تابع زیان لاینکس در پیش‌بینی

فرزانه جامعی

در برخی مسائل آماری استفاده از توابع زیان نامتقارن مناسب است. مناسب‌ترین تابع زیان نامتقارن در اینگونه موارد ، تابع زیان لاینکس است. در مورد مسئله برآورد با توجه به این تابع زیان مقالاتی از شفیعی و زلنر ارائه شده است. در مسائل عملی نیاز به پیش‌بینی توابعی تصادفی از پارامتر دیده می‌شود. این مقاله در مورد پیش‌بینی با توجه به تابع زیان لاینکس است. برای پیدا کردن این پیش‌بینی کننده‌ها احتیاج به تعمیم نظریه $UMVUE$ کلاسیک داریم. این نظریه شامل قضیه رائو– بلکول و شرایط لازم وکافی برای پیدا کردن پیش‌بینی کننده‌ای که تابع زیان را به‌طور یکنواخت در پارامتر می‌نیمم کند، می‌باشد.

واژگان کلیدی: مسئله پیش‌بینی ، L– نااریبی ، تابع زیان لاینکس ، آماره مناسب ، قضیه رائو– بلکول

۱ . مقدمه

در خیلی از موارد کاربردی برآورد و پیش‌بینی استفاده از تابع زیان متقارن نامناسب است. در این گونه موارد بهترین تابع زیان نامتقارن تابع زیان لاینکس ($LINear EXponential$) است که در سال ۱۹۷۵ توسط واریان معرفی شد. این تابع زیان به صورت زیر می‌باشد:

$$L(d,\theta) = \exp[\alpha(d-\theta)] - \alpha(d-\theta) - ۱ \qquad -\infty \leq \alpha \leq +\infty$$

تعمیمی از تعریف نااریبی که با توجه به تابع زیان است و L– نااریبی گفته می‌شود

توسط لهمن (۱۹۵۴)، کلبانوف (۱۹۷۳) و نوربلوچی ومیدن (۱۹۸۳) ارائه شده است. از دید کلبانوف L− ناارببی به این صورت تعریف شده است که $\delta(X)$ یک برآورد L− ناریب برای $\gamma(\theta)$ است اگر:

$$E_\theta\Big(L(\delta(X),\gamma(\theta))\Big) \le E_\theta\Big(L(\delta(X),\gamma'(\theta))\Big) \qquad \text{برای تمام مقادیر } \theta$$

کلبانوف (۱۹۷۴) قضیه رائو − بلکول را به دسته‌ای از توابع زیان که تابع زیان لاینکس یکی از این دسته توابع زیان می‌باشد تعمیم داد. زلنر (۱۹۸۶) نشان داد که میانگین نمونه‌ای برآوردی غیر مجاز (*inadmisible*) تحت تابع زیان لاینکس است. شفیعی (۱۹۹۵) نتیجه زلنر را برای خانواده‌های مکانی تعمیم داد. پارسیان نیز از دید بیزی به این مسئله پرداخته است و برای پارامترهای توزیع نرمال برآوردهای لاینکس− ناریب را به دست آورده است. یاتراکوس (۱۹۹۲) با توجه به تابع زیان مربع اشتباهات برای توابعی تصادفی از پارامتر و با تعمیم نظریه رائو − بلکول بهترین پیش‌بینی کننده در بین توابع زیان مربع اشتباهات را به دست آورد. ایکس‌یائو (۱۹۹۹) بهترین پیش‌بینی کننده لاینکس− ناریب را به دست آورد. در این مقاله ما برای توابعی تصادفی از پارامتر و با توجه به تابع زیان لاینکس بهترین پیش‌بینی کننده را به دست می آوریم.در بخش دوم این مقاله در مورد پیش‌بینی کننده‌های ریسک− ناریب بحث شده است. در بخش سوم بهترین پیش‌بینی کننده لاینکس− ناریب را به دست می‌آوریم.

