\left(T_N,\theta\right)$$

لذا طبق تعریف 3، برآوردگر T_N، کمینه - بیشینه است.

در ادامه مقاله کاربردهایی از قضیه 1 را در غالب چند نتیجه و مثال ارائه می‌دهیم.

کاربرد 1: تحت داده‌های سانسور شده از راست، بر اساس یک نمونه n تایی، فرض می‌کنیم متغیر برش عددی ثابت مانند $y_o>0, Y_i=y_o$ باشد. با تعریف:

$$N_k=\min\left\{m:\sum_{i=1}^{m} D_i\ge k\right\}\quad (3)$$

تابع درست‌نمایی به فرم زیر است:

$$L(\theta)=\prod_{i=1}^{N_k}\{f_\theta(x_i)\overline{G}(x_i)\}^{d_i}\{g(x_i)\overline{F}_\theta(x_i)\}^{1-d_i},$$
$$d_i=0,1$$

حال با فرض $N_k=n$ ، $0\le k\le n, k=\sum_{i=1}^{n} d_i$،
تحت مدل از راست سانسور شدهٔ:

$$W_i=\min(X_i,y_o)\ ,\ D_i=I(X_i\le y_o),\ y_o>0$$

تابع درست‌نمایی بالا به صورت زیر تبدیل خواهد شد:

$$L(\theta)=\begin{cases}\prod_{i=1}^{k} f_\theta(x_{(i)})\{\overline{F_\theta}(y_o)\}^{n-k} & 0<k\le n,\\ & x_{(1)}<\ldots<x_{(k)}<y_o\\ \{\overline{F_\theta}(y_o)\}^{n} & k=0,\ y_o<x_{(1)}\end{cases}$$

که در آن $X_{(i)}$ ها آماره‌های ترتیبی نمونهٔ $X_1,\ldots,X_n$ هستند.

فرض می‌کنیم چگالی X_i ها به فرم زیر باشد:

$$f_\theta(x)=\tfrac{1}{\theta}e^{-x/\theta}\ x>0,\theta\in\Theta=(0,\overline{\theta}),\overline{\theta}\le\infty$$

با قرار دادن تابع چگالی بالا در $L(\theta)$ داریم:

$$L(\theta)=\begin{cases}\frac{1}{\theta^k}e^{-\frac{1}{\theta}\left(\sum_{i=1}^{k}x_{(i)}+(n-k)y_o\right)} & 0<k\le n,\\ & x_{(1)}<\ldots<x_{(k)}<y_o\\ e^{-ny_o/\theta} & k=0,\ y_o<x_{(1)}\end{cases}$$

اگر $k=0$ باشد، تابع درست‌نمایی به طور یکنوا است و برآورد حداکثر درست‌نمایی وجود ندارد، ولی برای $k>0$ می‌توان نشان داد برآوردگر حداکثر درست‌نمایی θ برابر

1. Improper Cramer Rao bound

$$\tilde{\theta}_{N_K}=\tfrac{1}{k}\{\sum_{i=1}^{k}X_{(i),N_k}+(N_k-k)y_o\}$$

است. اکنون از (3) به سادگی ملاحظه می‌شود که N_k دارای توزیع احتمال دوجمله‌ای منفی با تابع توزیع احتمال و امید ریاضی و واریانس زیر است:

$$p(N_k=n)=\binom{n-1}{k-1}p^k(1-p)^{n-k},\ n=k,k+1,\ldots$$

$$E_p(N_k)=\tfrac{k}{p},Var_p(N_k)=\tfrac{kq}{p^2}.$$

حال با محاسبات مستقیم می‌توان نشان داد که آماره‌های مرتب $X_{(1),N_k},\ldots,X_{(k),N_k}$ به شرط $N_k=n$ دارای چگالی احتمال مساوی با چگالی توأم آماره‌های مرتب از یک نمونهٔ تصادفی شامل k مشاهده مستقل و هم‌توزیع از توزیع بریده‌شده[1] با تابع چگالی زیر هستند:

$$Q(u)=(\theta p)^{-1}e^{-\frac{u}{\theta}},0<u\le y_o$$

که در آن

$$\begin{aligned}p&=P(D=1)\\&=P(X\le y_o)\\&=1-e^{-y_o/\theta},\theta\in\Theta,y_o>0\end{aligned}$$

