فاصلهٔ اطمینان بیز تجربی ناوردا برای بردار میانگین توزیع نرمال و تعمیم آن برای خانواده نمایی

مهدی شمس؛ دانشگاه کاشان، گروه آمار

دریافت ۹۳/۰۲/۰۶ پذیرش ۹۵/۰۹/۱۵

چکیده

بر اساس یک مدل بیزی نرمال چندمتغیره با واریانس معلوم می‌توان یک فاصلهٔ اطمینان برای مؤلفه‌های میانگین که دارای توزیع نرمال‌اند پیدا کرد. این فاصله اطمینان را می‌توان به‌صورت مشروط به یک آمارهٔ کمکی نیز پیدا کرد که در هر دو حالت یک فاصلهٔ اطمینان بیز تجربی ناوردا با سطح اطمینان داده شده به‌دست آورد. در پایان می‌توان مسئله را در حالت کلی‌تر برای خانوادهٔ نمایی تعمیم داد.

واژه‌های کلیدی: بیز تجربی، برآوردگر بیز، فاصله اطمینان ناوردا.

مقدمه

روش بیز تجربی[1] (EB) یکی از شاخه‌های مهم آمار به‌شمار می‌رود که به‌ویژه برای مسئلهٔ برآورد هم‌زمان چند پارامتر به‌کار می‌رود. در این روش پارامترها با روش‌های کلاسیک مثل روش گشتاوری یا روش حداکثر درستنمایی برآورد می‌شوند و در نتیجه توزیع پیشین برآورد شده به‌منظور استنباط به‌کار می‌رود [۶]. در روش EB دو مسیر تصادفی داده‌ها و پارامترها به‌هم مربوط می‌شود. برای این منظور فرض کنید دادهٔ X از خانوادهٔ پارامتری $P=\{f_\theta(x):\theta\in\Theta\}$ است و پارامترهای $\theta\in\Theta$ دارای توزیع پیشین $\pi\in\Pi$ هستند (Π خانواده توزیع‌ها روی Θ است). هدف استنباط، تشخیص مقادیر θ است. از این رو، روند استنباط بستگی به اطلاع داشتن از خانوادهٔ Π دارد [۱۱]. برای مثال، اگر دلیلی برای این اعتقاد که θ_iها مستقل و هم‌توزیع و از یک توزیع پیشین $\pi_0(\lambda)$ هستند وجود داشته باشد (وقتی که توزیع π_0 مگر احتمالاً برای برخی از پارامترهای نامعلوم λ، معلوم باشد)، در روش EB پارامتری، λ با استفاده از توزیع کناری مشاهدات برآورد می‌شود [۱]. این روش در شاخه‌های مختلفی خصوصاً تجارت، اقتصاد، بیمه و نقشه‌برداری کاربرد دارد. برای آشنایی با کاربردهای روش‌های بیز تجربی در مسائل استنباط آماری بررسی مراجع [۵]، [۶] و [۱۱] می‌تواند مفید باشد.

کاکس [۳] اولین بنیان‌گذار ساختن فاصلهٔ اطمینان EB در ساده‌ترین حالت برای توزیع نرمال، برای ساختن فاصلهٔ اطمینان EB، شرطی کردن روی برخی آماره‌ها را توصیه کرد. دیدگاه دیگر را موریس[۱۱] ارائه کرد که در آن فاصلهٔ اطمینان بیز سلسله مراتبی[2] (HB) برای پارمترها ساخته می‌شد و در آخرین مرحله پارمترهای توزیع پیشین برآورد می‌شد. با توجه به چوله‌بودن توزیع پسین، موریس [۱۱] صدک‌هایی را که برای ساختن فاصلهٔ اطمینان EB به‌کار گرفته می‌شود به‌صورت تقریبی محاسبه کرد. بعد از آن لایرد و لوئیس [۹] فاصلهٔ اطمینان EB بوت‌استرپ را پیشنهاد

*نویسندهٔ مسئول mehdishams@kashanu.ac.ir

1. Empirical Bayes
2. Hierarchical Bayes

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

کردند و کارلین و گلفند [۲] به‌دنبال پیشنهاد افرون [۴]، درجه‌بندی کردن فاصلهٔ اطمینان EB بر اساس اصلاح اریبی و هم‌چنین شرطی کردن آن روی برخی آماره‌ها (که لزوماً آماره کمکی نیستند) را ارائه کردند. هیل [۷]یک چارچوب کلی برای ساختن فاصلهٔ اطمینان EB بر پایهٔ شرطی کردن روی برخی آماره‌های کمکی ساخت و در حالتی که مشاهدات و توزیع پیشین نرمال هستند نشان داد این فاصله دقیقاً بر فاصله اطمینان HB منطبق است ولی او برای ساختن فاصلهٔ اطمینان EB مجانبی هیچ روشی پیشنهاد نداد. ایدهٔ اصلی این مقاله بر اساس ترکیب روش کاکس [۳] و هیل [۷] در حالتی که مشاهدات و توزیع پیشین نرمال هستند، است. این فواصل اطمینان EB که ناوردا نیز هستند در دو حالت شرطی و غیرشرطی ساخته می‌شوند. سپس در حالت کلی‌تر برای مدلی که مشاهدات و توزیع پیشین از خانواده نمایی هستند مسئله تعمیم داده می‌شود. در انتها فواصل اطمینان EB در دو حالت شرطی و غیرشرطی در دو حالت مختلف از طریق داده‌های شبیه‌سازی شده مقایسه می‌شوند.

مخاطرهٔ بیز برای برآوردگرهای بیز تجربی

در ابتدا با ذکر یک مثال برآوردگر میانگین مربعات خطا برای برآوردگرهای EB محاسبه می‌شود. سپس در حالت کلی‌تر با محاسبه مخاطرهٔ بیز برای دو تابع زیان معروف می‌توان برآوردگر با کم‌ترین مخاطرهٔ را پیدا کرد.

مثال ۱. فرض کنید $X_1|\theta_1,...,X_k|\theta_k$ که $k\geq 3$ متغیرهای تصادفی مستقل با توزیع $N(\theta_i,V)$ باشند وقتی که $0<V$ معلوم است. هم‌چنین فرض کنید توزیع پیشین θ_iها که از یک‌دیگر مستقل هستند $N(\mathbf{z_i^T b},A)$ باشد، که در آن $\mathbf{z_i}$ها بردارهای رگرسیونی ستونی معلوم و با بعد $0\leq p\leq k-3$ هستند، و $\mathbf{b}$ یک بردار معلوم $r\times 1$ است. اگر $\mathbf{Z^T=(z_1,...,z_k)}$ و $rank(\mathbf{Z})=p$، آن‌گاه بر اساس درست‌نمایی و توزیع پیشین داده شده در بالا، توزیع پسین θ_iها مستقل از هم با توزیع $N\left((1-B)X_i+B\mathbf{z_i^T b},V(1-B)\right)$ است که در آن $B=V/(V+A)$ ($0\leq B\leq 1$را ضریب انقباضی[1] گویند). از این رو میانگین‌های توزیع پسین یعنی $E(\theta_i|X_i)$، برآوردگرهای بیز هستند [۱۱]:

