برآورد مینیماکس مجانبی-موضعی برای پارامترهای توزیع نرمال چند متغیره

مهدی شمس [۱] غلامرضا حسامیان [۲]

تاریخ دریافت: ۱۴۰۰/۱۱/۱۸
تاریخ پذیرش: ۱۴۰۱/۱۲/۱۵

چکیده:
نامساوی‌های اطلاع کاربرد فراوان در نظریه برآوردیابی و اتخاذ تصمیم‌های آماری دارند. در این مقاله کاربرد یک نامساوی اطلاع برای اتخاذ تصمیم مینیماکس در چارچوب نظریه بیز بیان می‌شود. بدین‌صورت که ابتدا یک نامساوی اساسی برای مخاطره بیز تحت تابع زیان مربع خطا معرفی می‌شود و سپس کاربردهایی از آن در تعیین برآوردگرهای مینیماکس مجانبی-موضعی در حالت یک متغیره و چند متغیره بیان می‌شود. در حالتی که مؤلفه‌های پارامتر متعامد باشند، برآوردگرهای مینیماکس مجانبی-موضعی، برای تابعی از بردار میانگین و ماتریس کواریانس در توزیع نرمال چند متغیره به دست می‌آید. در پایان کران‌های نامساوی اطلاع تحت یک تابع زیان عمومی محاسبه می‌شود.
واژه‌های کلیدی: برآوردگر مینیماکس، توزیع پیشین، تابع مخاطره، نامساوی اطلاع، پارامترهای متعامد.

۱ مقدمه

نامساوی‌های اطلاع در نظریه برآورد و نظریه تصمیم مورداستفاده قرار می‌گیرند. نامساوی اطلاع برای مخاطره بیز در پیشگوکننده‌ها [۱۰]، خانواده‌های نامنظم [۱]، پیشین‌های بریده‌شده [۲]، تحت زیان مقیاسی [۵] به کار گرفته‌شده و از طریق آن رویکردی برای یافتن برآوردگرهای مینیماکس پیدا می‌شود. در [۹] با استفاده از نامساوی کرامر-رائو یک کران پایین ناسره به دست آورده شد و از آن در دست‌یابی برآوردهای مینیماکس استفاده شد و با ارائه یک نامساوی اطلاع، دنباله‌ای از برآوردگرهای مینیماکس نرخ شکست در مدل نمایی سانسور شده و مدل نرخ شکست متناسب یافت شد. در [۱۹] مشابه نامساوی کرامر-رائو با تعمیم روش‌های [۸] تحت تابع زیان مربع خطا، یک کران پایین برای مخاطره برآوردگر دلخواه پارامتر ارائه و در مورد شرایط حصول این کران بحث شد و به دنبال روش‌های ذکرشده در [۸] و [۱۹] و با تعمیم نتایج آن‌ها در حالت چگالی مختلط مشابه کران کرامر- رائو، کران پایین مخاطره بیز محاسبه و شرایط حصول این کران ارائه شد. در [۱۱] نامساوی اطلاع و نامساوی باتاچاریا برای ماتریس کوواریانس وزنی تعمیم داده شد. در [۶] نسخه‌هایی از نامساوی کرامر-رائو که در متغیرهای تصادفی لگ-مقعر مورداستفاده قرار می‌گیرد تعمیم داده شد.

در این مقاله کاربرد نامساوی اطلاع برای اتخاذ تصمیم مینیماکس در چارچوب نظریه بیز بیان می‌شود. پس از معرفی نامساوی اساسی برای مخاطره بیز تحت تابع زیان مربع خطا، کاربردهایی از آن در تعیین برآوردگرهای مینیماکس مجانبی و موضعی در حالت یک متغیره و چند متغیره بیان می‌شود. به‌عنوان کاربردی از این مسئله، در حالتی که مؤلفه‌های پارامتر متعامد باشند، برآوردگرهای مینیماکس مجانبی-موضعی برای تابعی از بردار میانگین و ماتریس کواریانس در توزیع نرمال چند متغیره به دست آورده می‌شود. در پایان کران‌های نامساوی اطلاع تحت یک تابع زیان عمومی محاسبه می‌شود ([۱۳، ۱۵، ۱۶]).

[۱] گروه آمار دانشگاه کاشان (نویسنده مسئول: mehdishams@kashanu.ac.ir)
[۲] گروه آمار دانشگاه پیام نور

ابتدا مفاهیمی از نظریه تصمیم مانند تصمیم بیز و تصمیم مینیماکس مرور می‌شوند. فرض کنید $\boldsymbol{X} = (X_1, ..., X_n)$ متغیرهای تصادفی مستقل و هم‌توزیع در فضای اندازه‌پذیر $(\mathcal{X}, S)$ با اندازه احتمال $P_{\boldsymbol{\theta}}$ باشند که در آن $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^m$ و $m > 1$. همچنین خانواده توزیع $\{P_{\boldsymbol{\theta}}, \boldsymbol{\theta} \in \Theta\}$ نسبت به‌اندازه σ-متناهی μ مطلقاً پیوسته باشند و $f(., \boldsymbol{\theta}) = dP_{\boldsymbol{\theta}}/d\mu$ و به‌علاوه $q(\boldsymbol{\theta})$ یك چگالی احتمال پیشین برای $\boldsymbol{\theta}$ با تکیه‌گاه $S_q \subset \Theta$ باشد. در این مقاله امید ریاضی روی ناحیه $\mathcal{X}^n \times \Theta$ نسبت به چگالی $f_n(\boldsymbol{x}, \boldsymbol{\theta})\, q(\boldsymbol{\theta})$ را با E و امید ریاضی تحت چگالی $q(\boldsymbol{\theta})$ را با $E_{\boldsymbol{\theta}}(.)$ و امید ریاضی نسبت به چگالی $f_n(\boldsymbol{x}, \boldsymbol{\theta})$ را با $E_{\boldsymbol{X}|\boldsymbol{\theta}}$ نشان داده می‌شود. قاعده تصمیم (غیر تصادفی) $\boldsymbol{\theta}^*$ تابعی اندازه‌پذیر از فضای نمونه‌ای $\mathcal{X}$ به فضای کارهای $\mathcal{A}$ است، به‌طوری‌که به یک مشاهده نمونه تصادفی عملی را نسبت می‌دهد. فضای کارها معین می‌کند که مسئله تصمیم موردمطالعه یک مسئله برآورد نقطه‌ای و یا فاصله اطمینان و یا آزمون فرضیه است. برای مثال، در مسئله برآورد نقطه‌ای کارها حدس نقاط و در مسئله آزمون فرضیه دو کار پذیرش فرضیه صفر یا رد آن هستند. هم‌چنین در مسئله برآورد نقطه‌ای تابع تصمیم همان برآوردگر و در مسئله آزمون فرضیه تابع تصمیم همان تابع آزمون است. مجموعه کلیه توابع تصمیم ممکن فضای تصمیم نامیده می‌شود. در نظریه تصمیم هدف انتخاب بهینه یک تابع تصمیم است. تابع نامنفی $L(\boldsymbol{\theta}, a)$ از فضای حاصل‌ضرب دکارتی فضای پارامتر در فضای کارها به اعداد حقیقی مثبت را تابع زیان می‌نامند که نشان‌دهنده زیان ناشی از اخذ تصمیم a است هنگامی‌که در حالت واقعی $\boldsymbol{\theta}$ باشد [۱۲]. انتخاب تابع زیان بستگی به مسئله 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 تصمیم $\boldsymbol{\theta}^*$ می‌نامند و به‌صورت $R(\boldsymbol{\theta}, \boldsymbol{\theta}^*) = E_{\mathcal{X}^n|\boldsymbol{\theta}}(L(\boldsymbol{\theta}, \boldsymbol{\theta}^*(\boldsymbol{X})))$ تعریف می‌شود [۱۲]. در روش‌های کلاسیک پارامتر $\boldsymbol{\theta}$ یک مقدار ثابت و نامعلوم در نظر گرفته می‌شود و یک نمونه تصادفی $\boldsymbol{X}$ از جمعیتی که دارای توزیع $f_n(\boldsymbol{x}, \boldsymbol{\theta})$ است، جمع‌آوری می‌شود تا بر اساس آن در مورد پارامتر تصمیم‌گیری شود. در روش بیزی $\boldsymbol{\theta}$ کمیتی در نظر گرفته می‌شود که خود یک متغیر تصادفی است و تغییرات آن توسط یک توزیع پیشین $q(\boldsymbol{\theta})$ 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 $\boldsymbol{\theta}$ (توزیع پیشین) تصحیح می‌گردد. توزیع پیشین تصحیح‌شده را توزیع پسین $g_{\boldsymbol{X}|\boldsymbol{\theta}}(\boldsymbol{\theta})$ 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 $q(\theta)$ است. مخاطره بیز تابع تصمیم $\boldsymbol{\theta}^*$ با توجه به توزیع پیشین $q(\boldsymbol{\theta})$ به‌صورت امید ریاضی $R(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$ نسبت به $q(\boldsymbol{\theta})$ تعریف می‌شود و آن را با $BR(\boldsymbol{\theta}^*, \boldsymbol{\theta})$ نشان می‌دهیم. اصل مخاطره بیز به این معنی است که اطلاعات تابع مخاطره در یک عدد (مخاطره بیز) خلاصه می‌شود و دو تصمیم به‌وسیله مقایسه این مخاطره بیز با یکدیگر مقایسه می‌شوند. در واقع یک تصمیم بیز، مرتبط با توزیع پیشین $q(\boldsymbol{\theta})$ و تابع زیان L، تصمیمی است که مخاطره بیز را مینیمم کند.

رویکرد دیگر استفاده از تصمیم مینیماکس است. یک قاعده تصمیم $\boldsymbol{\theta}^*$ مینیماکس است هرگاه تابع مخاطره را مینیمم کند، به‌عبارت‌دیگر $R(\boldsymbol{\theta}, \boldsymbol{\theta}^*) = inf_{\delta} sup_{\boldsymbol{\theta}^* \in \Theta} R(\boldsymbol{\theta}, \delta)$. در حقیقت تصمیم مینیماکس آن تصمیمی است که در بدترین حالت تابع مخاطره $R(\boldsymbol{\theta}, \delta)$ نسبت به δ، بهترین حالت را انتخاب می‌کند و یک رویکرد محافظه‌کارانه برای انتخاب بهترین تصمیم است. برای کسب اطلاعات بیشتر در مورد تصمیم‌های بیز و مینیماکس به [۱۲] مراجعه کنید.

۲ یافتن برآورد مینیماکس با استفاده از نامساوی مخاطره بیزی

در [۱۸] بر اساس یک نمونه تصادفی، یک کران پایین برای مخاطره برآوردگر $g(\boldsymbol{\theta}^*)$ ($\boldsymbol{\theta}^*$ یک برآوردگر دلخواه $\boldsymbol{\theta}$ است) در تخمین پارامتر $g(\boldsymbol{\theta})$، یعنی $R(\boldsymbol{\theta}^*, \boldsymbol{\theta}) = E(g(\boldsymbol{\theta}^*) - g(\boldsymbol{\theta}))^2$ ارائه شد. در این بخش به بررسی نامساوی کرامر-رائو در حالتی که مؤلفه‌های پارامتر چندگانه است، می‌پردازیم و کران پایین برآورد مینیماکس مجانبی-موضعی در تخمین $g(\boldsymbol{\theta})$ توسط برآوردگر دلخواه $g(\boldsymbol{\theta}^*)$، بر اساس یک نمونه تصادفی به دست می‌آوریم و نشان می‌دهیم وقتی‌که مؤلفه‌های $\boldsymbol{\theta}$ متعامد باشند، کران‌های پایین ذکرشده قوی‌تر هستند. در ابتدا چند مفهوم از جبر خطی بیان می‌شود.

اگر $A_{m \times n} = (a_{ij})$ یك ماتریس از مرتبه $m \times n$ باشد و $A_{.j}$ها j-امین، $j = 1, ..., n$ ستون ماتریس A باشند، آنگاه $Vec(A)$ یك بردار از مرتبه $mn \times 1$ است و $Vec(A) = (A_{.1}, ..., A_{.n})^T$. برای دو ماتریس دلخواه $A_{p \times p}$ و $B_{p \times p}$ داریم $tr(AB) = (Vec(A))^T (Vec(B))$. اگر f تابعی از ماتریس $X = (x_{ij})_{p \times n}$ باشد، مشتق f نسبت

به X به‌صورت $\frac{d}{dX}f(X)=(\frac{\partial f(X)}{\partial x_{ij}})$ که در آن $i=1,2,...,p$ و $j=1,2,...,n$ تعریف می‌شود. به‌عنوان‌مثال اگر $\boldsymbol{x}$ یك بردار از مرتبه p باشد، داریم $\frac{d}{d\boldsymbol{x}}\boldsymbol{x}^T\boldsymbol{x}=2\boldsymbol{x}$ و $\frac{d}{d\boldsymbol{x}}\boldsymbol{a}^T\boldsymbol{x}=\boldsymbol{a}$ که در آن $\boldsymbol{a}_{p\times 1}$ یک بردار دلخواه است. اگر f و g توابعی دلخواه از $X=(x_{ij})_{p\times n}$ باشند، آنگاه $\frac{d}{dX}(f(X)g(X))=f(X)\frac{dg(X)}{dX}+\frac{df(X)}{dX}g(X)$. اگر $f(X)$ تابعی از یك ماتریس نامنفرد باشد ($|f(X)|\neq 0$)، آنگاه:

$$\frac{\partial}{\partial X}|f(X)|=-f(X)\frac{\partial}{\partial X}(f(X))^{-1}f(X).$$

اگر $A_{p\times p}$ یك ماتریس متقارن و $\boldsymbol{x}$ یك بردار دلخواه مخالف صفر از مرتبه p باشند، داریم $\frac{\partial}{\partial \boldsymbol{x}}\boldsymbol{x}^TA\boldsymbol{x}=2A\boldsymbol{x}$. اگر $\boldsymbol{x}$ و $\boldsymbol{y}$ دو بردار دلخواه به ترتیب از مرتبه n و p و $C_{n\times p}$ یك ماتریس باشد، آنگاه $\frac{\partial}{\partial C}\boldsymbol{x}^TC\boldsymbol{y}=\boldsymbol{x}\boldsymbol{y}^T$. اگر $X_{p\times p}$ یك ماتریس نامنفرد باشد، آنگاه $\frac{d}{dX}|X|=|X|\left(X^{-1}\right)^T$. اگر f تابعی از بردارهای p بعدی $\boldsymbol{x}$ و $\boldsymbol{y}$ باشد، مشتق جزئی f نسبت به $\boldsymbol{x}$ به‌صورت $\frac{\partial}{\partial\boldsymbol{x}}f=(Vec(\frac{\partial}{\partial\boldsymbol{y}}f))^T(Vec(\frac{\partial\boldsymbol{y}}{\partial\boldsymbol{x}}))=tr((\frac{\partial}{\partial\boldsymbol{y}}f)^T(\frac{\partial\boldsymbol{y}}{\partial\boldsymbol{x}}))$ تعریف می‌شود که قاعده مشتق زنجیره‌ای نامیده می‌شود.