به سادگی می‌توان نشان داد که وقتی U دارای چگالی Q است، آن گاه:

$$E_\theta(U)=\theta-y_o\tfrac{q}{p},Var_\theta(U)=\theta^2-y_o^2\tfrac{q}{p^2},$$

که در آن $q=1-p$. در نتیجه با تعریف:

$$\hat{\theta}_{N_K}=\tfrac{1}{k+1}\{\sum_{i=1}^{k}X_{(i),N_k}+(N_k-k)y_o\}$$

داریم:

$$\begin{aligned}E(\hat{\theta}_{N_K})&=E\left[E(\hat{\theta}_{N_K}|N_k)\right]\\&=E\left[\tfrac{1}{k+1}E\left(\sum_{i=1}^{k}X_i,N_k+(N_k-k)y_o\Big|N_k\right)\right]\\&=E\left[\tfrac{k}{k+1}E(U)+\tfrac{1}{k+1}(N_k-k)y_o\right]\\&=\tfrac{k}{k+1}\theta\end{aligned}$$

و

$$\begin{aligned}Var_\theta(\hat{\theta}_{N_K})&=E\left[Var_\theta(\hat{\theta}_{N_K}|N_k)\right]\\&+Var_\theta\left[E(\hat{\theta}_{N_K}|N_k)\right]\\&=E\left(Var_\theta\left(\tfrac{1}{k+1}\sum_{i=1}^{k}X_{i,N_k}+(N_k-k)y_o\Big|N_k\right)\right)\\&+Var\left[\tfrac{k}{k+1}(\theta-y_o\tfrac{p}{q})+(N_k-k)\right]\\&=E\left(\tfrac{k}{(k+1)^2}Var_\theta(U)\right)+\left(\tfrac{y_o}{k+1}\right)^2Var(N_k)\\&=\tfrac{k\theta^2}{(k+1)^2}.\end{aligned}$$

با توجه به روابط بالا داریم:

$$E(\hat{\theta}_{N_K}-\theta)^2=\tfrac{\theta^2}{k+1}\qquad(4)$$

حال فرض کنید $C(\theta)$ و N_K به ترتیب نشان دهندهٔ تابع هزینه و زمان توقف باشند، با توجه به مطالب ذکر شده در بالا نتیجه زیر را داریم:

نتیجه 1: اگر

$$\sup_{\Theta}\left(\tfrac{C(\theta)}{p}\right)=\lim_{\theta\to 0}\left(\tfrac{C(\theta)}{p}\right)=\tfrac{1}{(k+1)^2}$$

آن گاه $\hat{\theta}_{N_K}$ (بر پایه برآوردگر حداکثر درست‌نمایی $\tilde{\theta}_{N_k}$) دنباله‌ای از برآوردگرهای کمینه – بیشینه θ تحت تابع زیان $L(\tilde{\theta},\theta)=\theta^{-2}(\hat{\theta}_{N_K}-\theta)^2$ است. که در آن k یک مقدار صحیح مثبت و $0<\theta_1\le\infty,\Theta=(0,\theta_1)$.

اثبات: از رابطه (4) داریم.

---

1. Truncated

$$E\,(\hat{\theta}_{N_K}-\theta)^2=\frac{\theta^2}{k+1}$$

با جای‌گذاری عبارت بالا در (2) و به دلیل اینکه

از رابطهٔ: $\sup_{\Theta}\left(\frac{C(\theta)}{p}\right)=\frac{1}{(k+1)^2}$

$$R\left(\hat{\theta}_{N_k},\theta\right)=\frac{1}{k+1)}+\frac{kC(\theta)}{p}$$

داریم

$$\sup_{\Theta} R\left(\hat{\theta}_{N_K},\theta\right)=\frac{2k+1}{(k+1)^2} \qquad (5)$$

از طرفی چون G تابع توزیع متغیر تصادفی تباهیده Y در نقطه y_o است، با استفاده از اطلاع فیشر داده‌های سانسور شده از راست که در تعریف 1 ارائه شد داریم:

$$\begin{aligned}I_{(W_1,D_1)}(\theta)&=\int_{-\infty}^{+\infty}\left(\frac{\partial}{\partial\theta}\log f_\theta\right)^2\bar{G}f_\theta dx\\&+\int_{-\infty}^{+\infty}\left(\frac{\partial}{\partial\theta}\log\bar{F_\theta}\right)^2\bar{F_\theta}g dx\\&=\int_0^{y_o}\left(\frac{\partial}{\partial\theta}\log f_\theta\right)^2 f_\theta dx+\frac{\left(\frac{\partial}{\partial\theta}\bar{F_\theta}(y_o)\right)^2}{\bar{F_\theta}(y_o)}.\end{aligned}$$