$$\hat{\theta}_i^{BA}=(1-B)X_i+B\mathbf{z_i^T b} \qquad ; i=1,...,k .$$

در یک روش EB، $\mathbf{b}$ و A نامعلوم‌اند و باید از روی توزیع کناری X_iها برآورد شوند. X_iها از هم مستقل و با توزیع $N(\mathbf{z_i^T b},V+A)$ هستند [۱۱]. با قرار دادن $\mathbf{X}=(X_1,...,X_k)^T$ براساس توزیع کناری $\mathbf{X}$، آماره بسنده کامل برای $(\mathbf{b},A)$ برابر $(\hat{\mathbf{b}},S^2)$ به‌دست می‌آید که در آن $\hat{\mathbf{b}}=\mathbf{(Z^TZ)^{-1}Z^TX}$، برآوردگر کم‌ترین مربعات (و یا MLE) برای $\mathbf{b}$ است و

$$S^2=\left\|\mathbf{X}-\mathbf{Z}\mathbf{b}\right\|^2=\sum_{i=1}^{k}(X_i-\mathbf{z_i^T}\hat{\mathbf{b}})^2 .$$

هم‌چنین براساس توزیع کناری $\mathbf{X}$، $\hat{\mathbf{b}}\perp S^2$ و $\hat{\mathbf{b}}$ دارای توزیع $N\left(\mathbf{b},(V+A)\mathbf{(Z^TZ)^{-1}}\right)$ و S^2 دارای توزیع $(V+A)\chi^2_{k-p}$ است. بنا براین $\mathbf{b}$ به‌وسیلهٔ $\hat{\mathbf{b}}$ برآورد می‌شود. MLE برای B با $\min(kV/S^2,1)$ داده می‌شود، درصورتی که با توجه به رابطهٔ $E\left(V(k-p-2)/S^2\right)=B$، UMVUE برای B به‌صورت $\hat{B}_{js}=V(k-p-2)/S^2$ است [۱۱]. از این‌رو، یک برآوردگر EB عمومی برای θ_i بدین‌صورت است [۶]:

1. Shrinking Coefficient

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

$$\hat{\theta}_i = \left[1-\hat{B}(S^2)\right]X_i + \hat{B}(S^2)\mathbf{z_i^T\hat{b}}.$$

هدف یافتن برآوردگر کم‌ترین مربعات خطا یعنی $MSE(\hat{\boldsymbol{\theta}}) = E\left[(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})^{\mathbf{T}}\right]$ برای $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1,...,\hat{\theta}_k)^T$ است که در آن امید ریاضی روی توزیع توأم $\mathbf{X}$ و $\boldsymbol{\theta}$ گرفته شده است. با قرار دادن $\hat{\boldsymbol{\theta}}^{BA} = (1-B)\mathbf{X}+B\mathbf{Zb}$ داریم:

$$\begin{aligned} E\left[(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})^{\mathbf{T}}\right] &= E\left[(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}}+\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}}+\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})^{\mathbf{T}}\right] \\ &= E\left[(\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})^{\mathbf{T}}\right] + E\left[(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})^{\mathbf{T}}\right] \\ &\quad + E\left[(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})(\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})^{\mathbf{T}}\right] + E\left[(\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})^{\mathbf{T}}\right]. \end{aligned}$$

اکنون قسمت‌های مختلف سمت راست تساوی بالا محاسبه می‌شوند. چون $(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})(\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})^{\mathbf{T}}$ اسکالر است با ترانهاده خود برابر است از این‌رو:

$$\begin{aligned} E\left[(\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})^{\mathbf{T}}\right] &= E\left[(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})(\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})^{\mathbf{T}}\right] \\ &= E\left[E(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})(\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})^{\mathbf{T}}\,|\mathbf{X}\right] \\ &= E\left[E(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}}|\mathbf{X})(\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})^{\mathbf{T}}\right] \\ &= E\left[\left(E(\boldsymbol{\theta}-E(\boldsymbol{\theta}|\mathbf{X})|\mathbf{X}\right)(\hat{\boldsymbol{\theta}}^{\mathbf{BA}}-\hat{\boldsymbol{\theta}})^{\mathbf{T}}\right] = 0. \end{aligned}$$

هم‌چنین:

$$\begin{aligned} E\left[(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})^{\mathbf{T}}\right] &= E\left[E(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^{\mathbf{BA}})^{\mathbf{T}}\,|\mathbf{X}\right] \\ &= E\left[Var(\boldsymbol{\theta}|\mathbf{X})\right] \\ &= E\left[V(1-B)\mathbf{I_k}\right] \\ &= V(1-B)\mathbf{I_k} \end{aligned}$$

و بالاخره داریم:

$$\begin{aligned} \hat{\boldsymbol{\theta}}^{BA}-\hat{\boldsymbol{\theta}} &= (1-B)\mathbf{X}+B\mathbf{Zb}-\left[1-\hat{B}(S^2)\right]\mathbf{X}-\hat{B}(S^2)\mathbf{Z\hat{b}} \\ &= \left(\hat{B}(S^2)-B\right)\mathbf{X}+B\mathbf{Zb}+B\mathbf{Z\hat{b}}-B\mathbf{Z\hat{b}}-\hat{B}(S^2)\mathbf{Z\hat{b}} \\ &= \left(\hat{B}(S^2)-B\right)\mathbf{X}+B\mathbf{Z(b-\hat{b})}\left(\hat{B}(S^2)-B\right)\mathbf{Z\hat{b}} \\ &= \left(\hat{B}(S^2)-B\right)\left(\mathbf{X-Z\hat{b}}\right)-B\mathbf{Z(\hat{b}-b)}. \end{aligned}$$

با توجه به استقلال $\mathbf{\hat{b}}$ و $\mathbf{X-Z\hat{b}}$ و این که X_iها دارای توزیع $N(\mathbf{z_i}^T\mathbf{b}, V+A)$ هستند و نیز $B=\frac{V}{V+A}$ داریم:

$$\begin{aligned} E(\mathbf{\hat{b}}) &= E((\mathbf{Z^TZ})^{-1}\mathbf{Z^TX}) = (\mathbf{Z^TZ})^{-1}\mathbf{Z^TZb} = \mathbf{b}, \\ Var(\mathbf{\hat{b}}) &= Var\left[(\mathbf{Z^TZ})^{-1}\mathbf{Z^TX}\right] \\ &= (\mathbf{Z^TZ})^{-1}\mathbf{Z^T}Var(\mathbf{X})\mathbf{Z}(\mathbf{Z^TZ})^{-1} \\ &= \mathbf{Z}^{-1}(\mathbf{Z^T})^{-1}\mathbf{Z^T}(V+A)\mathbf{Z}(\mathbf{Z^TZ})^{-1} \\ &= (V+A)\mathbf{Z}^{-1}\mathbf{Z}(\mathbf{Z^TZ})^{-1} \end{aligned}$$