۲ . پیش‌بینی کننده ریسک ناریب

فرض کنید X و Y بردارهای تصادفی باشند که توزیع آنها به پارامتر مجهول $\theta \in \Theta(\subseteq R^k, k = ۱, ۲, \ldots)$ بستگی دارند. می‌خواهیم $h(y,\theta)$ را توسط x پیش‌بینی کنیم . تابع زیان نامنفی $L(d, h(y,\theta))$ زیان ناشی از پیش‌بینی $h(y,\theta)$ توسط d را نشان می‌دهد. فرض کنید$\delta(x)$ پیش‌بینی کننده $h(y,\theta)$ باشدآنگاه :

$$R(\theta, d) = E_\theta\left\{L(\delta(X), h(Y,\theta))\right\}$$

تابع ریسک است.

تعریف . اگر پیش‌بینی کننده $\delta(x)$ در رابطه زیر صدق کند

$$E_\theta \left\{ L(\delta(X), h(Y,\theta)) \right\} = \min_c E_\theta \left\{ L(\delta(X), h(Y,\theta) + c) \right\} \qquad (۱)$$

که c یک عدد حقیقی است، دراین صورت $\delta(x)$ پیش‌بینی کننده ریسک ناااریب گفته می‌شود.اگر پیش‌بینی کننده ریسک نااریب، ریسک را برای تمام θ ها می‌نیمم کند، آن بهترین پیش‌بینی کننده ریسک نااریب است .

قضیه ۱ . فرض کنید $L(d, h(y,\theta))$ تابع زیانی باشد که دارای مشتق مرتبه دوم است وبرای هر مقدار ثابت آرگومان اول درآرگومان دومش محدب است. فرض کنید یک پیش‌بینی کننده $\delta(x)$ وجود دارد به طوری که

$$E_\theta \left\{ \sup_{|c|<\epsilon} \frac{\partial^۲ L}{\partial y^۲} \left(\delta(X), h(Y,\theta) + c\right) \right\} < \infty \qquad \text{برای تمام مقادیر } \theta \quad (۲)$$

که ϵ مقدار ثابت است. شرط لازم و کافی برای اینکه $\delta(x)$ ریسک نااریب باشد این است که

$$E_\theta \left\{ \frac{\partial L}{\partial y} \left(\delta(X), h(Y,\theta)\right) \right\} = ۰ \qquad (۳)$$

اثبات. اگر $\delta(x)$ یک پیش‌بینی کننده ریسک نااریب برای $h(y,\theta)$ باشد آنگاه با استفاده از بسط تیلور و فرم لاگرانژی باقیمانده وبرای $|c| < \epsilon$ داریم

$$\begin{aligned} &E_\theta L\left(\delta(X), h(Y,\theta) + c\right) - E_\theta L\left(\delta(X), h(Y,\theta)\right) = \\ &cE_\theta \left\{ \frac{\partial L}{\partial y} \left(\delta(X), h(Y,\theta)\right) \right\} + \frac{۱}{۲} c^۲ E_\theta \left\{ \frac{\partial^۲ L}{\partial y^۲} \left(\delta(X), h(Y,\theta) + \nu(X, h(Y,\theta))c\right) \right\} \end{aligned}$$

که $\circ \le \nu(X, h(Y,\theta)) \le ۱$. اگر $\delta(X)$ یک پیش‌بینی کننده ریسک ناااریب برای $h(Y,\theta)$ باشد طرف چپ معادله فوق نامنفی است. از طرف دیگر سمت راست معادله فوق به عنوان تابع درجه دومی از c تنها هنگامی مثبت است که:

$$E_\theta \left\{ \frac{\partial L}{\partial y}\Big(\delta(X), h(Y,\theta)\Big) \right\} = \circ$$

برای اثبات عکس قضیه فرض کنید $E_\theta \left\{ \frac{\partial L}{\partial y}(\delta(X), h(Y,\theta)) \right\} = \circ$ آنگاه با استفاده از خاصیت محدب بودن تابع زیان L در دومین آرگومان داریم:

$$\begin{aligned} &E_\theta \left\{ L(\delta(X), h(Y,\theta) + c) \right\} - E_\theta \left\{ L(\delta(X), h(Y,\theta) \right\} = \\ &\qquad \frac{۱}{۲} c^۲ \left\{ \frac{\partial^۲ L}{\partial y^۲}\Big(\delta(x), h(y,\theta) + \nu(x, h(x, h(y,\theta) + c\Big) \right\} \ge \circ \end{aligned}$$