تجزیه طیفی ماتریس متقارن $A_{n\times n}$ نشان می‌دهد، ماتریس متعامد R وجود دارد به‌گونه‌ای که $RR^T=R^TR=I_p$ و $A=R\Lambda R^T$ که در آن $\Lambda=diag(\lambda_1,...,\lambda_n)$ و I_p ماتریس همانی از مرتبه p است و λ_iها، $i=1,...,n$ مقادیر ویژه ماتریس A هستند. ماتریس متقارن $A_{p\times p}$ نیمه معین مثبت ($A\geq 0$) نامیده می‌شود، اگر برای هر بردار $\boldsymbol{x}\in\mathbb{R}^p$، $\boldsymbol{x}\neq 0$، $Q=\boldsymbol{x}^TA\boldsymbol{x}$ نامنفی باشد و ماتریس متقارن $A_{p\times p}$ معین مثبت ($A>0$) نامیده می‌شود، هرگاه $Q>0$. همچنین:

۱) اگر $A>0$، آنگاه $A^{-1}>0$.

۲) $A>0$ اگر و فقط اگر ماتریس نامنفرد M وجود داشته باشد به‌طوری‌که $A=M^TM$.

۳) اگر λ_iها، $1\leq i\leq p$ مقادیر ویژه ماتریس معین مثبت A باشند، آنگاه $tr(A)=\sum_{i=1}^n\lambda_i$ و $\lambda_i>0$ که در آن $tr(A)$ اثر ماتریس A است.

قضیه زیر کران‌های یک فرم درجه دوم را تعیین می‌کند.

لم ۱.۲. *اگر λ_iها ($1\leq i\leq p$) مقادیر ویژه ماتریس متقارن A باشند، آنگاه:* $\min_{1\leq i\leq p}\lambda_i\leq\frac{\boldsymbol{x}^TA\boldsymbol{x}}{\boldsymbol{x}^T\boldsymbol{x}}\leq\max_{1\leq i\leq p}\lambda_i$.

لم ۲.۲. *([۱۴]) اگر $\boldsymbol{Z}=(Z_1,...,Z_n)$ یک بردار تصادفی با ماتریس واریانس-کواریانس معین مثبت Σ باشد که در آن برای هر $1\leq i\leq m$، $E(Z_i^2)<\infty$ و Y متغیر تصادفی دیگری باشد که $E(Y^2)<\infty$، با تعریف $\gamma_i=\text{cov}(Z_i,Y)$ که $i=1,...,m$ داریم:*

$$\text{var}(Y)\geq\gamma^T\Sigma^{-1}\gamma \qquad (1)$$

که در آن $\gamma^T=(\gamma_1,...,\gamma_n)$. به‌ویژه اگر λ_1 ماکزیمم مقادیر ویژه ماتریس Σ باشد، آنگاه $\text{var}(Y)\geq\frac{1}{\lambda_1}\sum_{i=1}^m{\gamma_i}^2\geq\frac{\sum_{i=1}^m\gamma_i^2}{tr(\Sigma)}$ که در آن $tr(\Sigma)$ مجموع درایه‌های قطری Σ است. بعلاوه اگر Σ یک ماتریس قطری با درایه‌های ${\sigma_i}^2$ باشد، آنگاه:

$$\text{var}(Y)\geq\sum_{i=1}^m\frac{{\gamma_i}^2}{{\sigma_i}^2}. \qquad (2)$$

برای اثبات نامساوی (۲) به ازای هر بردار دلخواه $\boldsymbol{x}\neq 0$ با استفاده از نامساوی کوشی- شوارتز داریم:

$$\left(\text{cov}\left(\boldsymbol{x}^T\boldsymbol{Z},Y\right)\right)^2\leq\text{var}\left(\boldsymbol{x}^T\boldsymbol{Z}\right).\text{var}(Y).$$

بنابراین طبق تعریف γ و لم ۱.۲ داریم:

$$\text{var}(Y)\geq\sup_{\boldsymbol{x}\neq 0}\frac{(\boldsymbol{x}^T\gamma)^2}{\boldsymbol{x}^T\Sigma\boldsymbol{x}}=\gamma^T\Sigma^{-1}\gamma$$

و با قرار دادن $\Sigma^{-1}=diag\{{\sigma_i}^{-2}\}$ در این نامساوی که در آن Σ یك ماتریس قطری با درایه‌های σ_i^2 است، نامساوی (۲) نتیجه می‌شود. از طرف دیگر داریم $tr(\Sigma)=\sum_{i=1}^m\text{var}(Z_i)$ و

$$\gamma^T\gamma=\sum_{i=1}^m{\gamma_i}^2=\rho^2\,\text{var}(Y)\sum_{i=1}^m\text{var}(Z_i)$$

که در آن ρ ضریب همبستگی بین Y و Z_i است. به دلیل اینکه $\rho^2\leq 1$ داریم $\text{var}(Y)\geq\frac{\sum_{i=1}^m\gamma_i^2}{tr(\Sigma)}$. حال با قرار دادن $A=\Sigma^{-1}$ در لم ۱.۲ داریم:

$$\inf_{\gamma\neq 0}\frac{\gamma^T\Sigma^{-1}\gamma}{\gamma^T\gamma}=\frac{1}{\lambda_1}. \qquad (3)$$

از رابطه (۳) نتیجه می‌شود:

$$\gamma^T\Sigma^{-1}\gamma\geq\frac{\sum_{i=1}^m\gamma_i^2}{\lambda_1}. \qquad (4)$$

اما به دلیل اینکه Σ^{-1} یک ماتریس معین مثبت است، λ_iها همگی مثبت هستند و ازاین‌رو $\lambda_1<tr(\Sigma)$. با توجه به این نکته اگر ماتریس قطری باشد، داریم $\sum_{i=1}^m\frac{\gamma_i^2}{\sigma_i^2}\geq\frac{\sum_{i=1}^m\gamma_i^2}{tr(\Sigma)}$. اگر Σ یك ماتریس قطری با درایه‌های σ_i^2 باشد، با قرار دادن $\Sigma^{-1}=diag\{{\sigma_i}^{-2}\}$ در (۳)، نامساوی (۲) نتیجه می‌شود. با مقایسه روابط (۲) و (۴) ملاحظه می‌شود که نامساوی (۴) قوی‌تر از نامساوی (۲) است. در [۱۴، ۱۵] تحت شرایط مطلوب زیر یک نامساوی اساسی برای مخاطره بیز پیدا می‌شود:

الف) توابع $k_i(\boldsymbol{x},\boldsymbol{\theta})$، $1 \le i \le m$ توابعی توأماً اندازه‌پذیر باشند که نسبت به‌اندازه حاصل‌ضرب $\lambda \times \mu^n$ روی $\Theta \times \mathcal{X}^n$ مطلقاً انتگرال‌پذیر هستند و λ اندازه لبگ در $\mathbb{R}^m$ است و برای هر $\boldsymbol{x} \in \mathcal{X}^m$، $\int_\Theta k_i(\boldsymbol{x},\boldsymbol{\theta})d\boldsymbol{\theta}=0$.

ب) فرض کنید $g(\boldsymbol{\theta})k_i(\boldsymbol{x},\boldsymbol{\theta})$ در $(\boldsymbol{x},\boldsymbol{\theta})$ توأماً اندازه‌پذیر باشد و نسبت به‌اندازه $\lambda\times\mu^n$ روی $\Theta\times\mathcal{X}^n$ به‌طور پیوسته انتگرال‌پذیر باشد.

ج) $g(.)$ نسبت به مؤلفه‌های $\boldsymbol{\theta}$ مشتق‌پذیر و $\frac{\partial}{\partial\boldsymbol{\theta}_i}g(\boldsymbol{\theta})$ در مؤلفه‌های $\boldsymbol{\theta}$ پیوسته باشد.

د) $h(.)$ تابع اندازه‌پذیر دلخواه و $f_n(\boldsymbol{x},\boldsymbol{\theta})h(\boldsymbol{\theta})$ نسبت به مؤلفه‌های $\boldsymbol{\theta}$ مشتق‌پذیر باشد. برای هر $\boldsymbol{x}\in\mathcal{X}^n$ و $\boldsymbol{\theta}\in S_q$ تعریف می‌شود $k_i(\boldsymbol{x},\boldsymbol{\theta})=\frac{\partial}{\partial\boldsymbol{\theta}_i}\left[f_n(\boldsymbol{x},\boldsymbol{\theta})h(\boldsymbol{\theta})\right]$.

به‌علاوه برای هر $\boldsymbol{x}\in\mathcal{X}^n$ فرض می‌شود $|f_n(\boldsymbol{x},\boldsymbol{\theta})h(\boldsymbol{\theta})|\to 0$ وقتی $|\boldsymbol{\theta}_i|\to\partial\Theta_i$ که در آن Θ_i تکیه‌گاه iامین مؤلفه مؤلفه‌های $\boldsymbol{\theta}$ و $\partial\Theta_i$ نقاط مرزی Θ_i است.

ه) جابه‌جایی مشتق نسبت به مؤلفه‌های $\boldsymbol{\theta}$ در $\int_{\mathcal{X}^n} f_n(\boldsymbol{x},\boldsymbol{\theta})\mu^n(d\boldsymbol{x})=1$ مجاز باشد.

و) $I_n(\boldsymbol{\theta})=\sum_{i=1}^m E\left[\frac{\partial}{\partial\boldsymbol{\theta}_i}\log f_n(\boldsymbol{X},\boldsymbol{\theta})\right]^2$ و $\frac{\partial}{\partial\boldsymbol{\theta}_i}I(\boldsymbol{\theta})$ موجود و در مؤلفه‌های $\boldsymbol{\theta}_i$ پیوسته باشد.

با تعریف $Y=g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})$، $Z_i=G_i(\boldsymbol{x},\boldsymbol{\theta})=\frac{k_i(\boldsymbol{x},\boldsymbol{\theta})}{f_n(\boldsymbol{x},\boldsymbol{\theta})q(\boldsymbol{\theta})}$ و با استفاده از قضیه فوبینی و فرض (الف) به‌سادگی می‌توان نشان داد $E(Z_i)=0$ و $\gamma_i=E(YZ_i)=-E\left[g(\boldsymbol{\theta})G_i(\boldsymbol{X},\boldsymbol{\theta})\right]$ و همچنین $\mathrm{cov}(Z_i,Z_j)=E[G_i(\boldsymbol{X},\boldsymbol{\theta})\,G_j(\boldsymbol{X},\boldsymbol{\theta})]$.

رابطه (۱) را در نظر بگیرید که در آن

$$\gamma_{m\times1}=\left(-E\left[g(\boldsymbol{\theta})G_1(\boldsymbol{X},\boldsymbol{\theta})\right],...,-E\left[g(\boldsymbol{\theta})G_m(\boldsymbol{X},\boldsymbol{\theta})\right]\right)^T$$

و $\Sigma=\left((E[G_i(\boldsymbol{X},\boldsymbol{\theta})G_j(\boldsymbol{X},\boldsymbol{\theta})])\right)_{m\times m}$. با توجه به مطالب بیان‌شده در بالا قضیه زیر نتیجه گرفته می‌شود:

قضیهٔ ۳.۲. *([۱۴]) تحت شرایط (الف)-(ج) داریم:*

$$E[g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})]^2 \ge \frac{\sum_{i=1}^m E\left[g(\boldsymbol{\theta})G_i(\boldsymbol{x},\boldsymbol{\theta})\right]}{\sum_{i=1}^m E[G_i(\boldsymbol{x},\boldsymbol{\theta})]^2} + \left[E\left(g(\boldsymbol{\theta}^*)\right)-g(\boldsymbol{\theta})\right]^2.$$

فرض کنید شرایط (ه) و (و) نیز برقرار باشند، در این صورت $G_i(\boldsymbol{x},\boldsymbol{\theta})$ را می‌توان به‌صورت $G_i(\boldsymbol{x},\boldsymbol{\theta})=\frac{\partial}{\partial\boldsymbol{\theta}_i}\log f_n(\boldsymbol{x},\boldsymbol{\theta})\frac{h(\boldsymbol{\theta})}{q(\boldsymbol{\theta})}+\frac{h^{(i)}(\boldsymbol{\theta})}{q(\boldsymbol{\theta})}$ نوشت که در آن $h^{(i)}(\boldsymbol{\theta})$ مشتق جزئی $h(\boldsymbol{\theta})$ نسبت به $\boldsymbol{\theta}_i$ است. با اضافه و کم کردن $\frac{\partial}{\partial\boldsymbol{\theta}_i}g(\boldsymbol{\theta})[f_n(\boldsymbol{x},\boldsymbol{\theta})h(\boldsymbol{\theta})]$ به عبارت $[g(\boldsymbol{\theta})G_i(\boldsymbol{x},\boldsymbol{\theta})]$، با توجه به شرط (د) و قضیه فوبینی رابطه $E\left[g(\boldsymbol{\theta})G_i(\boldsymbol{x},\boldsymbol{\theta})\right]=-E\left[g^{(i)}(\boldsymbol{\theta})\frac{h(\boldsymbol{\theta})}{q(\boldsymbol{\theta})}\right]$ نتیجه می‌شود که در آن $g^{(i)}(\boldsymbol{\theta})$ مشتق $g(\boldsymbol{\theta})$ نسبت به $\boldsymbol{\theta}_i$ است. همچنین از شرط (ح) داریم $E_\theta\left[\frac{\partial}{\partial\boldsymbol{\theta}_i}\log f_n(\boldsymbol{X},\boldsymbol{\theta})\right]=0$. بنابراین:

$$E[G_i^2(\boldsymbol{x},\boldsymbol{\theta})]=E_\theta\left[I_n^{(i)}(\boldsymbol{\theta})\frac{h^2(\boldsymbol{\theta})}{q^2(\boldsymbol{\theta})}\right]+E\left[\frac{h^{(i)}(\boldsymbol{\theta})}{q(\boldsymbol{\theta})}\right]^2$$

که در آن $I_n^{(i)}(\boldsymbol{\theta})=E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\frac{\partial}{\partial\boldsymbol{\theta}_i}\log {f_n}^{(i)}(\boldsymbol{X},\boldsymbol{\theta})\right]^2$. بنابراین قضیه زیر ثابت می‌شود.