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

$$= (V + A)(\mathbf{Z^T Z})^{-1}$$
$$= VB^{-1}(\mathbf{Z^T Z})^{-1}$$

و در پی آن

$$E\left[(\hat{\boldsymbol\theta}^{BA}-\hat{\boldsymbol\theta})(\hat{\boldsymbol\theta}^{BA}-\hat{\boldsymbol\theta})^{\mathbf T}\right]=E\left[\left(\left(\hat B(S^2)-B\right)(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})-B\mathbf{Z}(\hat{\mathbf b}-\mathbf b)\right)\left(\left(\hat B(S^2)-B\right)(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})-B\mathbf{Z}(\hat{\mathbf b}-\mathbf b)\right)^T\right]$$
$$=E\left[(\hat B(S^2)-B)^2(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})^{\mathbf T}\right]+0-0+B^2E\left(\mathbf{Z}(\hat{\mathbf b}-\mathbf b)(\hat{\mathbf b}-\mathbf b)^{\mathbf T}\mathbf{Z}^{\mathbf T}\right)$$
$$=E\left[(\hat B(S^2)-B)^2(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})^{\mathbf T}\right]+B^2\mathbf{Z}Var(\hat{\mathbf b})\mathbf{Z}^{\mathbf T}$$
$$=E\left[(\hat B(S^2)-B)^2(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})^{\mathbf T}\right]+B^2\mathbf{Z}VB^{-1}(\mathbf{Z^T Z})^{-1}\mathbf{Z}^{\mathbf T}$$
$$=VB\mathbf{Z}(\mathbf{Z^T Z})^{-1}\mathbf{Z}^{\mathbf T}.$$

با در نظر گرفتن محاسبات بالا ماتریس MSE بدین‌صورت در می‌آید:

$$E\left[(\hat{\boldsymbol\theta}-\boldsymbol\theta)(\hat{\boldsymbol\theta}-\boldsymbol\theta)^{\mathbf T}\right]=V(1-B)\mathbf{I_k}+VB\mathbf{Z}(\mathbf{Z^T Z})^{-1}\mathbf{Z}^{\mathbf T}+E\left[(\hat B(S^2)-B)^2(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})^{\mathbf T}\right].$$

باتوجه به این که $\mathbf X$ دارای توزیع $N\left(\mathbf{Zb},(V+A)\mathbf{I_k}\right)$ و همچنین $\mathbf{Z}\hat{\mathbf b}$ دارای توزیع $N\left(\mathbf{Zb},(V+A)\mathbf{Z}(\mathbf{Z^T Z})^{-1}\mathbf{Z}^T\right)$ است، بنا براین با توجه به خاصیت‌های توزیع نرمال $\mathbf{X}-\mathbf{Z}\hat{\mathbf b}$ دارای توزیع $N\left(\mathbf{0},(V+A)(\mathbf{I_k}-\mathbf{Z}(\mathbf{Z^T Z})^{-1}\mathbf{Z}^{\mathbf T})\right)$ است و از این رو،

$$(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})(V+A)^{-1}(\mathbf{I_k}-\mathbf{Z}(\mathbf{Z^T Z})^{-1}\mathbf{Z}^{\mathbf T})^{-1}(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})^{-1}$$

دارای توزیع χ_1^2 است. هم‌چنین چون $(V+A)^{-1}S^2$ دارای توزیع χ_{k-p}^2 است، از این رو، توزیع $\mathbf{U}=(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})^{\mathbf T}/S^2$ به پارامتر مجهول بستگی ندارد و بنا براین یک آمارهٔ کمکی است که با توجه به قضیهٔ باسو از S^2 (که تابعی از آماره بسنده کامل $(\hat{\mathbf b},S^2)$ است) مستقل است، یعنی $\mathbf U\perp S^2$. بنا براین:

$$E\left[\left(\hat B(S^2)-B\right)^2(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})^{\mathbf T}\right]=E\left[\left(\hat B(S^2)-B\right)^2S^2\mathbf U\right]=E\left[\left(\hat B(S^2)-B\right)^2S^2\right]E\left[\mathbf U\right]$$

که $E\left[\mathbf U\right]$ با توجه فرمول گشتاور نسبت‌ها بدین‌صورت محاسبه می‌شود:

$$E\left[\mathbf U\right]=\frac{E\left[(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})(\mathbf{X}-\mathbf{Z}\hat{\mathbf b})^{\mathbf T}\right]}{E(S^2)}=\frac{\mathbf{I_k}-\mathbf{Z}(\mathbf{Z^T Z})^{-1}\mathbf{Z}^{\mathbf T}}{k-p}$$

که با جای‌گذاری آن در عبارت قبل، ماتریس MSE بدین‌صورت ساده می‌شود:

$$E\left((\hat{\boldsymbol\theta}-\boldsymbol\theta)(\hat{\boldsymbol\theta}-\boldsymbol\theta)^{\mathbf T}\right)=V(1-B)\mathbf{I_k}+VB\mathbf{Z}(\mathbf{Z^T Z})^{-1}\mathbf{Z}^{\mathbf T}$$
$$+E\left[\left(\hat B(S^2)-B\right)^2S^2\right](k-p)^{-1}\left(\mathbf{I_k}-\mathbf{Z}(\mathbf{Z^T Z})^{-1}\mathbf{Z}^{\mathbf T}\right).$$

نکته ۱. با انتخاب $\hat B(S^2)=V(k-p-2)/S^2$ که $k>p+2$ اصلاح لیندلی[۱] [۱۰] از برآوردگر جیمز-اشتاین انقباضی[۲] [۸] عبارت $\mathrm E\left[\left(\hat B(S^2)-B\right)^2S^2\right]$ در بالا بدین‌صورت ساده می‌شود که درآن S^2 دارای توزیع $VB^{-1}\chi_{k-p}^2$ است:

1. Lindley
2. James – Stein Shrinker

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

$$
\begin{aligned}
E\left[\left(\hat{B}(S^2)-B\right)^2 S^2\right] &= E\left[\frac{\left(V(k-p-2)-BS^2\right)^2}{S^2}\right] \\
&= VBE\left[\frac{\left((k-p-2)-V^{-1}BS^2\right)^2}{V^{-1}BS^2}\right] \\
&= VBE\left[\frac{\left((k-p-2)-\chi^2_{k-p}\right)}{\chi^2_{k-p}}\right] \\
&= VB\left\{(k-p-2)^2 E\left[\frac{1}{\chi^2_{k-p}}\right]-2(k-p-2)+E\left[\chi^2_{k-p}\right]\right\} \\
&= VB\left\{(k-p-2)-2(k-p-2)+(k-p)\right\} = 2VB.
\end{aligned}
$$