با محاسبات مستقیم می‌توان نشان داد که:

$$I_{(W_1,D_1)}(\theta)=\frac{p}{\theta^2},\ p=1-e^{-y_o/\theta},\ y_o>0$$

حال با توجه با اینکه $\lim_{\theta\to 0}\left(\frac{C(\theta)}{p}\right)=(k+1)^{-2}$ و $I_{(W_1,D_1)}(\theta)=\theta^{-2}p$، از قسمت الف قضیه 1، به سادگی نامساوی زیر، وقتی که θ به سمت صفر میل می‌کند، نتیجه می‌شود:

(6)

$$\sup_{\Theta}[E\,(\hat{\theta}_M-\theta)^2\theta^{-2}+C(\theta)E_\theta(M)]\geq\frac{2k+1}{(k+1)^2}$$

که در آن $\tilde{\theta}_M$، یک برآوردگر دنباله‌ای در تخمین θ است.

با مقایسه روابط (5) و (6) داریم:

$$\inf_{\hat{\theta}_M}\sup_{\Theta}R(\hat{\theta}_M,\theta)=\frac{2k+1}{(k+1)^2}$$

بنابراین طبق قضیه 2، $\hat{\theta}_{N_K}$ دنباله‌ای از برآوردگرهای کمینه - بیشینه θ است.

کاربرد 2: فرض کنید $X_1,X_2,\ldots$ متغیرهای تصادفی مستقل و هم‌توزیع با تابع توزیع مطلقاً پیوسته F باشند که به وسیلهٔ متغیرهای مستقل و هم‌توزیع $Y_1,Y_2,\ldots$ با تابع توزیع G، از راست سانسور می‌شوند. همان طور که می‌دانیم تحت این مدل متغیرهای قابل مشاهده زوج‌های (W_i,D_i) هستند که به صورت زیر تعریف می‌شوند:

$$W_i=\min\left(X_i,Y_i\right),\ D_i=I\left(X_i\leq Y_i\right)$$

اگر (W_i,D_i) از مدل کوزیول - گرین تبعیت کنند یعنی عدد مثبت و ثابت $\beta>0$ موجود باشد به طوری که:

$$1-G_{\theta,\beta}(x)=\left(1-F_\theta(x)\right)^\beta$$

نتیجهٔ زیر را داریم:

نتیجه 2: فرض کنید تابع هزینه مقداری ثابت باشد $(0<C_o<1, C(p)=c_o)$ و با احتمال یک $N_o=1/\sqrt{c_o}-1$.

اگر فضای پارامتری به صورت $0<\bar{p}\leq 1, Q=(0,\bar{p})$ باشد، آن گاه:

$$\hat{p}_{N_o}=\frac{1}{N_o+1}\sum_{i=1}^{N_0}R_f(W_i)$$

دنباله‌ای از برآوردگرهای کمینه - بیشینه در تخمین p تحت تابع زیان $L(\tilde{p},p)=(\tilde{p}-p)^2/p^2$

نسبت به مدل کوزیول - گرین هستند، که در آن $p = P(D=1)$ نسبت مشاهدات سانسور نشده است و

$$R_{f_\theta}(x) = -\log \bar{F}_\theta(x),\ x \in \{x; f(x)>0\}.$$

اثبات: اگر $H_{W_1}(x)$ تابع توزیع $W_1=\min(X_1,Y_1)$ باشد تحت KGM ($\bar{G}_\theta = \bar{F}_\theta^{\beta}$، $\beta>0$) داریم:

$$\begin{aligned} H_w(x) &= 1-(\bar{F}_\theta \bar{G}_\theta)(x) \\ &= 1-\bar{F}_\theta^{(1+\beta)}(x) \\ &= 1-\left(\bar{F}_\theta\right)^{1/p} \end{aligned}$$