قضیهٔ ۴.۲. *([۱۴]) تحت برقراری شرایط (الف)-(ه) داریم:*

$$\begin{aligned}&E[g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})]^2 \ge \\ &\frac{\sum_{i=1}^m\left(E\left[g^{(i)}(\boldsymbol{\theta})\frac{h(\boldsymbol{\theta})}{q(\boldsymbol{\theta})}\right]\right)^2}{\sum_{i=1}^m E\left[I_n^{(i)}(\boldsymbol{\theta})\frac{h^2(\boldsymbol{\theta})}{q^2(\boldsymbol{\theta})}\right]+\sum_{i=1}^m E\left[\frac{h^{(i)}(\boldsymbol{\theta})}{q(\boldsymbol{\theta})}\right]^2}\\ &+\left[E\left(g(\theta^*)-g(\boldsymbol{\theta})\right)\right]^2\end{aligned}$$

به‌ویژه:

$$\begin{aligned}&E[g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})]^2 \ge \\ &\frac{\sum_{i=1}^m\left(E\left[g^{(i)}(\boldsymbol{\theta})\frac{h(\boldsymbol{\theta})}{q(\boldsymbol{\theta})}\right]\right)^2}{\sum_{i=1}^m E\left[I_n^{(i)}(\boldsymbol{\theta})\frac{h^2(\boldsymbol{\theta})}{q^2(\boldsymbol{\theta})}\right]+\sum_{i=1}^m E\left[\frac{h^{(i)}(\boldsymbol{\theta})}{q(\boldsymbol{\theta})}\right]^2}.\end{aligned} \tag{5}$$

با فرض برقراری شرایط مطلوب (ب) تا (ه)، در دو حالت خاص با انتخاب $h(\theta)=q(\theta)$ یا $h(\theta)=\frac{q(\theta)}{{I_n}^{(i)}(\theta)}$ در رابطه (۵) به ترتیب دو نتیجه زیر حاصل می‌شود.

نتیجه ۵.۲. *([۱۴]) تحت برقراری شرایط (الف)-(ه):*

$$\begin{aligned}E[g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})]^2 &\ge \frac{\sum_{i=1}^m\left(E\left[\frac{g^{(i)}(\boldsymbol{\theta})}{I_n(\boldsymbol{\theta})}\right]\right)^2}{E\left[\frac{1}{I_n(\boldsymbol{\theta})}\right]+\sum_{i=1}^m E\left[\frac{h^{(i)}(\boldsymbol{\theta})}{q(\boldsymbol{\theta})}\right]^2}\\ &=\frac{\frac{1}{n}\sum_{i=1}^m\left(E\left[\frac{g^{(i)}(\boldsymbol{\theta})}{I_1(\boldsymbol{\theta})}\right]\right)^2}{\frac{1}{n}E\left[\frac{1}{I_1(\boldsymbol{\theta})}\right]+\sum_{i=1}^m E\left[\frac{h^{(i)}(\boldsymbol{\theta})}{q(\boldsymbol{\theta})}\right]^2}\end{aligned}$$

که تساوی آخر بر اساس این حقیقت نوشته شده که:

$$\begin{aligned}I_n(\boldsymbol{\theta})&=\sum\nolimits_{i=1}^m I_n^{(i)}(\boldsymbol{\theta})\\ &=\sum\nolimits_{i=1}^m E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\frac{\partial}{\partial\boldsymbol{\theta}_i}\log f_n(\boldsymbol{x},\boldsymbol{\theta})\right]^2\\ &=nI_1(\boldsymbol{\theta}).\end{aligned}$$

نتیجه ۶.۲. *([۱۴]) تحت برقراری شرایط (الف)-(ه):*

$$\begin{aligned}&E[g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})]^2 \ge \\ &\frac{\sum_{i=1}^m\left(E\left[g^{(i)}(\boldsymbol{\theta})\right]\right)^2}{nE[I_1(\boldsymbol{\theta})]+\sum_{i=1}^m E\left[\frac{\partial}{\partial\boldsymbol{\theta}_i}\log q(\boldsymbol{\theta})\right]^2}.\end{aligned} \tag{6}$$

حال بر اساس رابطه (۶) نامساوی زیر که حالت حدی مخاطره مینیماکس است حاصل می‌شود:

قضیهٔ ۷.۲. *([۱۴]) تحت برقراری شرایط (الف)-(ه):*

$$\underline{\lim}_{n\to\infty} \inf_{\boldsymbol{\theta}^*}\sup_{\boldsymbol{\theta}\in J_n} E_\theta\left[\sqrt{n}\,(g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta}))\right]^2 \ge \frac{\sum_{i=1}^m \left(g^{(i)}(\boldsymbol{\theta}_o)\right)^2}{I_1(\boldsymbol{\theta}_o)}$$

که در آن و وقتی‌که $J_n=\{\boldsymbol{\theta}:\,|\boldsymbol{\theta}-\boldsymbol{\theta}_o|<\varepsilon_n\}$، $\varepsilon_n,\frac{1}{n\varepsilon_n^2}\to 0$.

اثبات: به ازای $\varepsilon_n>0$ ، $J_n\subset\Theta^o$ و $\boldsymbol{\theta}_o=(\theta_{1o},...,\theta_{mo})\in J_n$ که در آن Θ^o مجموعه نقاط درون Θ است. برای هر $\boldsymbol{\theta}\in J_n$:

$$E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})\right]^2 \le \sup_{\boldsymbol{\theta}\in J_n} E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})\right]^2 \qquad (۷)$$

که در آن $q(\boldsymbol{\theta})$ چگالی پیشین دلخواه در J_n است. اکنون با انتخاب تابع چگالی پیشین

$$q(\boldsymbol{\theta})=\prod_{i=1}^m\left\{\tfrac{1}{\varepsilon}\cos^2\tfrac{\pi(\theta_i-\theta_{i\circ})}{2\varepsilon}\right\} \qquad (۸)$$

در J_n، با محاسبات مستقیم، به‌سادگی ملاحظه می‌شود که $E\left[\frac{\partial}{\partial\theta_i}\log q(\boldsymbol{\theta})\right]^2=\pi^2\varepsilon_n{}^{-2}$، $1\le i\le m$. حال با جایگزین نمودن رابطه فوق در رابطه (۶) نامساوی زیر حاصل می‌شود:

$$E[g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})]^2 \ge \frac{\sum_{i=1}^m\left(E\left[g^{(i)}(\boldsymbol{\theta})\right]\right)^2}{n\,E[I(\boldsymbol{\theta})]+m\pi^2\varepsilon_n{}^{-2}}. \qquad (۹)$$

با تلفیق روابط (۷) و (۹) به‌راحتی مشاهده می‌شود:

$$\inf_{\boldsymbol{\theta}^*}\sup_{\boldsymbol{\theta}\in J_n} E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\sqrt{n}\,(g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta}))\right]^2 \ge \frac{\sum_{i=1}^m\left(E\left[g^{(i)}(\boldsymbol{\theta})\right]\right)^2}{E[I_1(\boldsymbol{\theta})]+m\pi^2\varepsilon_n{}^{-2}n^{-1}} \ge \frac{\inf_{\boldsymbol{\theta}\in J_n}\sum_{i=1}^m\left(g^{(i)}(\boldsymbol{\theta})\right)^2}{\sup_{\boldsymbol{\theta}\in J_n} I_1(\boldsymbol{\theta})+m\pi^2\varepsilon_n{}^{-2}n^{-1}}.$$

حال با استفاده از خواص $\underline{\lim}_{n\to\infty}$ و $\overline{\lim}_{n\to\infty}$ در [۳]، می‌توان نشان داد:

$$\underline{\lim}_{n\to\infty} \inf_{\boldsymbol{\theta}^*}\sup_{\boldsymbol{\theta}\in J_n} E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\sqrt{n}\,(g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta}))\right]^2 \ge \frac{\underline{\lim}_{n\to\infty}\inf_{\boldsymbol{\theta}\in J_n}\sum_{i=1}^m\left(g^{(i)}(\boldsymbol{\theta})\right)^2}{\overline{\lim}_{n\to\infty}\sup_{\boldsymbol{\theta}\in J_n} I_1(\boldsymbol{\theta})+\overline{\lim}_{n\to\infty} m\pi^2\varepsilon_n^{-2}n^{-1}} \qquad (۱۰)$$

که در آن $J_n=\{\boldsymbol{\theta}:\,|\boldsymbol{\theta}-\boldsymbol{\theta}_o|<\varepsilon_n\}\subset\Theta^o$.

اگر $\varepsilon_n\to 0$، با توجه به اینکه طبق شرایط (ج) و (و)، $I(\boldsymbol{\theta})$ و $g^{(i)}(\boldsymbol{\theta})$ها در $\boldsymbol{\theta}_o$ پیوسته‌اند، می‌توان تحقیق کرد که:

$$\underline{\lim}_{n\to\infty}\inf_{\boldsymbol{\theta}\in J_n}\sum_{i=1}^m g^{(i)}(\boldsymbol{\theta})=\sum_{i=1}^m g^{(i)}(\boldsymbol{\theta}_o),$$

$$\overline{\lim}_{n\to\infty}\sup_{\boldsymbol{\theta}\in J_n} I_1(\boldsymbol{\theta})=I_1(\boldsymbol{\theta}_o).$$

حال فرض کنید وقتی $n\to\infty$، $\frac{1}{n\varepsilon_n^2}\to 0$ (به‌عنوان‌مثال $\varepsilon_n=n^{-\alpha}$ و $0<\alpha<\frac{1}{2}$ را در نظر بگیرید)، در این صورت رابطه (۱۰) به‌صورت زیر یعنی حکم مسئله تبدیل می‌شود:

$$\underline{\lim}_{n\to\infty} \inf_{\boldsymbol{\theta}^*}\sup_{\boldsymbol{\theta}\in J_n} E_\theta\left[\sqrt{n}\,(g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta}))\right]^2 \ge \frac{\sum_{i=1}^m \left(g^{(i)}(\boldsymbol{\theta}_o)\right)^2}{I_1(\boldsymbol{\theta}_o)} \qquad (۱۱)$$

که در آن وقتی‌که $n\to\infty$، $\varepsilon_n,\frac{1}{n\varepsilon_n^2}\to 0$.

اکنون حالتی خاص از نامساوی بالا در نظرگرفته می‌شود هنگامی‌که مؤلفه‌های $\boldsymbol{\theta}=(\theta_1,...,\theta_m)$ از فضای پارامتر متعامد باشند ([۱۴])، یعنی $E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\frac{\partial}{\partial\theta_i}\log f(\boldsymbol{X},\boldsymbol{\theta})\frac{\partial}{\partial\theta_j}\log f(\boldsymbol{X},\boldsymbol{\theta})\right]=0$. توجه کنید که برای هر $i\neq j$ داریم:

$$E\left[G_i(\boldsymbol{X},\boldsymbol{\theta})G_j(\boldsymbol{X},\boldsymbol{\theta})\right]=E\left[\frac{k_i(\boldsymbol{X},\boldsymbol{\theta})k_j(\boldsymbol{X},\boldsymbol{\theta})}{f_n^2(\boldsymbol{X},\boldsymbol{\theta})q^2(\boldsymbol{\theta})}\right]$$
$$=\int_{\chi^n}\int_\Theta \tfrac{\partial}{\partial\theta_i}\log\left[f_n(\boldsymbol{x},\boldsymbol{\theta})q(\boldsymbol{\theta})\right] \times \tfrac{\partial}{\partial\theta_j}\log\left[f_n(\boldsymbol{x},\boldsymbol{\theta})q(\boldsymbol{\theta})\right]f_n(\boldsymbol{x},\boldsymbol{\theta})q(\boldsymbol{\theta})\mu^n(d\boldsymbol{x})d\boldsymbol{\theta}$$

اما با توجه به استقلال X_iها و شرط (د) و (ه)، مقدار فوق برابر صفر است. بنابراین ماتریس $\Sigma=((E(G_i(X,\theta)G_j(X,\theta)))$ یك ماتریس قطری است و ازاین‌رو با استفاده از لم ۲.۲ و با استدلال مشابه در قضیه ۲-۴ می‌توان نشان داد ([۱۴]):

$$E[g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})]^2 \ge \sum_{i=1}^m \frac{\left(E\left[g^{(i)}(\boldsymbol{\theta})\right]\right)^2}{E\left[I_n^{(i)}(\boldsymbol{\theta})\right]+E\left[\frac{\partial}{\partial\theta_i}\log q(\boldsymbol{\theta})\right]^2}.$$

حال قضیه زیر به‌سادگی از قضیه ۷.۲ حاصل می‌شود.

قضیهٔ ۸.۲. *([۱۴]) تحت برقراری شرایط (الف)-(و)، اگر مؤلفه‌های $\boldsymbol{\theta}$ متعامد باشند و $\varepsilon_n,\frac{1}{n\varepsilon_n^2}\to 0$، آنگاه:*

$$\underline{\lim}_{n\to\infty} \inf_{\boldsymbol{\theta}^*}\sup_{\boldsymbol{\theta}\in J_n} E_\theta\left[\sqrt{n}\,(g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta}))\right]^2 \ge \sum_{i=1}^m \frac{\left[g^{(i)}(\boldsymbol{\theta}_o)\right]^2}{I_1^{(i)}(\boldsymbol{\theta}_o)} \qquad (۱۲)$$

که در آن $J_n=\{\boldsymbol{\theta}:|\theta_i-\theta_{io}|<\varepsilon_n\}$.