بنابراین $\delta(X)$ یک پیش‌بینی کننده ریسک– نااریب برای $h(Y,\theta)$ است. $\square$

نتیجه ۱. اگر تابع زیان ما یک تابع زیان مربع اشتباهات باشد آنگاه $\delta(x)$ ریسک نااریب است اگر و فقط اگر

$$E_\theta\left(\delta(X)\right) = E_\theta\left(h(Y,\theta)\right) \qquad (۴)$$

نتیجه ۲. اگر $\delta(x)$ در رابطه زیر صدق کند

$$E_\theta\left\{\exp\left[\alpha(\delta(X) - h(Y,\theta)\right]\right\} \le \infty \qquad \text{برای تمام مقادیر } \theta \qquad (۵)$$

آنگاه $\delta(x)$ لاینکس نااریب است اگر و فقط اگر

$$E\{\exp[\alpha(\delta(X) - h(Y,\theta)]\} = ۱ \qquad (٦)$$

قضیه ۲ . اگر $\delta(x)$ لاینکس ناریب باشد و در ۵ صدق کند، نمی‌تواند میانگین ناریب باشد مگر آنکه $P(\delta(x) = h(y,\theta)) = ۱$ باشد. برعکس این رابطه نیز درست است.

اثبات. فرض کنید $\delta(x)$ میانگین ناریب نیز باشد پس توأماً در روابط ۴ و ٦ صدق می‌کند بنابراین داریم

$$R(\theta,\delta) = E_\theta \{\exp[\alpha(\delta(X) - h(Y,\theta)] - \alpha(\delta(X) - h(Y,\theta)) - ۱\} = ۰$$

واین نتیجه می‌دهد $P(\delta(x) = h(y,\theta)) = ۱$. بنابراین $\delta(x)$ نمی‌تواند میانگین ناریب باشد مگر آنکه داشته باشیم $P(\delta(x) = h(y,\theta)) = ۱$. □

۳ . پیش‌بینی کننده‌های لاینکس– ناریب

در این بخش پیش‌بینی کننده لاینکس– ناریب را به دست می‌آوریم. ابتدا آماره مناسب را تعریف می‌کنیم. این آماره مانند آماره کافی در مسئله برآورد نقش مهمی را در مسئله پیش‌بینی ایفا می‌کند.

تعریف . یک آماره $T = T(X)$، مناسب گفته می‌شود اگر X و Y به شرط T از هم مستقل باشند و T برای خانواده توزیع X کافی باشد.

قضیه ۳. اگر $f(x,y \mid \theta)$ تابع چگالی توأم X و Y، باشد آنگاه شرط لازم و کافی برای اینکه آماره $T(X)$، یک آماره مناسب باشد این است که توابع غیرمنفی h و g وجود داشته باشند به طوری که

$$f(x,y \mid \theta) = h(x)g(T(x),y,\theta) \tag{۷}$$

این قضیه دقیقاً شبیه قضیه تجزیه فیشر– نیمن اثبات می‌شود.

اگر $\delta(X)$ برآوردی میانگین– نااریب باشد و T آماره مناسب باشد آنگاه:

$$\delta^*(T) = E(\delta(X)|T)$$

نیز ریسک– نااریب است و ریسک δ^* کمتر یا مساوی ریسک δ است. ما در اینجا نتایج مشابهی را برای توابع زیان به شکل زیر به دست می‌آوریم.

$$L(d, h(y,\theta)) = \psi(d)A(h(y,\theta)) + B(h(y,\theta)) + \phi(d) \qquad (۸)$$

که $A(\cdot)$ و $B(\cdot)$ محدب و دارای مشتق مرتبه دوم هستند و (> ۰) $\psi(\cdot)$ یک تابع کاملاً یکنوا است و $\phi(\psi^{-۱}(\cdot))$ تابعی محدب است. تابع زیان لاینکس به این شکل است.