از این رو، ماتریس MSE بدین‌صورت است:

$$
\begin{aligned}
E\left(\hat{\boldsymbol\theta}-\boldsymbol\theta)(\hat{\boldsymbol\theta}-\boldsymbol\theta)^{\mathbf T}\right) &= V(1-B)\mathbf{I_k} + VB\mathbf{Z}(\mathbf{Z^T Z})^{-1}\mathbf{Z^T} + 2(k-p)^{-1}VB\left(\mathbf{I}-\mathbf{Z}(\mathbf{Z^TZ})^{-1}\mathbf{Z^T}\right) \\
&= VB\left\{(B^{-1}-1)\mathbf{I_k} + \mathbf{Z}(\mathbf{Z^TZ})^{-1}\mathbf{Z^T} + 2(k-p)^{-1}\left(\mathbf{I_k}-\mathbf{Z}(\mathbf{Z^TZ})^{-1}\mathbf{Z^T}\right)\right\} \\
&= VB\left\{\left(B^{-1}-\frac{k-p-2}{k-p}\right)\mathbf{I_k} + \frac{(k-p-2)}{k-p}\mathbf{Z}(\mathbf{Z^TZ})^{-1}\mathbf{Z^T}\right\} \\
&= V\mathbf{I_k} - VB\frac{k-p-2}{k-p}\left(\mathbf{I_k}-\mathbf{Z}(\mathbf{Z^TZ})^{-1}\mathbf{Z^T}\right).
\end{aligned}
$$

نکته ۲. تحت یک چارچوب مشابه، هیل[1] [۷] مثال‌هایی برای پیدا کردن $E\left[(\hat{\boldsymbol\theta}-\boldsymbol\theta)(\hat{\boldsymbol\theta}-\boldsymbol\theta)^{\mathbf T}\middle| S^2\right]$ مطرح شده است به‌طوری‌که $S^2=\left\|\mathbf{X}-\mathbf{Z}\hat{\mathbf{b}}\right\|^2$ یک آمارهٔ کمکی بر پایهٔ توزیع کناری X است. تنها تغییری که در MSE به وجود می‌آید این است که در عبارت $(k-p)^{-1}\left(\mathbf{I_k}-\mathbf{Z}(\mathbf{Z^TZ})^{-1}\mathbf{Z^T}\right)$ امید غیرشرطی روی S^2 جای خود را به خود S^2 می‌دهد، یعنی:

$$
\begin{aligned}
E\left(\left((\hat{\boldsymbol\theta}-\boldsymbol\theta)(\hat{\boldsymbol\theta}-\boldsymbol\theta)^{\mathbf T}\middle|S^2\right)\right. &= V(1-B)\mathbf{I_k} + VB\mathbf{Z}(\mathbf{Z^TZ})^{-1}\mathbf{Z^T} \\
&+ \left(\hat{B}(S^2)-B\right)^2 S^2(k-p)^{-1}\left(\mathbf{I_k}-\mathbf{Z}(\mathbf{Z^TZ})^{-1}\mathbf{Z^T}\right).
\end{aligned}
$$

اکنون مخاطرهٔ بیز برآوردگرهای EB نسبت به توابع زیان $L_1(\boldsymbol\theta,\mathbf{a})=\left\|\boldsymbol\theta-\mathbf{a}\right\|^2$ و $L_2(\boldsymbol\theta,\mathbf{a})=(\mathbf{a}-\boldsymbol\theta)^T\mathbf{Q}(\mathbf{a}-\boldsymbol\theta)$ محاسبه می‌شوند که در آن $\mathbf{Q}$ یک ماتریس معین مثبت معلوم است. برای این منظور در ابتدا مثال ۲ بررسی می‌شود.

مثال ۲. فرض کنید $\mathbf{X}|\boldsymbol\theta$ دارای توزیع $N(\boldsymbol\theta,\mathbf{I_p})$ و $\boldsymbol\theta$ دارای توزیع $N_p(\boldsymbol\mu,A\mathbf{I_p})$ باشد، به‌طوری‌که $p\geq 3$، $\boldsymbol\mu$ معلوم و A نامعلوم است. در این صورت کناری $\mathbf{X}$ دارای توزیع $N(\boldsymbol\mu,B^{-1}\mathbf{I_p})$ و پسین $\boldsymbol\theta|\mathbf{X}=\mathbf{x}$ دارای توزیع $N\left((1-B)\mathbf{x}+B\boldsymbol\mu,(1-B)\mathbf{I_p}\right)$ است که در آن $B=(A+1)^{-1}$ ($0\leq B\leq 1$ ضریب انقباضی است). بنا براین برآوردگر بیز (میانگین پسین) به‌صورت $E(\boldsymbol\theta|\mathbf{X}=\mathbf{x})=(1-B)\mathbf{x}+B\boldsymbol\mu$ خواهد بود. در یک روند EB، برخی یا تمام پارامترهای توزیع پیشین نامعلوم‌اند و این پارامترها را از توزیع کناری $\mathbf{X}$ برآورد می‌شوند. با توجه به توزیع کناری $\mathbf{X}$، $\left\|\mathbf{X}-\boldsymbol\mu\right\|^2$ یک آمارهٔ بسنده کامل و دارای توزیع $B^{-1}\chi^2_p$ است. هم‌چنین برای $p\geq 3$،

1. Hill

$(p-2)\big/\|\mathbf{X}-\boldsymbol{\mu}\|^2$ یک UMVUE برای B است که با جای‌گذاری این برآوردگر از B در برآوردگر بیز یک برآوردگر EB از $\boldsymbol{\theta}$ بدین‌صورت به‌دست می‌آید:

$$\hat{\boldsymbol{\theta}}_{\mathbf{EB}}=(1-\frac{p-2}{\|\mathbf{X}-\boldsymbol{\mu}\|^2})\mathbf{X}+\frac{p-2}{\|\mathbf{X}-\boldsymbol{\mu}\|^2}\boldsymbol{\mu}=\mathbf{X}-\frac{p-2}{\|\mathbf{X}-\boldsymbol{\mu}\|^2}(\mathbf{X}-\boldsymbol{\mu})$$

که به برآوردگر جیمز-اشتاین معروف است [۱۱]. از مزیت‌های این برآوردگر این است که مخاطرهٔ آن تحت تابع زیان L_1 از مخاطرهٔ $\mathbf{X}$ کوچک‌تر است، ولی تحت تابع زیان L_2 لزوماً این مورد برقرار نیست [۲]. برای اثبات این حقیقت بدین‌صورت عمل می‌شود:

$$\begin{aligned}E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]&=E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{B}})\right]+E\left[(\hat{\boldsymbol{\theta}}_{\mathbf{B}}-\hat{\boldsymbol{\theta}}_{\mathbf{EB}})(\hat{\boldsymbol{\theta}}_{\mathbf{B}}-\hat{\boldsymbol{\theta}}_{\mathbf{EB}})^T\right]\\&=E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{B}})\right]+E\left[(B-\frac{p-2}{\|\mathbf{X}-\boldsymbol{\mu}\|^2})^2(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T\right]\end{aligned}$$