که در آن $p=1/(1+\beta)$؛ بنابراین
$\bar{F}_\theta(x) = \left(1-H_{W_1}(x)\right)^p$

و در پی آن

$$\begin{aligned} &P\left(R_{f_\theta}(W_1)\le t\right) \\ &= P\left(-\log \bar{F}_\theta(W_1) \le t\right) \\ &= P\left(\bar{F}_\theta(W_1)\ge e^{-t}\right) \\ &= P\left(\left[\bar{F}_\theta(W_1)\right]^{\frac{1}{p}} \ge e^{-t/p}\right) \\ &= P\left(H(W_1)\le 1-e^{-t/p}\right) \\ &= 1-e^{-t/p}, t>0 \end{aligned}$$

زیرا توزیع $H(W_1)$ یکنواخت در بازه $(0,1)$ است، بنابراین از رابطهٔ بالا نتیجه می‌شود که $R_{f_\theta}(w_1)\sim E(p)$، بنابراین میانگین و واریانس $\tilde{p}_n$ به ترتیب برابر است با $E(\tilde{p}_n) = \frac{n}{n+1}p$ و $Var(\tilde{p}_n) = \frac{n}{(n+1)^2}p^2$. به سادگی می‌توان نشان داد که:

$$E(\tilde{p}_n - p)^2 p^{-2} = \frac{1}{n+1} \quad (7)$$

اما می‌توان نشان داد که تحت KGM ، W_1 مستقل از D_1 $(D_1\sim b(p))$ است؛ بنابراین:

$$I_{(W_1,D_1)}(p) = I_{W_1}(p)+I_{D_1}(p)$$

که در آن

$$I_{W_1}(p) = -E\left[\frac{\partial^2}{\partial\theta^2} h_{W_1}(X,\theta)\right], \quad I_{D_1}(p)=\frac{1}{p(1-p)}$$

به سادگی می‌توان نشان داد که $I_{W_1}(p)=\frac{1}{p^2}$ و بنابراین اطلاع فیشر تحت مدل کوزیول - گرین برابر است با

$$I_{(W_1,D_1)}(p) = \frac{1}{p^2(1-p)}.$$

اگر در قضیه 1، قسمت الف عبارت‌های زیر را جایگزین کنیم:

$$\begin{aligned} &W(p)=p^{-2}, \\ &C(p)=C_o, \\ &I_{(W_1,D_1)}(\theta) = \frac{1}{p^2(1-p)} \end{aligned}$$

برای هر مقدار کمینه – بیشینه $\tilde{P}_M$ در مسئله تخمین p داریم:

$$\inf_{\hat{\pi}_M} \sup_{\Theta} \left\{ E(\hat{\pi}_M - \pi)^2 \pi^{-2} + c_o E_\theta(M) \right\} \geq 2\sqrt{c_o} - c_o \qquad (8)$$

که در آن M زمان توقف است.

از طرف دیگر بنا به (7) داریم:

$$E(\hat{\pi}_N - \pi)^2 = \frac{\pi^2}{n_o + 1}$$

بنابراین سوپریمم مخاطره $\tilde{\pi}_{N_o}$ به صورت زیر است:

$$\sup_{\Theta} R(\hat{\pi}_{N_o}, \pi) = \sup_{\Theta} \left\{ E(\hat{\pi}_{N_o} - \pi)^2 \pi^{-2} + c_o E(N_o) \right\} = \frac{1}{n_o+1} + c_o n_o$$

اکنون با قرار دادن $n_o = 1/\sqrt{c_o} - 1$ در رابطه بالا و مقایسه با رابطه (8)، بنا بر قضیهٔ 2، $\tilde{\pi}_{N_o}$ دنباله‌ای از برآوردگرهای کمینه - بیشینه در تخمین π است.

کاربرد 3: فرض کنید $X_1, X_2, \ldots$ متغیرهای تصادفی مستقل و هم‌توزیع با توزیع احتمال زیر باشد.

$$P(X_i = 1) = p_1,$$
$$P(X_i = 0) = q_1 = 1 - p_1, \quad i = 1, 2, \ldots$$

که در آن $0 < \bar{p}_1 < 1, \Theta = (0, \bar{p}_1)$.

مسئله مورد علاقه، یافتن دنباله‌ای از برآوردگرهای کمینه - بیشینه معکوس احتمال موفقیت $(\theta = 1/p_1)$ تحت تابع زیان مربع خطا با تابع وزن $w(\theta) = \theta^{-2}$ است.