تعریف ۹.۲. *([۱۴]) اگر نامساوی (۱۲) برای هر $\boldsymbol{\theta}\in\Theta$ به تساوی تبدیل شود، $g(\boldsymbol{\theta}^*)$ گویند یك برآوردگر مینیماکس مجانبی-موضعی $g(\boldsymbol{\theta})$ (LAME) است.*

در ادامه با ذکر چند مثال، کاربرد قضایای ذکرشده را در مسئله اتخاذ تصمیم مینیماکس بیان می‌شود.

مثال ۱۰.۲. *([۱۴]) فرض کنید X و Y متغیرهای تصادفی مستقل باشند که $X \sim E(\lambda)$ و $Y \sim E(\psi)$. حال با تعریف $\boldsymbol{\theta} = (\theta_1, \theta_2)$ که در آن $\lambda = \theta_1\theta_2^{-\frac{1}{2}}$ و $\psi = \theta_1\theta_2^{\frac{1}{2}}$، هدف یافتن یک برآوردگر $LAME$ برای $\frac{\psi}{\lambda}$ بر اساس یک نمونه تصادفی به‌اندازه n از $Z_i = (X_i, Y_i)$ها است. برای این منظور ابتدا با محاسبات مستقیم می‌توان نشان داد که مؤلفه‌های $\boldsymbol{\theta}$ بر هم عمودند. حال با تعریف $g(\boldsymbol{\theta}) = \theta_2$، به‌سادگی ملاحظه می‌شود که $I_1^{(2)}(\boldsymbol{\theta}_o) = (2\boldsymbol{\theta}_o^2)^{-1}$، $g^{(2)}(\boldsymbol{\theta}_o) = 1$ و $g^{(1)}(\boldsymbol{\theta}_o) = 0$. با جایگذاری این مقادیر در کران پایین (۱۲) خواهیم داشت:*

$$\lim_{n\to\infty} \inf_{\boldsymbol{\theta}^*} \sup_{\boldsymbol{\theta}\in J_n} E\left[\sqrt{n}\left(g(\boldsymbol{\theta}^*) - g(\boldsymbol{\theta})\right)\right]^2 \geq \sum_{i=1}^{2} \frac{\left[g^{(i)}(\boldsymbol{\theta}_o)\right]^2}{I_1^{(i)}(\boldsymbol{\theta}_o)} = 2\boldsymbol{\theta}_o^2.$$

از طرفی، با توجه به اینکه $\frac{\lambda\sum_{i=1}^{n} X_i}{\psi\sum_{i=1}^{n} Y_i} \sim F(2n, 2n)$، به‌سادگی می‌توان نشان داد که $g(\boldsymbol{\theta}^*) = \frac{n-1}{n}\frac{\lambda\sum_{i=1}^{n} X_i}{\psi\sum_{i=1}^{n} Y_i}$ یک برآوردگر نااریب برای $g(\boldsymbol{\theta})$ است و

$$E[g(\boldsymbol{\theta}^*) - g(\boldsymbol{\theta})]^2 = \mathrm{var}\left(g(\boldsymbol{\theta}^*)\right) = \frac{2n-1}{n(n-2)}\left(\frac{\psi}{\lambda}\right)^2 = \frac{2n-1}{n(n-2)}\theta_2^2.$$

حال تحت برقراری فرض (و)، اگر $\varepsilon_n \to 0$ وقتی $n \to \infty$ داریم:

$$\lim_{n\to\infty} \inf_{\boldsymbol{\theta}^*} \sup_{\boldsymbol{\theta}\in J_n} E_{\boldsymbol{\theta}}\left[\sqrt{n}\left(g(\boldsymbol{\theta}^*) - g(\boldsymbol{\theta})\right)\right]^2 = \lim_{n\to\infty} \sup_{\boldsymbol{\theta}\in J_n}\left(\frac{1}{I_1^{(2)}(\boldsymbol{\theta})}\right) = 2\boldsymbol{\theta}_o^2.$$

لذا طبق تعریف ۹.۲، $g(\boldsymbol{\theta}^*)$ یک $LAME$ برای $g(\theta) = \frac{\psi}{\lambda}$ است.

مثال ۱۱.۲. *([۱۴]) فرض کنید $\boldsymbol{X} = (X_1, ..., X_n)$ یک نمونه تصادفی nتایی از توزیع $N(\mu, \sigma^2)$ باشند. هدف به دست آوردن یک $LAME$ برای $g(\boldsymbol{\theta}) = \mu + b\sigma^2$ است که در آن $\boldsymbol{\theta} = (\mu, \sigma^2)$ و b مقدار معلوم است. از این که $g(\boldsymbol{\theta}^*) = \overline{X} + \frac{b}{n+1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ یک برآورد طبیعی برای $g(\boldsymbol{\theta})$ است، به نظر می‌رسد که یک مقدار مناسب برای b مقداری باشد که میانگین توان دوم خطا (MSE) برای $g(\boldsymbol{\theta}^*)$ را مینیمم کند. به‌سادگی می‌توان نشان داد که $g^{(1)}(\boldsymbol{\theta}_o) = 1$، $g^{(2)}(\boldsymbol{\theta}_o) = b$، $I_1^{(1)}(\boldsymbol{\theta}) = \sigma^{-2}$ و $I_1^{(2)} = (2\sigma^4)^{-1}$. حال با جایگذاری مقادیر فوق در عبارت سمت راست رابطه (۱۲) می‌توان نوشت:*

$$\lim_{n\to\infty} \inf_{\boldsymbol{\theta}^*} \sup_{\boldsymbol{\theta}\in J_n} E_{X|\theta}\left[g(\boldsymbol{\theta}^*) - g(\boldsymbol{\theta})\right]^2 \geq \sigma_o^2 + 2b\sigma_o^4. \quad (۱۳)$$

از طرف دیگر ملاحظه می‌شود که

$$E[g(\boldsymbol{\theta}^*) - g(\boldsymbol{\theta})]^2 = \mathrm{var}(\overline{X} + bS^2) = \frac{\sigma^2}{n} + \frac{2b^2}{n-1}\sigma^4.$$

بنابراین با استفاده از شرط (و):

$$\begin{aligned}
&\lim_{n\to\infty} \inf_{\boldsymbol{\theta}^*} \sup_{\boldsymbol{\theta}\in J_n} E_{\boldsymbol{\theta}}\left[\sqrt{n}\left(g(\boldsymbol{\theta}^*) - g(\boldsymbol{\theta})\right)\right]^2 \\
&= \lim_{n\to\infty} \sup_{\boldsymbol{\theta}\in J_n}(\sigma^2 + 2b\sigma^4) \\
&= \lim_{n\to\infty} \sup_{\boldsymbol{\theta}\in J_n}\left(\frac{1}{I_1^{(1)}(\boldsymbol{\theta})} + \frac{b}{I_1^{(2)}(\boldsymbol{\theta})}\right) \\
&= \frac{1}{I_1^{(1)}(\boldsymbol{\theta}_o)} + \frac{1}{I_1^{(2)}(\boldsymbol{\theta}_o)} \\
&= \sigma_o^2 + 2b^2\sigma_o^4.
\end{aligned}$$

با مقایسه عبارت فوق و رابطه (۱۳) نتیجه می‌شود که $g(\boldsymbol{\theta}^*)$ بنا بر تعریف ۹.۲ یک برآوردگر مینیماکس مجانبی-موضعی برای $g(\boldsymbol{\theta})$ است. هم‌چنین می‌توان نشان داد که مؤلفه‌های $\boldsymbol{\theta} = (\tau/\sigma, \tau + \sigma)$ متعامد هستند و $g(\boldsymbol{\theta}^*) = \frac{\bar{X}}{S^2}$ یک $LAME$ در تخمین $g(\boldsymbol{\theta}) = \frac{\mu}{\sigma^2}$ است که در آن $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$.

۳ برآورد مینیماکس مجانبی-موضعی برای پارامترهای توزیع نرمال چندمتغیره

در این بخش حالت چندمتغیره مثال ۱۱.۲ بررسی می‌شود. فرض کنید $\boldsymbol{X_i}$، $1 \leq i \leq n$ بردارهای تصادفی k-بعدی از توزیع نرمال چند متغیره با میانگین $\beta\boldsymbol{a}$ و ماتریس واریانس-کواریانس $\Sigma = V(\psi)$ باشند به‌طوری‌که β و ψ پارامترهای ثابت مجهول و $\boldsymbol{a}$ و V به ترتیب بردار k-بعدی و ماتریس $k \times k$ معلوم باشند. اگر β و ψ همبستگی تابعی نداشته باشند، با انجام محاسبات مستقیم می‌توان نشان داد که پارامترهای $\boldsymbol{\beta}$ و ψ متعامد هستند. مسئله موردعلاقه در اینجا، برآورد تابعی از پارامتر به‌صورت زیر است:

$$g(\beta, \psi) = \beta\boldsymbol{\alpha}^T{}_a + \boldsymbol{\gamma}^T V(\psi)\boldsymbol{\delta} \quad (۱۴)$$

که در آن γ، α و δ بردارهای معلوم k-بعدی هستند. با تعریف $\boldsymbol{\theta} = (\beta, \psi)$ و با اعمال مشتق‌گیری برداری از طرفین رابطه (۱۴) داریم:

$$\begin{aligned}
g^{(1)}(\boldsymbol{\theta}) &= \frac{\partial}{\partial\beta}g(\beta, \psi) = \boldsymbol{\alpha}^T\boldsymbol{a}, \\
g^{(2)}(\boldsymbol{\theta}) &= \frac{\partial}{\partial\psi}g(\beta, \psi) = \boldsymbol{\gamma}^T V'(\psi)\boldsymbol{\delta}
\end{aligned} \quad (۱۵)$$

که در آن $V'(\psi)$ ماتریس حاصل‌شده از مشتق‌گیری از درایه‌های $V(\psi)$ نسبت به ψ است [۱۴]. توجه کنید در اینجا فرض شده که درایه‌های

$V(\psi)$ نسبت به ψ مشتق‌پذیر هستند. اکنون با جایگذاری مقادیر (۱۵) در رابطه (۱۲) ملاحظه می‌شود:

$$\begin{aligned}\underline{\lim}_{n\to\infty}\ \inf_{\boldsymbol{\theta}^*}\sup_{\boldsymbol{\theta}\in J_n} E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\sqrt{n}\left(g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})\right)\right]^2 &\geq \sum_{i=1}^{2}\frac{\left[g^{(i)}(\boldsymbol{\theta}_o)\right]^2}{I_1^{(i)}(\boldsymbol{\theta}_o)}\\ &= \frac{\left(\boldsymbol{\alpha}^T\boldsymbol{a}\right)^2}{I_1^{(1)}(\boldsymbol{\theta}_o)}+\frac{\left[\gamma^T V'(\psi)\boldsymbol{\delta}\right]^2}{I_1^{(1)}(\boldsymbol{\theta}_o)}\end{aligned} \tag{16}$$

که در آن $q(\boldsymbol{\theta})$ چگالی پیشین معرفی‌شده در (۸) برای $m=2$ است و همچنین $J_n=\{\boldsymbol{\theta}:|\boldsymbol{\theta}-\boldsymbol{\theta}_o|<\varepsilon_n\}$ که در آن $\varepsilon_n^{-2}n^{-1}\to 0$، $\varepsilon_n\to 0$ و $f(\boldsymbol{x},\boldsymbol{\theta})=\frac{1}{(2\Pi)^{k/2}|\Sigma|^{1/2}}\mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}$ که در آن $\boldsymbol{\mu}=\beta\mathbf{a}$ و $\Sigma=V(\psi)$. برای به دست آوردن کران پایین رابطه (۱۶) به‌صورت صریح ابتدا برای $i=1,2$ مقادیر

$$I^{(i)}(\boldsymbol{\theta})=E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\frac{\partial}{\partial\boldsymbol{\theta}_i}\log f_n(\boldsymbol{X},\boldsymbol{\theta})\right]^2$$

را محاسبه می‌کنیم. از تابع چگالی نرمال چندمتغیره شروع می‌کنیم.

$$\begin{aligned}\log f(\boldsymbol{x},\boldsymbol{\theta})=&-\frac{k}{2}\log(2\pi)+\frac{1}{2}\log\left|\Sigma^{-1}\right|\\ &-\frac{1}{2}\left(\boldsymbol{x}^T\Sigma^{-1}\boldsymbol{x}-2\boldsymbol{x}^T\Sigma^{-1}\boldsymbol{\mu}+\boldsymbol{\mu}^T\Sigma^{-1}\boldsymbol{\mu}\right).\end{aligned} \tag{17}$$

با مشتق‌گیری از طرفین رابطه (۱۷) داریم:

$$\begin{aligned}\frac{\partial}{\partial\boldsymbol{\mu}}\log f\left(\boldsymbol{x}\,,\,\boldsymbol{\theta}\right)&=\Sigma^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}\right),\\ \frac{\partial}{\partial\beta}\log f(\boldsymbol{x},\boldsymbol{\theta})&=\left[vec\left(\frac{\partial}{\partial\boldsymbol{\mu}}\log f\right)\right]^T\left[Vec\left(\frac{d\boldsymbol{\mu}}{d\beta}\right)\right]\\ &=tr\left[\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\boldsymbol{a}\right]=\boldsymbol{a}^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}).\end{aligned}$$

بنابراین

$$\begin{aligned}I_1^{(1)}(\boldsymbol{\theta})&=E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\frac{\partial}{\partial\boldsymbol{\theta}}\log f(\boldsymbol{x},\boldsymbol{\theta})\right]^2\\ &=E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\boldsymbol{a}^T\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^T\Sigma^{-1}\boldsymbol{a}\right]\\ &=\boldsymbol{a}^T\Sigma^{-1}E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^T\right]\Sigma^{-1}\boldsymbol{a}\\ &=\boldsymbol{a}^T\Sigma^{-1}\boldsymbol{a}=\boldsymbol{a}^TV^{-1}(\psi)\boldsymbol{a}.\end{aligned}$$

برای محاسبه $I_1^{(2)}(\theta)$ نیز به‌صورت زیر عمل می‌کنیم:

با اعمال مشتق‌گیری جزئی از طرفین (۱۷) نسبت به Σ^{-1} داریم $\frac{\partial}{\partial\Sigma^{-1}}\log f(\boldsymbol{x},\theta)=\frac{1}{2}(\Sigma-Z)$ که در آن $Z=(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^T$.