که درآن که E امید ریاضی روی توزیع توأم $\mathbf{X}$ و $\boldsymbol{\theta}$ است و $\hat{\boldsymbol{\theta}}_{\mathbf{B}}=(1-B)\mathbf{X}+B\boldsymbol{\mu}$. از این‌که $\|\mathbf{X}-\boldsymbol{\mu}\|^2$ آمارهٔ بسنده کامل و $\frac{(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T}{\|\mathbf{X}-\boldsymbol{\mu}\|^2}=\frac{B^{-1}(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T}{B^{-1}\|\mathbf{X}-\boldsymbol{\mu}\|^2}$ آماره کمکی است، قضیه باسو استقلال این دو آماره را تضمین می‌کند. بنا براین با توجه فرمول گشتاور نسبت‌ها و این‌که برای $p\geq 3$، $E\|\mathbf{X}-\boldsymbol{\mu}\|^2=B^1$ و $E\|\mathbf{X}-\boldsymbol{\mu}\|^{-2}=B(p-2)^{-1}$ داریم:

$$E\left[\frac{(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T}{\|\mathbf{X}-\boldsymbol{\mu}\|^2}\right]=\frac{E\left[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T\right]}{E\left(\|\mathbf{X}-\boldsymbol{\mu}\|^2\right)}=\frac{B^{-1}\mathbf{I}_{\mathbf{p}}}{B^{-1}p}=p^{-1}\mathbf{I}_{\mathbf{p}},$$

$$E\left[\frac{(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T}{\|\mathbf{X}-\boldsymbol{\mu}\|^4}\right]=E\left[\frac{(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T}{\|\mathbf{X}-\boldsymbol{\mu}\|^2}\right]E\left[\|\mathbf{X}-\boldsymbol{\mu}\|^{-2}\right]=\left(p^{-1}\mathbf{I}_{\mathbf{p}}\right)B(p-2)^{-1}=\frac{B\mathbf{I}_{\mathbf{p}}}{p(p-2)}$$

و در پی آن

$$\begin{aligned}E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]&=E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{B}})\right]+B^2E\left[(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T\right]-2B(p-2)E\left[\frac{(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T}{\|\mathbf{X}-\boldsymbol{\mu}\|^2}\right]+(p-2)^2E\left[\frac{(\mathbf{X}-\boldsymbol{\mu})(\mathbf{X}-\boldsymbol{\mu})^T}{\|\mathbf{X}-\boldsymbol{\mu}\|^4}\right]\\&=E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{B}})\right]+B^2(B^{-1}\mathbf{I}_{\mathbf{p}})-2B(p-2)(p^{-1}\mathbf{I}_{\mathbf{p}})+(p-2)^2\left(p^{-1}(p-2)^{-1}B\,\mathbf{I}_{\mathbf{p}}\right)\\&=E\left[L_1(\boldsymbol{\theta}+\hat{\boldsymbol{\theta}}_{\mathbf{B}})\right]+B\,\mathbf{I}_{\mathbf{p}}-B(p-2)p^{-1}\mathbf{I}_{\mathbf{p}}.\end{aligned}$$

از طرف دیگر از این‌که

$$E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{B}})\right]=E\left[L_1(\boldsymbol{\theta},E(\boldsymbol{\theta}|\mathbf{X}))\right]=E\left[Var(\boldsymbol{\theta}|\mathbf{X})\right]=E\left[(1-B)\mathbf{I}_{\mathbf{p}}\right]=(1-B)\mathbf{I}_{\mathbf{p}},$$

نتیجه می‌شود که

$$E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]=(1-B)\mathbf{I}_{\mathbf{p}}+B\mathbf{I}_{\mathbf{p}}-B(p-2)p^{-1}\mathbf{I}_{\mathbf{p}}=\mathbf{I}_{\mathbf{p}}-B(p-2)p^{-1}\mathbf{I}_{\mathbf{p}}.$$

ذکر این نکته ضروری است که با توجه به رابطهٔ بالا و از این که برای $p\geq 3$، $1-B(p-2)p^{-1}\leq 1$ و با توجه به این که

$$E\left[L_1(\boldsymbol{\theta},\mathbf{X})\right)=E\left[(\mathbf{X}-\boldsymbol{\theta})(\mathbf{X}-\boldsymbol{\theta})^T\right]=E\left[\mathbf{I}_{\mathbf{p}}\right]=\mathbf{I}_{\mathbf{p}}$$

می‌توان نتیجه گرفت که $E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]\le E\left[L_1(\boldsymbol{\theta},\mathbf{X})\right]$. این نامساوی بیان‌گر این است که مخاطرهٔ برآوردگر جیمز- اشتاین تحت SEL از مخاطرهٔ $\mathbf{X}$ کوچک‌تر است.

برای پیدا کردن مخاطرهٔ بیز $\hat{\boldsymbol{\theta}}_{\mathbf{EB}}$ تحت تابع زیان L_2 با توجه به این موضوع که

$$L_2(\boldsymbol{\theta},\mathbf{a})=(\boldsymbol{\theta}-\mathbf{a})^T\mathbf{Q}(\boldsymbol{\theta}-\mathbf{a})=tr\left[(\boldsymbol{\theta}-\mathbf{a})^T\mathbf{Q}(\boldsymbol{\theta}-\mathbf{a})\right]=tr\left[\mathbf{Q}(\boldsymbol{\theta}-\mathbf{a})^T(\boldsymbol{\theta}-\mathbf{a})\right]=tr\left[\mathbf{Q}L_1(\boldsymbol{\theta},\mathbf{a})\right]$$

داریم:

$$\begin{aligned}E\left[L_2(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]&=E\left[tr\left(\mathbf{Q}L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right)\right]\\&=E\left[tr(\mathbf{Q})L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]\\&=tr(\mathbf{Q})E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]\\&=tr(\mathbf{Q})\left[\mathbf{I_p}-B(p-2)p^{-1}\mathbf{I_p}\right]\\&=tr(\mathbf{Q})-B(p-2)p^{-1}tr(\mathbf{Q}).\end{aligned}$$

مثال ۳. اگر در مثال ۲ فرض کنید عناصر بردار میانگین $\boldsymbol{\mu}$ همگی یک‌سان و برابر μ باشند، یعنی $\boldsymbol{\theta}$ دارای توزیع $N(\mu\mathbf{1_p},A\mathbf{I_p})$ باشد به‌طوری‌که $\mu\in R$ و $A>0$ هر دو نامعلوم‌اند. توزیع کناری $\mathbf{X}$ به‌صورت $N(\mu\mathbf{1_p},B^{-1}\mathbf{I_p})$ است و از این رو، $\left(\overline{X},\sum_{i=1}^{p}(X_i-\overline{X})^2\right)$ یک آمارهٔ بسنده کامل است و بنا براین UMVUE برای μ و B به‌ترتیب برابر $\overline{X}$ و $(p-3)\Big/\sum_{i=1}^{p}(X_i-\overline{X})^2$ است. با جای‌گذاری این برآوردگرها در $E(\boldsymbol{\theta}|\mathbf{X}=\mathbf{x})=(1-B)\mathbf{x}+B\boldsymbol{\mu}$، برآوردگر EB برای $\boldsymbol{\theta}$ بدین‌صورت به‌دست می‌آید:

$$\hat{\boldsymbol{\theta}}_{\mathbf{EB}}=\left(1-\frac{p-3}{\sum_{i=1}^{p}(X_i-\overline{X})^2}\right)\mathbf{X}+\frac{p-3}{\sum_{i=1}^{p}(X_i-\overline{X})^2}\overline{X}\,\mathbf{1_p}=\mathbf{X}-\frac{p-3}{\sum_{i=1}^{p}(X_i-\overline{X})^2}(\mathbf{X}-\overline{X}\,\mathbf{1_p}).$$