فرض کنید هزینه هر مشاهده به صورت زیر باشد.

$$\xi(X_i) = \begin{cases} 0 & X_i = 0 \\ c_1 & X_i = 1 \end{cases}$$

که در آن $c_1 \in (0, 1)$ مقداری ثابت است. از این رو متوسط هزینه مشاهدات با زمان توقف M بنا بر لم والد برابر است با:

$$K(\theta) = E\left[\sum_{i=1}^{M} \xi(X_i)\right] = (c_1/\theta) E_\theta(M)$$

و مخاطره کل $\tilde{\theta}_M$ از رابطه (1) به صورت زیر است:

$$R(\tilde{\theta}_M, \theta) = E(\tilde{\theta}_M - \theta)^2 \theta^{-2} + \left(\frac{c_1}{\theta}\right) E_\theta(M)$$

فرض کنید زمان توقف را به صورت زیر تعریف کنیم

$$N_r = \min\left\{ m : \sum_{i=1}^{m} X_i = r \right\} \qquad r = 1, 2, \ldots$$

که در آن r مقدار ثابتی است. حال با تعریف

$$\hat{\theta}_{N_r} = N_r / (r+1)$$

بنا به این حقیقت که $N_r \sim NB(r, 1/\theta)$، به سادگی می‌توان نشان داد که

$$R(\tilde{\theta}_{N_r}, \theta) = \frac{r(1 - \theta^{-1}) + 1}{(1+r)^2} + c_1 r$$

بنابراین:

$$\sup_{\theta > 1/\bar{p}_1} R(\hat{\theta}_{N_r}, \theta) = \frac{1}{r+1} + c_1 r \qquad (9)$$

از طرف دیگر، به دلیل اینکه

$$C(\theta)=E_\theta(\xi(X_1))=c_1/\theta,$$
$$I(\theta)=1/\theta^2(\theta-1)$$

از قسمت الف قضیه 1 وقتی که θ به سمت بی‌نهایت میل می‌کند، داریم:

$$\inf_{\tilde\theta_M}\sup_{\theta>1/\bar p_1} R(\tilde\theta_M,\theta)\geq\sqrt{c_1}(2-\sqrt{c_1}) \tag{10}$$

حال با مقایسه روابط (9)و (10) برآوردگر $\tilde\theta_{N_\rho}$ بنا بر قضیه 2، دنباله‌ای از برآوردگرهای کمینه - بیشینه پارامتر $\theta=1/p_1$ است هرگاه $\rho=(1-\sqrt{c_1})/\sqrt{c_1}$.

کاربرد 4: در مثال قبل، علاقه‌مندیم که یک برآوردگر کمینه - بیشینه برای نسبت بخت‌ها[1] p_1/q_1 را وقتی که $q_1\in(0,\bar q_1)$ است، به دست آوریم. اما به دلیل اینکه $X_i'=1-X_i, \theta=1/q_1-1$ دارای توزیع برنولی با احتمال موفقیت $q_1=P(X_i'=1)$ هستند، مسئله تبدیل به برآورد عکس احتمال موفقیت $(\theta=1/q_1)$ در مثال قبلی بر اساس دنباله‌ای از متغیرهای مستقل و هم‌توزیع $X_1,X_2,\ldots$ است؛ بنابراین آماره:

$$\tilde\theta_{N_\rho}=\frac{N'_\rho}{\rho+1}-1$$

با زمان توقف:

$$N'_\rho=\min\left\{m:\sum_{i=1}^{m}(1-X_i)=\rho\right\}$$

1. Odd ratio

یک برآوردگر کمینه – بیشینه p_1/q_1 تحت تابع زیان مربع خطا با تابع وزن $w(\theta)=q_1^2=(1+\theta)^{-2}$ و تابع هزینه $C(\theta)=c_1q_1=c_1(1+\theta)^{-1}$ است که در آن $\rho=(1-\sqrt{c_1})/\sqrt{c_1}$ و هزینه مشاهده X_i به صورت زیر است:

$$\xi(X_i)=\begin{cases}c_1 & X_i=0\\ 0 & X_i=1, i=1,2,\ldots\end{cases}$$

کاربرد 5: فرض کنید $X_1,X_2,\ldots$ دنباله‌ای از متغیرهای تصادفی مستقل و هم‌توزیع از خانواده توزیع‌های نمایی به فرم