از روابط بالا نتیجه می‌شود که:

$$\begin{aligned}\frac{\partial}{\partial\psi}\log f(\boldsymbol{x},\boldsymbol{\theta})&=\left[Vec\left(\frac{\partial}{\partial\Sigma^{-1}}\log f(\boldsymbol{x}\,,\boldsymbol{\theta})\right)\right]^T\left[Vec\left(\frac{d\Sigma^{-1}}{d\psi}\right)\right]\\ &=-\frac{1}{2}tr\left[(Z-\Sigma)\Sigma^{-1}\frac{d\Sigma}{d\psi}\Sigma^{-1}\right]\\ &=-\frac{1}{2}tr\left[(Z-V)V^{-1}V'V^{-1}\right]\end{aligned} \tag{18}$$

که در آن $V=\Sigma$ و $V'=\frac{\partial}{\partial\psi}\Sigma$. از طرفی از رابطه (۱۸) داریم:

$$E\left[\frac{\partial}{\partial\psi}\log f(\boldsymbol{x},\boldsymbol{\theta})\right]=-\frac{1}{2}tr\left[\left(E_{\boldsymbol{X}|\boldsymbol{\theta}}\left(Z\right)-V\right)\ V^{-1}V'V^{-1}\right]=0.$$

و در پی آن $I_1^{(2)}(\theta)=E_{\boldsymbol{x}|\theta}\left[\frac{\partial}{\partial\psi}\log f\right]^2=Var_\theta\left[\frac{\partial}{\partial\psi}\log f\right]$ و از رابطه (۱۸) داریم:

$$\begin{aligned}&Var\left\{\frac{1}{2}tr\left[(z-V)V^{-1}V'V^{-1}\right]\right\}\\ &=\frac{1}{4}Var_\theta\left\{tr\left[ZV^{-1}V'V^{-1}\right]\right\}\\ &=\frac{1}{4}Var_\theta\left\{tr\left[V^{\frac{1}{2}}V^{-\frac{1}{2}}ZV^{-\frac{1}{2}}V^{\frac{1}{2}}V^{-1}V'VV^{-1}\right]\right\}\\ &=\frac{1}{4}Var_\theta\left\{tr\left[V^{\frac{1}{2}}WV^{\frac{1}{2}}V^{-1}V'V^{-1}\right]\right\}\\ &=\frac{1}{4}Var_\theta\left\{tr\left[WV^{-\frac{1}{2}}V'V^{-\frac{1}{2}}\right]\right\}\\ &=\frac{1}{4}Var_\theta\left\{tr(AW)\right\}\end{aligned} \tag{19}$$

که در آن $A=V^{-\frac{1}{2}}V'V^{-\frac{1}{2}}$ و $W=V^{-\frac{1}{2}}ZV^{-\frac{1}{2}}$. اما طبق تجزیه طیفی ماتریس متقارن A داریم:

$$\begin{aligned}&H^TH=HH^T=I,\\ &\Omega=diag\left\{\lambda_1,...,\lambda_n\right\}\end{aligned} \tag{20}$$

که در آن H یك ماتریس متعامد از مرتبه k است و λ_i ها مقادیر ویژه ماتریس A هستند. با قرار دادن (۲۰) در رابطه (۱۹) داریم:

$$\begin{aligned}I_1^{(2)}(\theta)&=\frac{1}{4}Var_\theta\left[tr(H^T\Omega HH)\right]\\ &=\frac{1}{4}Var_\theta\left[tr(\Omega HWH^T)\right]\\ &=\frac{1}{4}Var_\theta\left[tr(\Omega\tilde{W})\right]\end{aligned} \tag{21}$$

که در آن $\tilde{W}=HWA^T$.

با توجه به اینکه $\boldsymbol{Y}=V^{-\frac{1}{2}}(\boldsymbol{X}-\boldsymbol{\mu})$ دارای توزیع نرمال با میانگین صفر و ماتریس واریانس-کوواریانس $I_{k\times k}$ (ماتریس همانی) است، با توجه به خواص توزیع ویشارت $W=YY^T\sim Wishart(I,k,1)$ و $\tilde{W}=H^TWH\sim Wishart(I,k,1)$. علاوه بر آن $\tilde{W}_{ij}$ها به ازای $1\leq i,j\leq k$ ناهم‌بسته‌اند و همچنین $Var(w_{ii})=2$. با توجه به روابط ذکرشده در بالا، رابطه (۲۱) به‌صورت زیر ساده می‌شود:

$$\begin{aligned}I_1^{(2)}(\boldsymbol{\theta})&=\frac{1}{4}Var_\theta\left[\sum_{i=1}^{k}\lambda_i\tilde{w}_{ii}\right]\\ &=\frac{1}{4}\sum_{i=1}^{k}\lambda_i^2Var_\theta(\tilde{w}_{ii})=\frac{1}{2}tr(A^2)\end{aligned}$$

زیرا از بسط طیفی ماتریس متقارن A داریم $A^2=H^T\Omega^2H$. بنابراین

$$tr(A^2)=tr(\Omega^2HH^T)=tr(\Omega^2)=\sum_{i=1}^{k}\lambda_i^2.$$

اما با توجه به تعریف A داریم $tr(A^2) = \frac{1}{2}tr(V^{-1}V^T)^2$ و ازاین‌رو

$$I_1^{(2)}(\boldsymbol{\theta}) = E_{\boldsymbol{x}|\boldsymbol{\theta}}\left[\frac{\partial}{\partial\psi}\log f(\boldsymbol{x},\boldsymbol{\theta})\right]^2 = \frac{1}{2}tr(V^{-1}V')^2.$$

حال با جایگذاری روابط بالا در رابطه (۱۶) وقتی $n \to \infty$، $\varepsilon_n \to 0$ و ${\varepsilon_n}^{-2}n^{-1} \to 0$ نتیجه زیر حاصل می‌شود:

$$\lim_{n\to\infty} \inf_{\boldsymbol{\theta}^*}\sup_{\boldsymbol{\theta}\in J_n} E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\sqrt{n}(g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta}))\right]^2 \geq \frac{(\boldsymbol{\alpha}^T\boldsymbol{a})^2}{\boldsymbol{a}^TV^{-1}(\psi_o)\boldsymbol{a}} + \frac{2\left[\gamma^TV'(\psi_o)\gamma\right]^2}{tr[V^{-1}(\psi_o)V'(\psi_o)]^2} \tag{22}$$

به‌طوری‌که $J_n = \{\boldsymbol{\theta} : |\boldsymbol{\theta}-\boldsymbol{\theta}_o| < \varepsilon_n\}$ [۱۴]. در حالت خاص فرض کنید $V(\psi)$ یك تابع خطی به‌صورت $V(\psi) = \Delta\psi$ باشد که در آن $\psi > 0$ و $\Delta_{k\times k}$ یك ماتریس معین مثبت است. با جایگذاری $V(\psi) = \Delta\psi$ در رابطه (۲۲) داریم [۱۴]:

$$\lim_{n\to\infty} \inf_{\boldsymbol{\theta}^*}\sup_{\boldsymbol{\theta}\in J_n} E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\sqrt{n}\,(g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta}))\right]^2 \geq \frac{\psi_o(\boldsymbol{\alpha}^T\boldsymbol{a})^2}{\boldsymbol{a}^T\Delta^{-1}\boldsymbol{a}} + \frac{2(\gamma^T\Delta\boldsymbol{\delta})^2\psi_o^2}{k}. \tag{23}$$

حال یك برآوردگر طبیعی $g(\boldsymbol{\theta}) = \beta\boldsymbol{\alpha}^Ta + \psi\gamma^T\Delta\boldsymbol{\delta}$ را به‌صورت $g(\boldsymbol{\theta}^*) = \boldsymbol{\alpha}^T\bar{X} + c\gamma^T\tilde{A}\boldsymbol{\delta}$ در نظر بگیرید که در آن c یك مقدار ثابت و $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ و $\tilde{A} = \frac{1}{n}\sum_{i=1}^n(\mathrm{X}_i-\bar{\mathrm{X}})(\mathrm{X}_i-\bar{\mathrm{X}})^T$ $\tilde{A} = \frac{1}{n}\sum_{i=1}^n(\boldsymbol{X}_i-\overline{\boldsymbol{X}})(\boldsymbol{X}_i-\overline{\boldsymbol{X}})^T$ و برآوردگرهای درستنمایی μ و Σ، برای $n > k$ است. هم‌چنین $\bar{X}$ و $\frac{n\tilde{A}}{n-1}$ به ترتیب برآوردگرهای نااریب μ و Σ هستند. اکنون با استفاده از این نکته که $\bar{X}$ و $\tilde{A}$ از هم مستقل هستند، می‌توان نشان داد:

$$\begin{aligned} E_{\boldsymbol{X}|\boldsymbol{\theta}}[g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})]^2 &= \frac{1}{n}\psi(\boldsymbol{\alpha}^T\Delta\boldsymbol{\alpha}) \\ &+ \psi^2(\gamma^T\Delta\boldsymbol{\delta})^2 \times \left\{\frac{c^2(n-1)^2}{n^2} - \frac{2c(n-1)}{n} + 1\right\} \\ &+ \psi^2\frac{(n-1)c^2}{n^2} \times \left[(\gamma^T\Delta\boldsymbol{\delta})^2 + (\gamma^T\Delta\gamma)(\boldsymbol{\delta}^T\Delta\boldsymbol{\delta})\right]. \end{aligned} \tag{24}$$

تابع خطی $V(\psi) = \Delta\psi + \Gamma$، ماتریس نیمه معین مثبت معلوم Γ، ماتریس معین مثبت $\Delta_{k\times k}$ و $\psi > 0$ را در نظر بگیرید. حال اگر $\Gamma = 0$، با جایگذاری $V(\psi) = \Delta\psi$ در رابطه (۲۲) خواهیم داشت [۱۴]:

$$\lim_{n\to\infty} \inf_{\boldsymbol{\theta}^*}\sup_{\boldsymbol{\theta}\in J_n} E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\sqrt{n}\,(g(\theta^*)-g(\boldsymbol{\theta}))\right]^2 \geq \psi_o\frac{(\boldsymbol{\alpha}^T\boldsymbol{a})^2}{\boldsymbol{a}^T\Delta^{-1}\boldsymbol{a}} + \frac{2(\gamma^T\Delta\gamma)^2\psi_o^2}{k^2}.$$

توجه کنید که $\overline{\boldsymbol{X}}$ و $\tilde{A}$ برآوردگرهای درستنمایی $\boldsymbol{\mu}$ و Σ هستند که در آن $n > k$ و هم‌چنین $\overline{\boldsymbol{X}}$ و $\frac{n\tilde{A}}{n-1}$ برآوردگرهای نااریب $\boldsymbol{\mu}$ و Σ هستند.

با استفاده از این نکته و اینکه $\overline{\boldsymbol{X}}$ و $\tilde{A}$ از هم مستقل هستند، ملاحظه می‌شود که ([۱۴]):

$$\begin{aligned} E_{\boldsymbol{X}|\theta}[g(\theta^*)-g(\theta)]^2 &= Var_\theta[g(\theta^*)] \\ &+ [E_{x|\theta}\,(g(\theta^*)) - g(\theta)]^2 \end{aligned} \tag{25}$$

$$E_{\boldsymbol{X}|\theta}\,(g(\theta^*)) = \beta\,\boldsymbol{\alpha}\,\boldsymbol{a} + c\left(\frac{n-1}{n}\right)\psi\,\boldsymbol{\gamma}^T\Delta\boldsymbol{\delta} \tag{26}$$

$$Var_\theta\,(g(\theta^*)) = \frac{1}{n}\psi\,\boldsymbol{\alpha}^T\Delta\boldsymbol{\alpha} + c^2Var_\theta(\boldsymbol{\gamma}^T\tilde{A}\boldsymbol{\delta}). \tag{27}$$

با تعریف $S = \sum_{i=1}^n(\boldsymbol{X}_i-\overline{\boldsymbol{X}})(\boldsymbol{X}_i-\overline{\boldsymbol{X}})^T$، جمله دوم سمت راست رابطه (۲۷) را می‌توان به‌صورت زیر محاسبه کرد:

$$Var_{\boldsymbol{\theta}}(\boldsymbol{\delta}^T\tilde{A}\boldsymbol{\delta}) = \frac{1}{n^2}Var_{\boldsymbol{\theta}}(\boldsymbol{\delta}^TS\boldsymbol{\delta}) = [tr(BS)] \tag{28}$$

که در آن $B = \boldsymbol{\delta}\gamma^T$ است. توجه کنید چون A متقارن است:

$$tr(BS) = tr\left[\left(\frac{B+B^T}{2}\right)S\right]. \tag{29}$$

در ادامه با انتخاب $D = \frac{B+B^T}{2}$ ابتدا لم مفید زیر را بیان می‌کنیم.

لم ۰.۱.۳ *([۱۴]) برای ماتریس متقارن D داریم*

$$Var_\theta\,[tr(DS)] = 2(n-1)tr(D\Sigma D\Sigma).$$

اثبات: با توجه به تعریف اثر ماتریس، برای دو ماتریس $D = (d_{ij})_{k\times k}$ و $\Sigma = (\sigma_{ij})_{k\times k}$ داریم:

$$tr(D\Sigma D\Sigma) = \sum_{ijpl}^k d_{ij}\sigma_{jp}d_{pl}\sigma_{\ell i}$$

و چون $D\Sigma D\Sigma$ متقارن است:

$$tr(D\Sigma D\Sigma) = \sum_{i,j}^k {d^2}_{ij}{\sigma^2}_{ij} + 2\sum_{i<j<p<e} d_{ij}\sigma_{jp}d_{pl}\sigma_{ei}.$$

از طرف دیگر داریم:

$$\begin{aligned} Var_\theta\,[tr(DS)] &= \sum_{i=1}^k\sum_{j=1}^k {d_{ij}}^2 Var(s_{ij}) \\ &+ 2\sum_{i<..<e}\mathrm{cov}(a_{ij}s_{ij}, a_{\rho e}s_{pe}) \end{aligned} \tag{30}$$

و $S \sim Wishart_k(\Sigma, n-1)$ و $Var_\theta(s_{ij}) = (n-1)({\sigma_{ij}}^2 + \sigma_{ii}\sigma_{jj})$ و $Cov_\theta(s_{ij}, s_{pe}) = (n-1)(\sigma_{ip}\sigma_{je} + \sigma_{ie}\sigma_{jp})$. با جایگذاری روابط دو رابطه اخیر در (۳۰)، ملاحظه می‌شود که:

$$\begin{aligned} &Var_\theta[tr(DS)] = \\ &2(n-1)\left[\sum_{i,j}^k {d^2}_{ij}{\sigma^2}_{ij} + 2\sum_{i<j<p<e} d_{ij}\sigma_{jp}d_{pl}\sigma_{ei}\right]. \end{aligned}$$

حال با مقایسه رابطه بالا و رابطه (۳۰) اثبات لم کامل است.