این تعدیل برآوردگر جیمز- اشتاین را لیندلی [۶] مطرح کرده است[۱۰]. برآوردگر $\hat{\boldsymbol{\theta}}_{\mathbf{EB}}$ از این رو اهمیت دارد که برای $p\ge 4$ بر $\mathbf{X}$ غلبه[1] می‌کند، یعنی مخاطرهٔ بیز آن (تحت تابع زیان L_1) از مخاطرهٔ $\mathbf{X}$ کم‌تر است. برای اثبات این حقیقت مخاطرهٔ بیز $\hat{\boldsymbol{\theta}}_{\mathbf{EB}}$ تحت دو تابع زیان L_1 و L_2 بدین‌صورت به‌دست می‌آید:

$$\begin{aligned}E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]&=\mathbf{I_p}-B(p-3)(p-1)^{-1}(\mathbf{I_p}-p^{-1}\mathbf{J_p})\\E\left[L_2(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]&=tr(\mathbf{Q})-B(p-3)(p-1)^{-1}tr\left[\mathbf{Q}(\mathbf{I_p}-p^{-1}\mathbf{J_p}\right]\end{aligned}$$

به‌طوری‌که $\mathbf{J_p}=\mathbf{1_p}\mathbf{1_p}^T$. و هم‌چنین برای $p\ge 4$ و تحت تابع زیان L_1، $\hat{\boldsymbol{\theta}}_{\mathbf{EB}}$ بر $\mathbf{X}$ غلبه می‌کند، یعنی [۶]:

$$E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]\le E\left[L_1(\boldsymbol{\theta},\mathbf{X})\right].$$

بنا براین $E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]=E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]+E(\hat{\boldsymbol{\theta}}_{\mathbf{EB}}-\hat{\boldsymbol{\theta}}_{\mathbf{B}})(\hat{\boldsymbol{\theta}}_{\mathbf{EB}}-\hat{\boldsymbol{\theta}}_{\mathbf{B}})^T$ که در آن

$$\hat{\boldsymbol{\theta}}_{\mathbf{B}}=(1-B)\mathbf{X}+B\mu\mathbf{1_p}.$$

با توجه به این که:

1. Dominate

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

$$\hat{\boldsymbol{\theta}}_{\mathbf{EB}}-\hat{\boldsymbol{\theta}}_{\mathbf{B}}=\left(B-\frac{p-3}{\sum_{i=1}^{p}(X_i-\overline{X})^2}\right)(\mathbf{X}-\overline{X}\mathbf{1_p})+B(\overline{X}-\mu)\mathbf{1_p}$$

و هم‌چنین این که $\overline{X}$ دارای توزیع $N\left(\mu,(Bp)^{-1}\right)$ است و $\mathbf{X}-\overline{X}\mathbf{1_p}\perp\overline{X}$ (قضیهٔ باسو)، داریم:

$$E\left[\left(\hat{\boldsymbol{\theta}}_{\mathbf{EB}}-\hat{\boldsymbol{\theta}}_{\mathbf{B}}\right)\left(\hat{\boldsymbol{\theta}}_{\mathbf{EB}}-\hat{\boldsymbol{\theta}}_{\mathbf{B}}\right)^T\right]=E\left[\left(B-\frac{p-3}{\sum_{i=1}^{p}(X_i-\overline{X})}\right)^2(\mathbf{X}-\overline{X}\mathbf{1_p})(\mathbf{X}-\overline{X}\mathbf{1_p})^T\right]+B^2(Bp)^{-1}\mathbf{J_p}.$$

دوباره با استفاده از قضیهٔ باسو داریم:

$$\frac{(\mathbf{X}-\overline{X}\mathbf{1_p})(\mathbf{X}-\overline{X}\mathbf{1_p})^T}{\sum_{i=1}^{p}(X_i-\overline{X})^2}\perp\sum_{i=1}^{p}(X_i-\overline{X})^2$$

و نیز با توجه به این‌که $E\left[(\mathbf{X}-\overline{X}\mathbf{1_p})(\mathbf{X}-\overline{X}\mathbf{1_p})^T\right]=B^{-1}(\mathbf{I_p}-p^{-1}\mathbf{J_p})$ و چون $\sum_{i=1}^{p}(X_i-\overline{X})^2$ دارای توزیع $B^{-1}\chi^2_{p-1}$ است، از این رو، برای هر $p\geq 4$ داریم:

$$\begin{aligned}E\left[(\hat{\boldsymbol{\theta}}_{\mathbf{EB}}-\hat{\boldsymbol{\theta}}_{\mathbf{B}})(\hat{\boldsymbol{\theta}}_{\mathbf{EB}}-\hat{\boldsymbol{\theta}}_{\mathbf{B}})^T\right]&=B^2B^{-1}(\mathbf{I_p}-p^{-1}\mathbf{J_p})-2B(p-3)(p-1)^{-1}(\mathbf{I_p}-p^{-1}\mathbf{J_p})\\&\quad+(p-3)^2B(p-3)^{-1}(p-1)^{-1}(\mathbf{I_p}-p^{-1}\mathbf{J_p})+Bp^{-1}\mathbf{J_p}\\&=B\mathbf{I_p}B(p-3)(p-1)^{-1}(\mathbf{I_p}-p^{-1}\mathbf{J_p}).\end{aligned}$$

بنا براین مخاطرهٔ بیز $\hat{\boldsymbol{\theta}}_{\mathbf{EB}}$ بدین‌صورت ساده می‌شود:

$$E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]=E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{B}})\right]+B\mathbf{I_p}-B(p-3)(p-1)^{-1}(\mathbf{I_p}-P^{-1}\mathbf{J_p}).$$

از طرف دیگر چنان‌که در مثال ۲ به‌دست آوردیم $E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{B}})\right]=(1-B)\mathbf{I_p}$ که با جای‌گذاری آن در مخاطرهٔ بیز $\hat{\boldsymbol{\theta}}_{\mathbf{EB}}$داریم:

$$\begin{aligned}E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]&=(1-B)\mathbf{I_p}+B\mathbf{I_p}-B(p-3)(p-1)^{-1}(\mathbf{I_p}-p^{-1}\mathbf{J_p})\\&=\mathbf{I_p}-B(p-3)(p-1)^{-1}(\mathbf{I_p}-p^{-1}\mathbf{J_p}).\end{aligned}$$

مشابه مثال ۲ با توجه به‌این‌که $L_2(\boldsymbol{\theta},\mathbf{a})=tr\left(QL_1(\boldsymbol{\theta},\mathbf{a})\right)$، مخاطرهٔ بیز $\hat{\boldsymbol{\theta}}_{\mathbf{EB}}$ تحت تابع زیان L_2 نیز به‌دست می‌آید.