$$f_\theta(x)=W(x)\exp\{h(\theta)T(x)+a(\theta)\},$$
$$E(T(X))=\theta$$

و با تابع هزینه ثابت $C(\theta)=c$، $c\in R^+$ و زمان توقف N باشند، در این صورت به سادگی می‌توان نشان داد که مخاطره برآوردگر $\tilde\theta_{N,b}=\frac{1}{n+b}\sum_{i=1}^{N}T(X_i)$ تحت تابع زیان خطای با وزن، به صورت زیر است:

$$\begin{aligned}&R(\tilde\theta_{n,b},\theta)\\&=E_\theta(\tilde\theta_{n,b}-\theta)^2 I(\theta)+cE_\theta(N)\\&=\frac{n_o+b^2\theta^2 I(\theta)}{(n_o+b)^2}+cn_o\end{aligned}$$

و نتیجه زیر برقرار است:

نتیجه 3: فرض کنید فضای پارامتری به فرم زیر باشد:

$$\Theta=(0,\theta_1),\ 0<\theta_1\le\infty$$
$$(\Theta=(\theta_2,\infty),\ 0\le\theta_2<\infty)$$

الف) اگر $\lim_{\theta\to 0}\theta^2 I(\theta)=\infty$ ($\lim_{\theta\to\infty}\theta^2 I(\theta)=\infty$) و با احتمال یک زمان توقف N مساوی $1/\sqrt{c}$ باشد، آن گاه $\tilde{\theta}_{N,o}=\frac{1}{N}\sum_{i=1}^{N}T(X_i)$ برآوردگر کمینه - بیشینه $\theta\in\Theta$ است.

ب) اگر $\sup_{\theta\in\Theta}\theta^2 I(\theta)=\lim_{\theta\to 0}\theta^2 I(\theta)=\frac{1}{b}$ ($\sup_{\theta\in\Theta}\theta^2 I(\theta)=\lim_{\theta\to\infty}\theta^2 I(\theta)=\frac{1}{b}$) و $0<c<1/b^2$،

و برای بعضی از مقادیر $b>0$، زمان توقف N با احتمال یک $1/\sqrt{c}-1$ باشد آن گاه:

$$\tilde{\theta}_{N,b}=\frac{1}{N+b}\sum_{i=1}^{N}T(X_i)$$

یک برآوردگر کمینه – بیشینه $E(T(X))=\theta$ است.

مثال 1: فرض کنید $X_1,X_2,\ldots$ دنباله‌ای از متغیرهای مستقل و هم‌توزیع با تابع چگالی احتمال زیر باشند:

$$P_\theta(X=x)=\frac{e^{-\theta}\theta^{x}}{x},\ x=0,1,2,\ldots,$$

که در آن $\theta\in(\theta_2,\infty)$ به طوری که $0\le\theta_2<\infty$. می‌دانیم که برآوردگر $T(X)=X$ یک برآوردگر ناریب میانگین توزیع پواسون $(E_\theta(X)=\theta)$ است. از این رو و به دلیل اینکه

$$\lim_{\theta\to\infty}\theta^2 I(\theta)=\lim_{\theta\to\infty}\theta^2\tfrac{1}{\theta}=\infty$$

بنا بر قسمت الف نتیجهٔ 3، $\tilde{\theta}_{N,o}$ برآوردگر کمینه - بیشینهٔ θ است.

مثال 2: فرض کنید $X_1,X_2,\ldots$ دنباله‌ای از متغیرهای مستقل و از توزیع نرمال با میانگین صفر و واریانس θ باشند که در آن فضای پارامتر به صورت $\Theta=(0,\theta_1)$ که $0<\theta_1\le\infty$ یا $\Theta=(\theta_2,\infty)$ که $0\le\theta_2<\infty$ باشد، با تعریف $T(X)=X^2$ و با توجه به اینکه $I(\theta)=1/2\theta^2$، از قسمت ب نتیجهٔ 3 می‌توان نشان داد که اگر $c<1/4$، آن گاه $\tilde{\theta}_{N,2}$ یک برآورگر کمینه – بیشینه $E(T(X))=\theta$ است.