قضیه ۴. اگر δ یک ریسک نااریب با توجه به تابع زیان فوق باشد ودر رابطه ۲ صدق کند و T نیز یک آماره مناسب باشد آنگاه

$$\delta^*(T) = \psi^{-۱}(E[\psi(\delta(X) \mid T]) \qquad (۹)$$

نیز در رابطه ۲ صدق می‌کند و ریسک نااریب است. ریسک δ^* کوچکتر یا مساوی ریسک δ است.

اثبات. شرط ۲ هم ارز است با

$$E\left\{\psi(\delta(X)) \sup_{|c|\leq\epsilon} A''(h(Y,\theta)+c)\right\} < \infty$$

و

$$E\left\{\sup_{|c|\leq\epsilon} B''(h(Y,\theta)+c)\right\} < \infty$$

از آنجاییکه T آماره‌ای مناسب است

$$\begin{aligned} E_\theta\left\{\psi(\delta^*(T)) \sup_{|c|\leq\epsilon} A''(h(Y,\theta)+c)\right\} &= E_\theta\left\{E(\psi(\delta(X))|T) \sup_{|c|\leq\epsilon} A''(h(Y,\theta)+c)\right\} \\ &= E_\theta\left\{\psi(\delta(X)) \sup_{|c|\leq\epsilon} A''(h(Y,\theta)+c)\right\} \end{aligned}$$

بنابراین شرط ۲ برای δ^* صادق است. طبق قضیه ۱ باید نشان دهیم δ^* در رابطه ۳ صدق می‌کند. با توجه به نااریبی δ و رابطه ۳ داریم:

$$\begin{aligned}
E_\theta\left\{\frac{\partial}{\partial y}L(\delta^*(T),h(Y,\theta))\right\} &= E_\theta\{\psi(\delta^*(T))A'(h(Y,\theta))+B'(h(Y,\theta))\}\\
&= E_\theta\{E(\psi(\delta(X))|T)A'(h(Y,\theta))+B'(h(Y,\theta))\}\\
&= E_\theta\{E(\psi(\delta(X))A'(h(Y,\theta)|T)+B'(h(Y,\theta))\}\\
&= E_\theta\{\psi(\delta(X))A'(h(Y,\theta))+B'(h(Y,\theta))\}\\
&= E_\theta\left\{\frac{\partial}{\partial y}L(\delta(X),h(y,\theta))\right\}\\
&= 0
\end{aligned}$$

بنابراین δ^* ریسک— نااریب است.

در قسمت بعد باید نشان دهیم ریسک δ^* از ریسک δ کمتر یا با آن مساوی است. با استفاده از نامساوی جنسن داریم:

$$\begin{aligned}
\phi(\delta^*(T)) &= \phi(\psi^{-1}(E[\psi(\delta(X))|T]))\\
&\le E(\phi(\delta(X))|T)
\end{aligned}$$

پس:

$$E\phi(\delta^*(T)) \le E\phi(\delta(X))$$

بنابراین:

$$\begin{aligned}
R(\theta,\delta^*) &= E_\theta\{\psi(\delta^*(T))A(h(y,\theta))+B(h(y,\theta))+\phi(\delta^*(T))\}\\
&= E_\theta\{E(\psi(\delta(X))|T)A(h(y,\theta))+B(h(y,\theta))+\phi(\delta^*(T))\}\\
&= E_\theta\{\psi(\delta(X))A(h(y,\theta))+B(h(y,\theta))+\phi(\delta^*(T))\}\\
&\le E_\theta\{\psi(\delta(X))A(h(y,\theta))+B(h(y,\theta))+\phi(\delta(X))\}\\
&= R(\theta,\delta)
\end{aligned}$$

□

قضیه ۵. اگر T مناسب و کامل باشد و

$$E_\theta(A'(h(Y,\theta)) \mid T) = Q(\theta)K(T) \qquad (۱۰)$$

برای بعضی از توابع غیر صفر Q و K برقرار باشد آنگاه δ^* رابطه ۹، بهترین پیش‌بینی کننده ریسک ناااریب در بین پیش‌بینی کننده‌هایی که در رابطه ۲ صدق می‌کنند است.