با توجه به این که برای $p\geq 4$، $-B(p-3)(p-1)^{-1})\leq 0$ و نیز $E\left[L_1(\boldsymbol{\theta},\mathbf{X})\right]=\mathbf{I_p}$ داریم:

$$E\left[L_1(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}_{\mathbf{EB}})\right]\leq E\left[L_1(\boldsymbol{\theta},\mathbf{X})\right].$$

بنا براین مخاطرهٔ برآوردگر تعدیلی جیمز- اشتاین(تحت SEL) از مخاطرهٔ $\mathbf{X}$ کوچک‌تر است.

فواصل اطمینان برپایهٔ احتمال‌های غیرشرطی

مدل بیزی را که در آن $\mathbf{Y}|\boldsymbol{\theta}$ دارای توزیع $N(\boldsymbol{\theta},\sigma^2I_m)$ و $\boldsymbol{\theta}$ دارای توزیع $N(\mu,\tau^2I_m)$ باشد در نظر بگیرید (σ^2 معلوم است و $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_m)^T$ و $\mathbf{Y}=(Y_1,\ldots,Y_m)^T$). بنابراین برآوردگر بیز (میانگین توزیع پسین) برای θ برابر است با $(1-B)\mathbf{Y}+B\boldsymbol{\mu}$ که درآن واریانس پسین وابسته برابر $\sigma^2(1-B)\mathbf{I_m}$ است وقتی

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

که $B=\frac{\sigma^2}{\sigma^2+\tau^2}$ و هم‌چنین متغیر تصادفی کناری $\mathbf{Y}$ دارای توزیع $N(\boldsymbol{\mu},\frac{\sigma^2}{B}I_m)$ است. حال اگر $\boldsymbol{\mu}$ نامعلوم باشد، آن‌گاه به‌طور معمول $\mathbf{Y}$ را به‌صورت برآوردگر $\boldsymbol{\mu}$ در نظر بگیرید که برآوردگر EB برای $\boldsymbol{\theta}$ نیز هست. فرض کنید $\boldsymbol{\mu}=\mu\mathbf{1_m}$ که در آن $\mathbf{1_m}$ یک بردار m مؤلفه‌ای با مختصات یک است. بنا براین در روش EB تحت فرضی که مدل با میانگین ثابت باشد، μ و τ^2 نامعلوم هستند و هر دو از کناری $\mathbf{Y}$ که دارای توزیع $N(\mu\mathbf{1_m},\frac{\sigma^2}{B}\mathbf{I_m})$ است، برآورد می‌شوند. بر پایهٔ این توزیع $(\bar{Y},S^2)$ یک آمارهٔ بسنده کامل برای (μ,τ^2) است که در آن $S^2=\frac{1}{m-1}\sum_{i=1}^m(Y_i-\bar{Y})^2$. مشابه مثال ۲ یک برآوردگر عمومی EB برای $\boldsymbol{\theta}$ به‌صورت $\hat{\theta}^{EB}=(1-\hat{B}(S))Y+\hat{B}(S)\bar{Y}1_m$ است که در آن $\hat{B}(S)$ یک برآوردگر برای B است [۵]. برآوردگر جیمز-اشتاین معمول برای B برابر است با $\hat{B}_{js}=\frac{(m-3)\sigma^2}{(m-1)S^2}$، در صورتی که قسمت مثبت برآوردگر انقباضی جیمز-اشتاین به‌صورت $\hat{B}(S)=\min\left(\frac{(m-3)\sigma^2}{(m-1)S^2},1\right)$ است. هدف تعیین یک فاصله اطمینان EB برای یکی از مؤلفه‌های $\boldsymbol{\theta}$ مثل θ_1 است. اگر μ و τ^2 معلوم باشند، $(1-B)Y_1+B\mu\pm z_{\alpha/2}\sigma(1-B)^{1/2}$ یک فاصله معتبر بیزی $100(1-\alpha)\%$ خواهد بود (که $z_{\alpha/2}$ نقطه‌ای است که در آن سطح زیر نمودار $N(0,1)$ از آن نقطه به بعد برابر $\alpha/2$ است.) در صورتی که اگر (μ,τ^2) معلوم نباشد، یک فاصلهٔ اطمینان برای θ_1 به‌صورت $\left(1-\hat{B}(S)\right)Y_1+\hat{B}(S)\bar{Y}\pm z_{\alpha/2}\sigma\left(1-\hat{B}(S)\right)^{1/2}$ است که در آن $0<\hat{B}(S)<1$ که این فاصله در $\bar{Y}$ تباهیده می‌شود اگر $\hat{B}(S)=1$. موریس [۱۱] فاصلهٔ اطمینان بالا را با تقریب برآوردگرهای HB برای میانگین و واریانس $\boldsymbol{\theta}$ بنا ساخت. لایرد و لوئیس [۹] فاصلهٔ اطمینان EB غیرشرطی بوت‌استرپ را بر اساس تقریب موریس پیشنهاد کردند در صورتی که کارلین و گلفند [۲] تقریب بوت‌استرپ برای فاصله اطمینان EB شرطی روی Y_1 را توصیه کردند. در این قسمت بر اساس چارچوب کاکس [۳] یک فاصلهٔ اطمینان EB بر اساس اصلاح اریبی ارائه می‌شود. برای این منظور یک درجه‌بندی برای فاصله اطمینان EB به‌منظور بهتر کردن احتمالات پوشش غیرشرطی مجانبی معرفی می‌شود.

قضیهٔ ۱. اگر Φ و ϕ به‌ترتیب تابع توزیع و تابع چگالی احتمال توزیع نرمال استاندارد باشند و برای $d>1$ که به m بستگی ندارد، $\hat{B}_d(S)=\min\left\{\frac{(m-d)\sigma^2}{(m-1)S^2},\frac{m-d}{m-1}\right\}$، آن‌گاه برای mهای بزرگ:

$$P\left[\theta_1\in\left(1-\hat{B}_d(S)\right)Y_1+\hat{B}_d(S)\bar{Y}\pm t\sigma\sqrt{1-\hat{B}_d(S)}\right]=2\Phi(t)-1-\frac{t\phi(t)}{m}\left[\frac{(6-d)B}{1-B}+\frac{(1+t^2)B^2}{2(1-B)^2}\right]+O(m^{-3/2}).$$

اثبات. با در نظر گرفتن آمارهٔ کمکی $A=\left(Y_1-\bar{Y}\right)/S$ و هم‌چنین این تعریف:

$$Z(t)=\frac{t\sigma\left(\sqrt{1-\hat{B}}-\sqrt{1-B}\right)+B(\bar{Y}-\mu)+(B-\hat{B})SA}{\sigma\sqrt{1-B}}$$