حال از روابط (۲۸) و (۲۹) نتیجه می‌شود:

$$\begin{aligned} Var_\theta[tr(BS)] &= 2(n-1)tr\left[\left(\tfrac{B+B^T}{2}\right)\Sigma\left(\tfrac{B+B^T}{2}\right)\Sigma\right] \\ &= 2(n-1)[tr(B\Sigma B\Sigma)+tr(B\Sigma B^T\Sigma)]. \end{aligned}$$

با قرار دادن $B=\boldsymbol{\delta}\gamma^T$ و $\Sigma=\Psi\Delta$ در عبارت فوق و سپس با جایگذاری مقدار حاصل در (۲۸)، رابطه

$$Var_\theta[\gamma^T\tilde{A}\boldsymbol{\delta}]=(\tfrac{n-1}{n^2})\Psi^2[(\gamma^T\Delta\boldsymbol{\delta})+(\gamma^T\Delta\gamma)(\boldsymbol{\delta}^T\Delta\boldsymbol{\delta})$$

به دست می‌آید. با قرار دادن رابطه فوق در رابطه (۲۶)، رابطه

$$\begin{aligned} &Var_\theta\left(g(\theta^*)\right)=\tfrac{1}{n}\psi\,\boldsymbol{\alpha}^T\Delta\boldsymbol{\alpha} \\ &\quad+c^2(\tfrac{n-1}{n^2})\Psi^2[(\gamma^T\Delta\boldsymbol{\delta})+(\gamma^T\Delta\gamma)(\boldsymbol{\delta}^T\Delta\boldsymbol{\delta}) \end{aligned}$$

به دست می‌آید. حال با قرار دادن عبارت فوق و رابطه (۲۶) در (۲۵) داریم ([۱۴]):

$$\begin{aligned} &E_{x|\theta}\left[g(\theta^*)-g(\theta)\right]^2= \\ &\tfrac{1}{n}\psi(\boldsymbol{\alpha}^T\Delta\boldsymbol{\alpha})+\psi^2(\gamma^T\Delta\boldsymbol{\delta})^2\left\{\tfrac{c^2(n-1)^2}{n^2}-\tfrac{2c(n-1)}{n}+1\right\} \\ &+\psi^2\tfrac{(n-1)c^2}{n^2}\left[(\gamma^T\Delta\boldsymbol{\delta})^2+(\gamma^T\Delta\gamma)(\boldsymbol{\delta}^T\Delta\boldsymbol{\delta})\right]. \end{aligned}$$

همچنین می‌توان نشان داد که مجموع مربعات خطای $g(\boldsymbol{\theta}^*)$ از $g(\boldsymbol{\theta})$ مینیمم خود را در $c=c^*$ اختیار می‌کند که در آن:

$$c^*=\frac{(\gamma^T\Delta\boldsymbol{\delta})^2}{(\gamma^T\Delta\boldsymbol{\delta})^2+\frac{1}{n}(\gamma^T\Delta\gamma)(\boldsymbol{\delta}^T\Delta\boldsymbol{\delta})}.$$

بنابراین $g_1(\boldsymbol{\theta}^*)=\boldsymbol{\alpha}^T\times c^*\left(\gamma^T\widetilde{A}\boldsymbol{\delta}\right)$ یک برآوردگر $g(\boldsymbol{\theta})$ است که میانگین مجموع مربعات خطا را مینیمم می‌کند. با قرار دادن در رابطه (۲۴) به‌سادگی ملاحظه می‌شود که ([۱۴]):

$$\begin{aligned} &\lim_{n\to\infty}E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\sqrt{n}(g_1(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta}))\right]^2= \\ &\psi(\boldsymbol{\alpha}^T\Delta\boldsymbol{\alpha})+\psi^2\left\{(\gamma^T\Delta\gamma)^2(\boldsymbol{\delta}^T\Delta\boldsymbol{\delta})+\gamma^T\Delta\boldsymbol{\delta})^2\right\}. \end{aligned}$$

توجه کنید که اگرچه $g_1(\boldsymbol{\theta}^*)$ یک برآورد بهینه از این نظر است که میانگین توان دوم خطا را مینیمم می‌کند اما واضح است که حد بالا در کران پایین نامساوی (۲۳) را به دست نمی‌دهد. لذا طبق تعریف ۹.۲، $g_1(\boldsymbol{\theta}^*)$ یک $LAME$ برای $g(\boldsymbol{\theta})$ نیست، اما همان‌گونه که ملاحظه شد، $g_1(\boldsymbol{\theta}^*)$ در حالت یک‌بعدی در مثال ۱۰.۲ یک برآوردگر $LAME$ است. لذا مسئله موردعلاقه، یافتن یک $LAME$ در تخمین پارامترهایی به‌صورت $g(\boldsymbol{\theta})$ است. برای این منظور فرض کنید $\boldsymbol{X}=(\boldsymbol{X}_1,...,\boldsymbol{X}_n)$ یک ماتریس باشد که در آن $\boldsymbol{X}_i$ بردارهای مستقل و هم‌توزیع از توزیع نرمال چند متغیره $N_k\left(\boldsymbol{\mu}=\beta\boldsymbol{a},\psi\Delta\right)$ باشند. مدل رگرسیونی $Y=\mathrm{z}\boldsymbol{\mu}^T+U$ را در نظر بگیرید که در آن $Y=(\boldsymbol{X}_1,...,\boldsymbol{X}_k)^T$، $\boldsymbol{Z}=1_n$ و $U=(\boldsymbol{u}_1,...,\boldsymbol{u}_n)^T$ که $\boldsymbol{u}_i\sim N_k(0,\psi\Delta)$، $i=1,...,k$. از طرفی می‌دانیم که برآوردگر رگرسیونی پارامتر μ به‌صورت

$$\boldsymbol{\mu}^T=\hat{\beta}_{\mathrm{a}}{}^T=\left(\left({}^T_n 1_n\right)^{-1}1_n^TY=\bar{X}^T\right.$$

و میانگین توان دوم خطا برابر با $E=\sum_{i=1}^n(\boldsymbol{X}_i-\overline{\boldsymbol{X}})(\boldsymbol{X}_i-\overline{\boldsymbol{X}})^T$ است. اکنون با تعریف $\hat{\beta}=\frac{\boldsymbol{a}^T\Delta^{-1}\overline{\boldsymbol{x}}}{\boldsymbol{a}^T\Delta\boldsymbol{a}}$ و $\hat{\psi}=\frac{tr(\Delta^{-1}S)}{(n-1)k}$، با توجه به [۸]، $\mathrm{tr}\left(\Delta^{-1}E\right)\sim\chi^2_{k(n-1)}$ و به‌سادگی می‌توان نشان داد که $\widehat{\psi}$ و $\widehat{\beta}$ برآوردهای نااریب ψ و β هستند. اکنون پارامتر موردعلاقه و برآوردگرهای متناظر را به‌صورت $g(\boldsymbol{\theta})=\beta\left(\boldsymbol{\alpha}^T\mathrm{a}\right)+\psi\left(\gamma^T\Delta\boldsymbol{\delta}\right)$ و همچنین $g_2\left(\boldsymbol{\theta}^*\right)=\hat{\beta}\left(\boldsymbol{\alpha}^T\mathbf{a}\right)+\hat{\psi}\left(\gamma^T\Delta\boldsymbol{\delta}\right)$ در نظر بگیرید. با محاسبات ساده می‌توان نشان داد که نامساوی (۲) را می‌توان به‌صورت زیر نوشت:

$$\begin{aligned} &\lim_{n\to\infty}\inf\nolimits_{\boldsymbol{\theta}^*}\sup\nolimits_{\boldsymbol{\theta}\in J_n}E_{X|\boldsymbol{\theta}}\left[\sqrt{n}\left(g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})\right)\right]^2 \\ &\geq\frac{(\boldsymbol{\alpha}^T\boldsymbol{a})^2}{I_1^{(1)}(\boldsymbol{\theta}_o)}+\frac{(\gamma^T\Delta\boldsymbol{\delta})^2}{I_1^{(2)}(\boldsymbol{\theta}_o)} \end{aligned} \tag{31}$$

که در آن $I_1^{(1)}(\boldsymbol{\theta}_o)=\frac{\boldsymbol{\alpha}^T\Delta^{-1}\boldsymbol{a}}{\psi_o}$، $I_1^{(2)}(\boldsymbol{\theta}_o)=\frac{k}{2\psi_o^2}$ و $\boldsymbol{\theta}_o=(\beta_o,\psi_o)$. از طرف دیگر به دلیل استقلال $\hat{\beta}$ و $\hat{\psi}$ و اینکه $g_2(\boldsymbol{\theta}^*)$ یک برآورد نااریب $g(\boldsymbol{\theta})$ است، می‌توان نوشت:

$$\begin{aligned} E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[g_2(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})\right]^2&=\frac{\psi(\boldsymbol{\alpha}^T\boldsymbol{a})^2}{n(\boldsymbol{\alpha}^T\Delta^{-1}\boldsymbol{a})}+\frac{2\psi^2}{(n-1)k}(\gamma^T\Delta\boldsymbol{\delta}) \\ &=\frac{(\boldsymbol{\alpha}^T\boldsymbol{a})}{I_1^{(1)}(\boldsymbol{\theta})}+\frac{(\gamma^T\Delta\boldsymbol{\delta})}{I_1^{(2)}(\boldsymbol{\theta})}. \end{aligned}$$

اما با استفاده از شرط (و) می‌توان نشان داد که

$$\begin{aligned} &\lim_{n\to\infty}\inf\nolimits_{\boldsymbol{\theta}^*}\sup\nolimits_{\boldsymbol{\theta}\in J_n}E_{\boldsymbol{X}|\boldsymbol{\theta}}\left[\sqrt{n}\left(g(\boldsymbol{\theta}^*)-g(\boldsymbol{\theta})\right)\right]^2 \\ &=\lim_{n\to\infty}\sup\nolimits_{\boldsymbol{\theta}\in J_n}\left(\frac{(\boldsymbol{\alpha}^T\boldsymbol{a})}{I_1^{(1)}(\boldsymbol{\theta})}+\frac{(\gamma^T\Delta\boldsymbol{\delta})}{I_1^{(2)}(\boldsymbol{\theta})}\right) \\ &=\frac{(\boldsymbol{\alpha}^T\boldsymbol{a})}{I_1^{(1)}(\boldsymbol{\theta}_o)}+\frac{(\gamma^T\Delta\boldsymbol{\delta})}{I_1^{(2)}(\boldsymbol{\theta}_o)} \\ &=\frac{\psi_o(\boldsymbol{\alpha}^T\boldsymbol{a})^2}{n(\boldsymbol{\alpha}^T\Delta^{-1}\boldsymbol{a})}+\frac{2\psi_o{}^2}{(n-1)k}(\gamma^T\Delta\boldsymbol{\delta}) \end{aligned} \tag{32}$$

که در آن $J_n=\{\boldsymbol{\theta}:|\boldsymbol{\theta}-\boldsymbol{\theta}_o|<\varepsilon_n\}$ و با $n\to\infty$، $\frac{1}{n\varepsilon_n^2}\to 0$، ε_n. با مقایسه روابط (۳۱) و (۳۲) بنا به تعریف (۹.۲) ملاحظه می‌شود که $g_2(\boldsymbol{\theta}^*)$ یک $LAME$ در تخمین مؤلفه‌های $g(\boldsymbol{\theta})$ است.

۴ نامساوی کرامر-رائو برای توابع زیان به فرم $g(x)=c|x|^e$

در بخش ۳ یک کران پایین برای مخاطره بیز کلاسی از برآوردگرهای پارامتر چندبعدی، بر اساس تابع زیان مربع خطا به دست آوردیم. در این بخش می‌خواهیم یك کران پایین برای مخاطره بیز تحت تابع زیان به‌صورت $L(\theta^*,\theta)=|g(\theta^*)-g(\theta)|^e$ که $e>1$ را به دست آوریم و به‌عنوان یك کاربردی از آن، یك کران پایین مینیماکس مجانبی-موضعی برای چنین توابع زیانی به دست آورده می‌شود و در واقع این بخش تعمیم بخش ۳ در حالت θ یک‌بعدی است. نمادهایی که در بخش ۳ بیان شد، در این بخش موردنیاز خواهد بود. فرض کنید شرایط مطلوب (الف)-(و) برقرار باشند.