اثبات. کافی است نشان دهیم پیش‌بینی کننده رابطه ۹ بهترین پیش‌بینی کننده بر اساس T است. فرض کنید $\delta(T)$ یک پیش‌بینی کننده ریسک نااریب دیگر بر اساس T باشد، از رابطه ۳ داریم

$$E_\theta\left\{\psi(\delta^*(T))A'(h(Y,\theta))\right\} = E_\theta\left\{\psi(\delta(T))A'(h(Y,\theta))\right\}$$

$$E_\theta\left\{\psi(\delta^*(T))E_\theta(A'(h(Y,\theta)) \mid T)\right\} = E_\theta\left\{\psi(\delta(T))E_\theta(A'(h(Y,\theta)) \mid T)\right\}$$

بنابراین از رابطه ۱۰ داریم

$$Q(\theta)E_\theta\left\{K(T)\psi(\delta^*(T))\right\} = Q(\theta)E_\theta\left\{K(T)\psi(\delta(T))\right\}$$

از کامل بودن T و یکنوایی اکید ψ داریم

$$P_\theta(\delta^*(T) = \delta(T)) = ۱$$

□

مثال ۱. فرض کنید $X_۱, X_۲, \ldots, X_n$ متغیرهای تصادفی مستقل با توزیع پواسن و پارامتر مجهول θ باشند. می‌خواهیم بر اساس $(m < n)$ $X = (X_۱, X_۲, \ldots, X_m)$ برای $h(Y,\theta) = Y + \theta$ که $Y = \sum_{i=۱}^{n} X_i$ است بهترین پیش‌بینی کننده لاینکس–نااریب را پیدا کنیم.

حل: می‌توان با استفاده از قضیه ۳ نشان داد $\sum_{i=۱}^{m} X_i$ آماره مناسب است و کامل نیز است. طبق قضیه ۵ رابطه ۱۰ نیز برای این تابع تصادفی از پارامتر برقرار است.

بنابراین بهترین پیش‌بینی کننده لاینکس– نااریب برای $h(Y,\theta) = Y + \theta$ عبارت است از:

$$c = m\bar{X}\left(\frac{۱}{\alpha}\ln[\frac{\alpha}{m} + (\frac{n}{m} - ۱)(۱ - e^{-\alpha}) + ۱] + ۱\right)$$

مثال ۲. فرض کنید $X_۱, X_۲, \ldots, X_n$ متغیرهای تصادفی مستقل با توزیع $N(۲\mu, \sigma^۲)$ و Y دارای توزیع $N(\mu, \sigma^۲)$ باشند. کواریانس بین Y هر کدام از X_i ها $(i = ۱, \ldots, n)$ برابر $\rho\sigma^۲$ است که ρ معلوم و $\frac{۱}{n} < \sigma^۲$. می‌خواهیم بر اساس $X_۱, X_۲, \ldots, X_n$ برای $h(Y,\theta) = Y + \theta$ بهترین پیش‌بینی کننده لاینکس– نااریب را به دست آوریم.

حل: با استفاده از قضیه ۳ می‌بینیم که $\bar{X}$ آماره مناسب است. $\bar{X}$ کامل نیز است. می‌توان به آسانی نشان داد که شرط قضیه ۵ برقرار است. بنابراین بهترین پیش‌بینی‌کننده لاینکس– نااریب عبارت است از

$$\delta(\bar{X}) = \bar{X} + \frac{۱}{۲}(۲\rho - \frac{n + ۱}{n})\alpha\sigma^۲$$

مراجع

[1] Yannis G.Yatracos (1992), "On Prediction and Mean Squared Error", The canadian Journal of Statistics, vol.20, no.2, pages 187-200.

[2] Yushan Xiao, "Linex Unbiasedness in a Prediction Problem", The Annals of Institute of Statistical Mathematics, underprint.