با استفاده از قضیه 1 نتایج مشابه فراوانی را مانند نتایج قبلی در حالتی که تابع هزینه وابسته به پارامتر مجهول باشد، می‌توان به دست آورد. به‌عنوان مثال، نتیجه زیر را به سادگی می‌توان نشان داد.

نتیجه 4: اگر فضای پارامتری به فرم زیر باشد:

$$\Theta=(0,\theta_1),\ 0<\theta_1\le\infty$$
$$(\Theta=(\theta_2,\infty),\ 0\le\theta_2<\infty)$$

و رابطهٔ زیر برقرار باشد:

$$\sup_{\Theta}C(\theta)=\lim_{\theta\to 0}C(\theta)=c\ \ ,c>0$$
$$(\sup_{\Theta}C(\theta)=\lim_{\theta\to\infty}C(\theta)=c)$$

آن گاه $\tilde{\theta}_{N,o}=\frac{1}{N}\sum_{i=1}^{N}T(X_i)$ دنباله‌ای از برآوردگرهای کمینه - بیشینهٔ $E(T(X))=\theta\in\Theta$ است که در آن زمان توقف N با احتمال یک برابر $1/\sqrt{c}$ است.

نتیجه‌گیری

در این مقاله کاربردهایی از نامساوی والد – کرامر - رائو (که یک کران پایین برای مخاطره کمینه بیشینه بر پایهٔ یک نمونه تصادفی با زمان توقف تصادفی، در حالتی که فضای پارامتر بریده باشد، پیدا می‌کند)، در محاسبه برآوردگرهای کمینه - بیشینه تحت داده‌های سانسور شده و مدل کوزیول - گرین و خانواده توزیع‌های نمایی را معرفی می‌کنیم. محققان می‌توانند از کاربردهای دیگر این نامساوی در شاخه‌های دیگر آمار بهره جویند.

منابع

[1] Akahira, M. and Ohyauchi, N. (2002). Information Inequalities for the Bayes Risk for a Family of Non-Regular Distributions. Annals of the Institute of Statistical Mathematics, **54**(4), 806–815.

[2] Akahira, M. and Sato, M. (1996). An information inequality for the Bayes risk. The Annals of Statistics, **24**(5), 2288-2295.

[3] Alvo, M. (1977). Bayesian sequential estimation, Ann. Statist., **5**, 955–968.

[4] Brown, L. D. and Gajek, L. (1990). Information Inequalities for the Bayes Risk. The Annals of Statistics, **18**(4), 1578-1594.

[5] Eubank, R. L and Lariccia, V. N. (1982). Location and Scale parameter estimatiom from random censored data. Comm. Stat. A-Theory Methods, **11**, pp. 2869-2888.

[6] Gajek, L. (1987). On minimax value in the scale model with truncated data. Ann. Statist, **16**, pp. 669-677.

[7] Gajek, L. and Ghater, U. (1991). Estimating the scale parameter under random censorship. Statistics, **22**, pp. 529-549.

[8] Gardiner, J. C. and Susarla, V. (1984). Risk-efficient estimation of the mean exponential survival time under random censoring, Proc. Nat. Acad. Sci., U.S.A., **81**, 5906–5909.

[9] Gardiner, J. C. and Susarla, V. (1991). Some asymptotic distribution results in time-sequential estimation of the mean exponential survival time. Canad. J. Statist., **19**, 425-436 .

[10] Gardiner, J. C. and Susarla, V. and Ryzin, J. van. (1986). Time sequential estimation of the exponential mean under random withdrawals, Ann. Statist., 14, 607–618.

[11] Kaluszka, M. (2007). Information inequalities for the Bayes risk of predictors, Probability and Mathematical Statistics, **27**(2), 167-179.

[12] Koziol, J. A. and Green, S. B. (1976). A Cramer-Rao Mises statistic for randomly censored data. Biometrika, **63**, 465-474.

[13] Lehmann, E. L., and Casella, G. (1998). Theory of Point Estimation, 2nd edition. Springer-Verlag, New York.

[14] Tahir, M. (1988). Asymptotically optimal Bayesian sequential point estimation with censored data, Sequential Anal., **7**, 227–237.

[15] Magiera, R. (1977). On sequential minimax estimation for the exponential class of processes. Zastos. Mat., **15**, 445–454.