قضیهٔ ۱.۴. *([۱۴]) اگر بین دو عدد حقیقی γ و $e>1$ رابطه $\frac{1}{\gamma}+\frac{1}{e}=1$ برقرار باشد و $E\,|g(\infty)G(x,\theta)|<\infty$ و $0<E|G(x,\infty)|^\gamma<\infty$ و $E\,|g(\theta)G(x,\theta)|<\infty$، آنگاه:*

$$\inf_{\theta^*}BR(g(\theta^*),g(\theta))=E|g(\theta^*)-g(\theta)|^e \qquad (۳۳)$$
$$\geq \frac{\left(E\,|g(\theta)G(x,\theta)|\right)^e}{\left(E|G(x,\theta)|^\gamma\right)^{e-1}}.$$

اثبات: به‌سادگی می‌توان تحقیق کرد که

$$\epsilon\left(\left(g(\theta^*)-g(\theta)\right)G(x,\theta)\right)=-\epsilon\left(g(\theta)\,G(x,\theta)\right).$$

با بکار بردن نامساوی هولدر ([۳]) برای عبارت سمت راست تساوی بالا داریم:

$$E\,|g(\theta)G(x,\theta)|\leq\left\{E|g(\theta^*)-g(\theta)|^e\right\}^{\frac{1}{e}}\left\{E|G(x,\theta)|^\gamma\right\}^{\frac{1}{\gamma}}$$

و از اینجا رابطه (۳۳) به دست می‌آید. مشابه محاسبات بخش ۲، می‌توان نشان داد که $E\left[g(\theta)G(X,\theta)\right]=-E\left[g'(\theta)\frac{h(\theta)}{q(\theta)}\right]$ و

$$E|G(X,\theta)|^\gamma=E\left[\left|\frac{h'(\theta)}{q(\theta)}+\frac{\partial}{\partial\theta}\log f_n(X,\theta),\,\frac{h(\theta)}{q(\theta)}\right|^\gamma\right].$$

حال با جایگذاری دو رابطه اخیر در (۳۳) داریم:

$$BR(\theta^*,\theta)\geq\frac{\left|E\left(\frac{h(\theta)}{q(\theta)}\right)\right|^e}{\left\{E\left|\frac{h'(\theta)}{q(\theta)}+\frac{\partial}{\partial\theta}\log f_n(X,\theta)\frac{h(\theta)}{q(\theta)}\right|^\gamma\right\}^{e-1}}. \qquad (۳۴)$$

نتیجه ۲.۴. *([۱۴]) اگر در رابطه (۳۴) به ازای $\gamma=e=2$، $h(\theta)$ را به ترتیب $\frac{q(\theta)}{I_n(\theta)}$ و $q(\theta)$ انتخاب کنیم، آنگاه طبق نتیجه ۵.۲ و ۶.۲ به ازای $m=1$ به ترتیب روابط زیر حاصل می‌شوند:*

$$E[g(\theta^*)-g(\theta)]^2\geq\frac{\frac{1}{n}\left(E\left[\frac{g'(\theta)}{I_1(\theta)}\right]\right)^2}{\frac{1}{n}E\left[\frac{1}{I_1(\theta)}\right]+\epsilon\left[\frac{h'(\theta)}{q(\theta)}\right]^2},$$
$$E[g(\theta^*)-g(\theta)]^2\geq\frac{E\left[g'(\theta)\right]}{nE[I_1(\theta)]+E\left[\frac{\partial}{\partial\theta}\log q(\theta)\right]^2}.$$

رابطه زیر همواره برقرار است:

$$\sup_{\theta\in j_n}E_{X|\theta}\left|\sqrt{n}(g(\theta^*)-g(\theta))\right|^e\geq E_{X|\theta}\left|\sqrt{n}(g(\theta^*)-g(\theta))\right|^e$$

که در آن Θ_o مجموعه نقاط درونی Θ است و

$$j_n=\{\theta:|\theta-\theta_o|<\varepsilon_n\,,\,\theta_o\in\Theta^o\}.$$

حال اگر $q(\theta)$ یك چگالی پیشین دلخواه تعریف‌شده در j_n باشد، برای $\theta\in j_n$ از رابطه بالا نتیجه می‌شود:

$$\sup_{\theta\in j_n}E_{X|\theta}\left|\sqrt{n}(g(\theta^*)-g(\theta))\right|^e\geq$$
$$\int_{J_n}E_{x|\theta}\left|\sqrt{n}(g(\theta^*)-g(\theta))\right|^e q(\theta)d\theta$$
$$=E\left|\sqrt{n}(g(\theta^*)-g(\theta))\right|^e.$$

حال با توجه به استقلال X_iها و رابطه $e=\gamma(e-1)$، اگر $h(\theta)=q(\theta)$ در رابطه (۲۷) اختیار شود، با مقایسه آن و رابطه بالا، نامساوی زیر حاصل می‌شود:

$$\inf_{\theta^*}\sup_{\theta\in j_n}E_{X|\theta}\left|\sqrt{n}(g(\theta^*)-g(\theta))\right|^e\geq \qquad (۳۵)$$
$$\left\{\frac{\left(E\left[g'(\theta)\right]\right)^\gamma}{E\left|\frac{\partial}{\partial\theta}\log q(\theta)+\sum_{i=1}^n\frac{\partial}{\partial\theta}\log f(x_i,\theta)\right|^\gamma}\right\}^{e-1}$$
$$\geq\left\{\frac{\inf\limits_{\theta\in j_n}\left[g'(\theta)\right]^\gamma}{E\left|\frac{\partial}{\partial\theta}\log q(\theta)+\sum_{i=1}^n\frac{\partial}{\partial\theta}\log f(x_i,\theta)\right|^\gamma}\right\}^{e-1}$$

به ازای مقادیر مختلف e و γ دو حالت را در نظر می‌گیریم:

حالت اول- *فرض کنید $\gamma=2k$، $k\geq1$ ($e=\frac{2k}{2k-1}\leq2$). فرض کنید $q(.)$ یك چگالی پیشین در j_n باشد که وقتی $n\to\infty$، $\frac{1}{n}E\left|\frac{\partial}{\partial\theta}\log q(\theta)\right|^\gamma\to0$. بعلاوه لم مفید زیر را در نظر بگیرید.*

لم ۳.۴. *([۷]) به ازای متغیرهای تصادفی مستقل $X_1,...,X_n$ با میانگین صفر و برای $p>2$، $E|s_n|^p\leq F_pn^{\frac{p}{2}-1}\sum_{i=1}^nE|X_i|^p$ که $s_n=\sum_{i=1}^nX_i$، $F_p=\frac{1}{2}p(p-1)\max(1,2^{p-3})\left[1+2p^{-1}D_{2m}^{\frac{p-2}{2m}}\right]$ و $D_o=0$، $D_{2m}=\sum_{i=1}^m\frac{K^{2m-1}}{(k-1)!}$، $m=1,2,...$ و $2m\leq p\leq 2m+2$.*

اثبات: تعاریف $\beta_{\rho,n}=\frac{1}{n}\sum_{i=1}^nE|X_i|^p$ و $\gamma_{p,i}=E|X_i|^p$، $i=1,2,...,n$ را در نظر بگیرید. با توجه به اینکه $\sum_{i=1}^nE|X_i|^p=$

$n\beta_{\rho,n}$، اگر $\beta_{p,n} = \infty$، اثبات کامل است. بنابراین فرض می‌کنیم $\beta_{n,p} < \infty$. بسط مک‌لورن $|s_{n-1} + x_n|$ به‌صورت زیر است:

$$|s_n|^\rho = |s_{n-1}|^\rho + p\; sgn(s_n - 1)|s_{n-1}|^{\rho-1}x_n + \frac{p(p-1)}{2}|s_{n-1} + \theta\, x_n|^{p-2}{x_n}^2, \theta \in (0, 1).$$

با توجه به استقلال X_iها و نامساوی مثلثی

$$|x + y|^p \le \max(1, 2^{p-1})(|x|^p + |y|^p), p > 0, x, y \in \mathbb{R}$$

داریم $E|\Delta_n(p)|^p \le 2^{-1}p\delta_p \left\{E|s_{n-1}|^{p-2}\gamma_{2,n} + \gamma_{p,n}\right\}$ که در آن $\Delta_n(p) = E(|sn|^p - |s_{n-1}|^p)$ و $\delta_p = (p-1)\max(1, 2^{p-3})$. از طرفی با استفاده از نامساوی لیاپانوف برای $2m \ge p - 2$، داریم:

$$E|s_{n-1}|^{p-2} \le \left(E|s_{n-1}|^{2m}\right)^{\frac{p-2}{2m}}. \tag{36}$$

حال قضیه زیر را در نظر بگیرید.

قضیهٔ ۴.۴. *اگر برای هر مقدار صحیح $p \ge 1$ و برای هر انتخاب از اعداد صحیح مثبت $k_1, ..., k_p, i_1, .., i_p$ با شرط $\min(k_1, ..., k_p) = 1$، نتیجه دهد که $E(X_{i_1}^{k_1}...X_{i_p}^{k_p}) = 0$، آنگاه برای.. $m = 1, 2, ..$ داریم:*

$$E|s_n|^{2m} \le D_{2m}n^{m-1}\sum_{i=1}^n E|X_i|^{2m},$$
$$D_{2m} = \sum_{i=1}^m \tfrac{k^{2m-1}}{(k-1)!}.$$

رابطه (۳۶) را با توجه به قضیه ۴.۴ می‌توان به‌صورت

$$E|s_{n-1}|^{2m} \le D_{2m}(n-1)^m\beta_{2m,2n}$$

ساده نمود. بنابراین

$$E|s_{n-1}|^{p-2} \le D_{2m}^{\frac{p-2}{2m}}(n-1)^{\frac{p-2}{2}}\beta_{2m,n-1}^{\frac{p-2}{2m}}. \tag{37}$$

از طرفی با استفاده از (۳۶) می‌توان نشان داد که $\beta_{2m,n-1} \le \beta_{p,n-1}^{2m/p}$، $2m \le p$ و $\gamma_{2,n} \le \gamma_{p,n}^{\frac{2}{p}}$، $p \ge 2$. حال با تلفیق روابط (۳۶) و (۳۷) با دو رابطه اخیر ملاحظه می‌شود که:

$$\Delta_n(p) \le \tfrac{1}{2}p\delta_p\left\{D_{2m}^{\frac{p-2}{2m}}(n-1)^{\frac{p-2}{p}}\beta_{p,n-1}^{\frac{p-2}{p}}\gamma_{p,n}^{\frac{2}{p}} + \gamma_{p,n}\right\}.$$

بنابراین با جایگذاری $x = \frac{p}{2}$، $z_n = \beta_{p,n}$ و $y_n = \gamma_{p,n}$ داریم:

$$\begin{aligned}
&E|s_n|^p\\
&= \textstyle\sum_{j=1}^n \Delta_p(j)\left\{D_{2m}^{\frac{p-2}{2m}}(j-1)^{\frac{p-2}{p}}\beta_{p,j-1}^{\frac{p-2}{p}}\gamma_{p,j}^{\frac{2}{p}} + \gamma_{n,j}\right\}\\
&\le \tfrac{1}{2}p\delta_p\textstyle\sum_{j=1}^n\left\{D_{2m}^{\frac{p-2}{2m}}(j-1)^{\frac{p-2}{p}}\beta_{p,j-1}^{\frac{p-2}{p}}\gamma_{p,j}^{\frac{2}{p}} + \gamma_{n,j}\right\}\\
&= \tfrac{1}{2}p\delta_p D_{2m}^{\frac{p-2}{2m}}\textstyle\sum_{j=1}^n(j-1)^{\frac{p-2}{p}}\beta_{p,j-1}^{\frac{p-2}{p}}\gamma_{p,j}^{\frac{2}{p}}\\
&+ \tfrac{1}{2}n\,p\delta_p\beta_{p,n}\\
&\le \tfrac{1}{2}p\delta_p\left(D_{2m}^{\frac{p-2}{p}}n^{\frac{p}{2}}\tfrac{2}{p}\beta_{p,n}\right) + \tfrac{n}{2}p\delta_p\beta_{p,n}\\
&= n^{\frac{p}{2}}\delta_p\beta_{p,n}\left(D_{2m}^{\frac{p-2}{2m}} + \tfrac{1}{2}pn^{1-\frac{p}{2}}\right)\\
&\le \tfrac{1}{2}p\delta_p n^{\frac{p}{2}}\beta_{p,n}\left(1 + \tfrac{2}{p}D_{2m}^{\frac{p-2}{2m}}\right).
\end{aligned}$$

با تعریف $F_p = \frac{1}{2}p\delta_p\left(1 + \frac{2}{p}D_{2m}^{\frac{p-2}{2m}}\right)$ اثبات کامل می‌شود.
حال اگر $p = \gamma$ و $s_{n+1} = \sum_{i=1}^{n+1}Y_i$ که در آن $Y_i = \frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta}\log f(x_i, \theta)$، $i = 1, ..., n$ و $Y_{n+1} = \frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta}\log q(\theta)$، آنگاه با توجه به لم ۳.۴ داریم:

$$E|S_{n+1}|^\gamma = E\Big|\tfrac{1}{\sqrt{n}}\tfrac{\partial}{\partial\theta}\log q(\theta) + \tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^n\tfrac{\partial}{\partial\theta}\log f(X_i, \theta)\Big|^\gamma \tag{38}$$

$$\le F_\gamma\frac{(n+1)^{\frac{\gamma}{2}-1}}{n^{\frac{\gamma}{2}}}\left\{E\Big|\tfrac{\partial}{\partial\theta}\log q(\theta)\Big|^\gamma + \textstyle\sum_{i=1}^n E\Big|\tfrac{\partial}{\partial\theta}\log f(x_i, \theta)\Big|^\gamma\right\}.$$

با تعریف $I_\gamma(\theta) = E_{X_1|\theta}\left|\frac{\partial}{\partial\theta}\log f(X_1, \theta)\right|^\gamma$ چون X_iها هم‌توزیع هستند، داریم $\sum_{i=1}^n E\left|\frac{\partial}{\partial\theta}\log F(x_i, \theta)\right|^\gamma = nE_\theta\left[I_\gamma(\theta)\right]$. حال با قرار دادن مقدار فوق در عبارت سمت راست رابطه (۳۸) خواهیم داشت:

$$\begin{aligned}
&\overline{\lim_{n\to\infty}}\left\{E\left|\tfrac{1}{\sqrt{n}}\tfrac{\partial}{\partial\theta}\log g(\theta) + \tfrac{1}{\sqrt{n}}\sum_{i=1}^n\tfrac{\partial}{\partial\theta}\log F(X_i, \theta)\right|^\gamma\right\}^{e-1} &(39)\\
&\le \overline{\lim_{n\to\infty}}\left\{F_\gamma\tfrac{(n+1)^{\frac{\gamma}{2}-1}}{n^{\frac{\gamma}{2}}}\left[nB_\theta\left[I_\gamma(\theta)\right] + E\left|\tfrac{\partial}{\partial\theta}\right|\log q(\theta)^\gamma\right]\right\}^{e-1}\\
&\le F_\gamma^{\ell-1}\left\{\overline{\lim_{n\to\infty}}\left(\tfrac{n+1}{n}\right)^{\frac{\gamma}{2}}\tfrac{n}{n+1}E_\theta[I_\gamma(\theta)] + \overline{\lim_{n\to\infty}}\left(\tfrac{n+1}{n}\right)^{\frac{\gamma}{2}}\tfrac{1}{n}E\left|\tfrac{\partial}{\partial\theta}\log q(\theta)\right|^\gamma\right\}^{e-1}\\
&\le F_\gamma^{\ell-1}\left\{\overline{\lim_{n\to\infty}}\sup_{\theta\in j_n}I_\gamma(\theta)\right\}^{e-1}.
\end{aligned}$$

با استدلال مشابه بخش ۲، می‌توان نشان داد که اگر $\varepsilon_n \to 0$، آنگاه:

$$\begin{aligned}
&\overline{\lim_{n\to\infty}}\sup\nolimits_{\theta\in j_n}I_\gamma(\theta) = I_\gamma(\theta_o), &(40)\\
&\underline{\lim_{n\to\infty}}\inf\nolimits_{\theta\in j_n}\left[g'(\theta)\right]^\gamma = g'(\theta_o)^\gamma.
\end{aligned}$$

حال اگر از طرفین رابطه (۳۵) حد بگیریم و سپس آن را با روابط (۳۹)

و (۴۰) تلفیق کنیم، خواهیم داشت:

$$\begin{aligned}&\underline{\lim}_{n\to\infty}\ \inf_{\theta^*}\sup_{\theta\in j_n} E_{X|\theta}\left[\sqrt{n}(g(\theta^*)-g(\theta))\right]^{\frac{2k}{2k-1}}\\ &\ge \frac{|g'(\theta_o)|^e}{\{F_\gamma\, I_\gamma(\theta_o)\}^{e-1}}\end{aligned}$$

به‌طوری‌که $\gamma=2k$ و $e=\frac{2k}{2k-1}$ و $k\ge 1$.

نتیجه ۵.۴. *به ازای $k=1$، F_2 برابر یک است و نامساوی فوق رابطه (۳۷) را در حالت یک‌بعدی، نتیجه می‌دهد.*

$$\underline{\lim}_{n\to\infty}\ \inf_{\theta^*}\sup_{\theta\in j_n} E_{x|\theta}\left[\sqrt{n}(g(\theta^*)-g(\theta))\right]^2 \ge \frac{[g'(\theta^*)]^2}{I_1(\theta_o)}$$

که در آن $j_n=\{\theta:|\theta-\theta_o|<\varepsilon_n\}$ و وقتی $n\to\infty$، $\varepsilon_n{}^{-2}n^{-1}\to 0$ و $\varepsilon_n\to 0$.

حالت دوم- *فرض کنید $e\ge 2$ $(\gamma\le 2)$. در این صورت یک کران پایین قابل‌محاسبه برای مخاطره برآورد مینیماکس نتیجه ۴.۲ را می‌توان به‌صورت زیر به دست آورد:*

از نامساوی اول (۳۴) با تعریف

$$Z=\tfrac{1}{\sqrt{n}}\tfrac{\partial}{\partial\theta}\log q(\theta)+\tfrac{1}{\sqrt{n}}\sum_{i=1}^{n}\tfrac{\partial}{\partial\theta}\log f(x_i,\theta)$$

داریم:

$$\begin{aligned}&\underline{\lim}_{n\to\infty}\ \inf_{\theta^*}\sup_{\theta\in j_n} E\left[\sqrt{n}(g(\theta^*)-g(\theta))\right]^e\\ &\ge \frac{(E\,[g'(\theta)])^e}{\{E|Z|^\gamma\}^{e-1}}\end{aligned} \tag{۴۱}$$

اما طبق فرض قضیه ۱.۴، $E(Z^2)<\infty$، لذا طبق نامساوی لیاپانوف یعنی (۳۶) برای $1<\gamma\le 2$ داریم $E|Z|^\gamma\le (E(Z^2))^{\frac{\gamma}{2}}$. از طرفی تحت برقراری شرط مطلوب (و)، مقدار عبارت $E_{X|\theta}\left[\frac{\partial}{\partial\theta}\log f(X_i,\theta)\right]$ صفر است. با استفاده از این نکته و استقلال X_iها می‌توان نشان داد که:

$$E(Z^2)=E(\tfrac{1}{\sqrt{n}}\tfrac{\partial}{\partial\theta}\log q(\theta))^2+E_\theta(I_{X_1}(\theta)).$$

حال با استفاده از نامساوی مثلثی $|x+y|^p\le 2^{p-1}(|x|^p+|y|^p)$، $p>0$ رابطه زیر نتیجه می‌شود:

$$E(z^2)^{\frac{e}{2}}\le 2^{\frac{e}{2}-1}\left\{(E(\tfrac{1}{\sqrt{n}}\tfrac{\partial}{\partial\theta}\log q(\theta))^2)^{\frac{e}{2}}+(E(I_{X_1}(\theta))^{\frac{e}{2}}\right\}^{\frac{e}{2}}.$$

با تلفیق رابطه بالا و رابطه (۴۱) داریم:

$$\begin{aligned}&\underline{\lim}_{n\to\infty}\ \inf_{\theta^*}\sup_{\theta\in j_n} E\left[\sqrt{n}(g(\theta^*)-g(\theta))\right]^e\\ &\ge \frac{2^{1-\frac{e}{2}}\,(E\,[g'(\theta)])^e}{\overline{\lim}_{n\to\infty}\left\{\left[E(\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta}\log q(\theta))^2\right]^{\frac{e}{2}}+E(I_{X_1}(\theta))^{\frac{e}{2}}\right\}}\\ &\ge \frac{2^{1-\frac{e}{2}}\ \underline{\lim}_{n\to\infty}\ \inf_{\theta\in j_n}[g'(\theta)]^e}{\overline{\lim}_{n\to\infty}\left[E(\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta}\log q(\theta))^2\right]^{\frac{1}{2}}+\overline{\lim}_{n\to\infty}\sup_{\theta\in j_n}(I_{X_1}(\theta))^{\frac{e}{2}}}.\end{aligned} \tag{۴۲}$$

با انتخاب چگالی پیشین $q(\theta)=\frac{1}{\varepsilon_n}\cos^2\frac{\pi(\theta-\theta_o)}{2\varepsilon_n}$، $\theta\in j_n$ در j_n، می‌توان نشان داد $E(\frac{\partial}{\partial\theta}\log q(\theta))^2=\pi^2\varepsilon_n^{-2}$. با انتخاب $\varepsilon_n=n^{-\alpha}$ و $0<\alpha<1$ از رابطه بالا داریم:

$$\lim_{n\to\infty}\tfrac{1}{n}E(\tfrac{\partial}{\partial\theta}\log q(\theta))^2=0 \tag{۴۳}$$

و از طرف دیگر با ارائه دلایل مشابه در دستیابی رابطه (۳۷) می‌توان نشان داد که اگر $\varepsilon_n\to 0$، آنگاه:

$$\begin{aligned}&\overline{\lim}_{n\to\infty}\sup_{\theta\in j_n}(I_{X_1}(\theta))^{\frac{e}{2}}=(I_{X_1}(\theta_o))^{\frac{e}{2}},\theta_o\in\Theta^o,\\ &\underline{\lim}_{n\to\infty}\ \inf_{\theta\in j_n}(g'(\theta))^e=(g'(\theta_o))^e.\end{aligned}$$

با قرار دادن رابطه (۴۳) و نیز دو رابطه بالا در نامساوی (۴۲) داریم:

$$\begin{aligned}&\overline{\lim}_{n\to\infty}\inf_{\theta^*}\sup_{\theta\in j_n}E\left[\sqrt{n}(g(\theta^*)-g(\theta))\right]^e\\ &\ge\frac{2^{1-\frac{e}{2}}\,(E\,[g'(\theta_o)])^e}{(I_{X_1}(\theta_o))^{\frac{e}{2}}}\end{aligned}$$

که در آن $j_n=\{\theta:|\theta-\theta_o|<\varepsilon_n\}$ و وقتی $n\to\infty$، $\varepsilon_n{}^{-2}n^{-1}\to 0$ و $\varepsilon_n\to 0$.

۵ بحث و نتیجه‌گیری

در این مقاله تصمیم‌های مینیماکس از دیدگاه بیز برای حالت بزرگ نمونه‌ای بررسی شدند و در چارچوب نظریه بیز، تصمیم مینیماکس توسط نامساوی اطلاع برای مخاطره بیز به دست آمد. در این روش ابتدا یک نامساوی اساسی برای مخاطره بیز تابعی از یک برآوردگر دلخواه تحت تابع زیان مربع خطا در نظر گرفته شد و به کمک آن برآوردگرهای مینیماکس مجانبی-موضعی در حالت یک متغیره و چند متغیره به دست آمدند. همچنین در حالتی که مؤلفه‌های پارامتر در توزیع نرمال چند متغیره، متعامد هستند، کاربرد این روش در تعیین برآوردگرهای مینیماکس مجانبی-موضعی برای تابعی از بردار میانگین و ماتریس کواریانس مطرح شد. در پایان کران‌های نامساوی اطلاع تحت یک

تابع زیان عمومی محاسبه شد. در [۲۱] نسخه توسعه‌یافته‌ای از نامساوی کوشی-شوارتز ارائه داده شد که نسخه احتمالی آن به این صورت است که برای متغیرهای تصادفی X و Y که روی یک فضای احتمال تعریف‌شده و گشتاورهای دوم متناهی دارند، داریم

$$|E(XY)|^۲ \leq E(X^۲)E(Y^۲) - \left[|E(X)|\sqrt{Var(Y)} - |E(Y)|\sqrt{Var(X)}\right]^۲$$

که برای حالت‌هایی که $E(X) \neq ۰$ یا $E(Y) \neq ۰$ نامساوی کوشی-شوارتز را به‌صورت اکید توسعه می‌دهد. در حالت خاص $E(Y) = ۰$ داریم $Var(X) \geq \frac{|E(XY)|^۲}{E(Y^۲)}$ که در [۱۶] بر اساس نامساوی اخیر کران‌های پایین جدید میانگین مربعات خطا برای برآوردگرهای پارامتر که خود یک متغیر تصادفی است معرفی می‌شود. در [۱۸] تعمیم‌های دیگری از نامساوی کرامر-رائو مرور می‌شود. محققان می‌توانند در ادامه مقاله حاضر از مفاهیم ارائه‌شده در [۱۶، ۱۸] استفاده کرده و مفاهیم ذکرشده در این مقاله را از دیدگاه تعمیم‌های نامساوی کوشی-شوارتز و نامساوی کرامر-رائو گسترش دهند.

مراجع

[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] Apostel, T.M., (1974). *Mathematical Analysis.* Addison-Wesley Publishing Co., Reading, Mass.

[4] Bobrovsky B.Z., Mayer-Wolf, E, and Zakai, M. (1987). Some classes of global Cramer-Rao bounds, *Annals of Mathematical Statistics.* **15**, 1421-1438.

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

[6] Cianchi, A., Lutwak, E., Yang, D., and Zhang, G. (2013). A Unified Approach to Cramér-Rao Inequalities, *IEEE Transactions on Information Theory.* **60**, 643-650.

[7] Dharmadhikari, S.W., and Jogdeo, K. (1968). Bounds on Moments of certain random variable, *Annals of Mathematical Statistics.* **40**, 1506-1508.

[8] Ferriera, P. E. (1981). Extending Fisher's measure of information, *Biometrika.* **68**, 695-698.

[9] Gajek, L. (1987). An improper Cramer-Rao lower bound, *Zastos. Math.* **19**, 241-256.

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

[11] Kelbert, M., and Mozgunov, P. (2017). Generalization of Cramer-Rao and Bhattacharyya inequalities for the weighted covariance matrix, *Mathematical Communications.* **22**, 25-40.

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

[13] Miller. R.G. (1983). What price Kaplan- Meier?, *Biometrics.* **39**, 1077-1081.

[14] Prakasa Rao, B.L.S. (1992). Cramer-Rao type integral inequalities for functions of multidimensional parameter, *Sankhya Ser. A.* **31**, 54-73.

[15] Prakasa Rao, B.L.S. (2001). Cramer-Rao Type Integral Inequalities for General Loss Function, *Test.* **10**, 105-120.

[16] Prakasa Rao, B.L.S. (2018). Improved Cramer-Rao Type Integral Inequalities or Bayesian Cramer-Rao Bounds. *Journal of the Indian Society for Probability and Statistics.* **19(1)**, 1-7.

[17] Prakasa Rao, B.L.S. (2020). Cramer-Rao inequality revisited. *Proceedings-Mathematical Sciences.* **130(1)**, 1-11.

[18] Rao, C.R., and Toutenburg, H. (1999). Linear Models. *Springer series in statistics.* New York.

[19] Targhtta. M.L. (1984). On Bayesian analogues to Bhattacharya's lower bounds, *Arab Gulf Journal of Scientific Research.* **2**, 583-59.

[20] Targhtta. M.L. (1988). On the attainment of the lower bound for the Bayes risk in estimating a parametric function, *Statistics.* **19**, 233-239.

[21] Walker S.G. (2017). A self-improvement to the Cauchy–Schwarz inequality. *Statistics & Probability Letters.* **122**, 86–90.

Asymptomatically-Locally minimax estimation for multivariate normal distribution parameters

Mehdi Shams [1], Gholamreza Hesamian [2]

Abstract:

Information inequalities have many applications in estimation theory and statistical decision making. This paper describes the application of an information inequality to make the minimax decision in the framework of Bayesian theory. In this way, first a fundamental inequality for Bayesian risk is introduced under the square error loss function and then its applications are expressed in determining asymptotically and locally minimax estimators in the case of univariate and multivariate. In the case that the parameter components are orthogonal, the asymptotic-local minimax estimators are obtained for a function of the mean vector and the covariance matrix in the multivariate normal distribution. In the end, the bounds of information inequality are calculated under a general loss function.

Keywords: Minimax estimator, prior distribution, risk function, information inequality, orthogonal parameters.

[1] **Department of Statistics, Faculty of Mathematical Sciences, University of Kashan, Kashan, Iran**
[2] **Department of Statistics, Faculty of Mathematical Sciences, Payame Noor University, Tehran, Iran**