داریم:

$$P\left[\theta_1\le(1-\hat{B})Y_1+\hat{B}\bar{Y}+t\sigma\sqrt{1-\hat{B}}\right]=E\left\{P\left[\frac{\theta_1-\left((1-B)Y_1+B\mu\right)}{\sigma\sqrt{1-B}}\le t+Z(t)\middle|Y_1=y_1\right]\right\}=E\left[\Phi\left(t+Z(t)\right)\right] \qquad (۱)$$

که E، امید ریاضی روی توزیع کناری Y است. با بسط سری تیلور تابع Φ در نقطه $t+Z(t)$ داریم:

$$\Phi\left(t+Z(t)\right)=\Phi(t)+Z(t)\phi(t)+\tfrac{1}{2}Z^2(t)\phi'(t)+\tfrac{1}{2}\int_t^{t+Z(t)}\left(t+Z(t)-x\right)^2\phi''(x)dx. \qquad (۲)$$

از طرف دیگر با به‌کارگیری بسط تیلور داریم:

$$\begin{aligned} E\left(Z(t)\right)&=\frac{t}{\sqrt{1-B}}E\left[\sqrt{1-\hat{B}}-\sqrt{1-B}\right]\\ &=tE\left[\frac{B-\hat{B}}{2(1-B)}-\frac{(B-\hat{B})^2}{8(1-B)^2}+\frac{3}{16}\int_0^{\frac{B-\hat{B}}{1-B}}\left(\frac{B-\hat{B}}{1-B}-x\right)^2(1+x)^{-5/2}dx\right]. \end{aligned} \qquad (۳)$$

اکنون با قرار دادن $\tilde{B}_d\equiv\tilde{B}_d(S)=\frac{(m-d)\sigma^2}{(m-1)S^2}$ داریم:

$$\begin{aligned} E(B-\hat{B})&=E(B-\tilde{B}_d)+E(\tilde{B}_d-\hat{B})\\ &=B-B(m-d)E(\frac{\sigma^2}{(m-1)BS^2})+E(\tilde{B}_d-\hat{B})\\ &=\frac{d-3}{m}B+E(\tilde{B}_d-\hat{B})+O(m^{-2}) \end{aligned} \qquad (۴)$$

و به‌طور مشابه:

$$\begin{aligned} E(B-\hat{B})^2&=E(B-\tilde{B}_d)^2+E(\tilde{B}_d-\hat{B})^2+2E(B-\tilde{B}_d)(\tilde{B}_d-\hat{B})\\ &=\frac{B^2\left(2m+d^2+15\right)}{\left(m-3\right)\left(m-5\right)}+E(\tilde{B}_d-\hat{B})^2+2E(B-\tilde{B}_d)(\tilde{B}_d-\hat{B})\\ &=2B^2/m+E(\tilde{B}_d-\hat{B})^2+2E(B-\tilde{B}_d)(\tilde{B}_d-\hat{B})+O(m^{-2}). \end{aligned} \qquad (۵)$$

حال برای هر $1\le r$ ثابت و دل‌خواه داریم:

$$\begin{aligned} E\left|\tilde{B}_d-\hat{B}\right|^r&=E\left[\left|\frac{(m-d)\sigma^2}{(m-1)S^2}-\frac{m-d}{m-1}\right|^r I_{(\frac{1}{m-1},\infty)}(\frac{\sigma^2}{(m-1)S^2})\right]\\ &\le(m-d)^r\sqrt{E\left|\frac{B}{\chi^2_{m-1}}-\frac{1}{m-1}\right|^{2r}}\sqrt{P\left[\frac{\chi^2_{m-1}}{B}<m-1\right]}\\ &\le KP^{1/2}\left[\left|\chi^2_{m-1}-(m-1)\right|<-(m-1)(1-B)\right]\quad;\quad\left(K=(m-d)^rO(m^{-r})\right)\\ &\le O(1)P^{1/2}\left[\left|\chi^2_{m-1}-(m-1)\right|>(m-1)(1-B)\right]\\ &\le\frac{O(1)E^{1/2}\left|\chi^2_{m-1}-(m-1)\right|^{2u}}{\left[(m-1)(1-B)\right]^u}\\ &=O(1)O(m^{-u/2})\\ &=O(m^{-u/2}). \end{aligned}$$

عبارت بالا برای هر $0<u$ ثابت و دل‌خواه برقرار است. بنا براین برای هر $1\le r$ با انتخاب u به‌گونه‌ای که $r\le u$ داریم، $\max\left\{m^{-u/2},m^{-r/2}\right\}=m^{-r/2}$ و از این رو،

$$E\left|B-\hat{B}\right|^r\le2^{r-1}\left[E\left|B-\tilde{B}_d\right|^r+E\left|\tilde{B}_d-\hat{B}\right|^r\right]\le2^{r-1}\left[O(m^{-u/2})+O(m^{-r/2})\right]=O(m^{-r/2}) \qquad (۶)$$

و سرانجام با توجه به این‌که $I_{(-\infty,B]}(\hat{B})+I_{(B,\infty)}(\hat{B})=1$ داریم:

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

$$\left|\int_0^{\frac{B-\hat{B}}{1-B}} \left(\tfrac{B-\hat{B}}{1-B}-x\right)^2 (1+x)^{-5/2} dx\right| = I_1 + I_2 \qquad (۷)$$

که در آن

$$I_1 = \left\{\int_0^{\frac{B-\hat{B}}{1-B}} \left(\tfrac{B-\hat{B}}{1-B}-x\right)^2 (1+x)^{-5/2} dx\right\} I_{(-\infty,B]}(\hat{B})$$

و

$$I_2 = \left\{\int_{\frac{B-\hat{B}}{1-B}}^{0} \left(\tfrac{B-\hat{B}}{1-B}-x\right)^2 (1+x)^{-5/2} dx\right\} I_{(B,\infty)}(\hat{B}) .$$

اما با توجه به این که برای هر $a>0$،

$$\int_0^a (a-x)^2 (1+x)^{-5/2} dx \le a^2 \int_0^a (1+x)^{-5/2} dx = a^2\left[1-(1+a)^{-3/2}\right] \le a^3$$

بدین‌صورت داریم:

$$E(I_1) \le E\left[\left(\tfrac{B-\hat{B}}{1-B}\right)^2 \tfrac{B-\hat{B}}{1-B} I_{(-\infty,B]}(\hat{B})\right] \le \frac{E\left|B-\hat{B}\right|^3}{(1-B)^3} = O(m^{-3/2}). \qquad (۸)$$

هم‌چنین با انتخاب $\varepsilon_m = m^{-\alpha}$ که $0<\alpha<\frac{1}{2}$ و نیز با توجه به این که برای هر $a \in \left(\frac{-\varepsilon_m}{1-B}, 0\right)$ داریم:

$$\int_a^0 (a-x)^2 (1+x)^{-5/2} dx \le \left(1-\tfrac{\varepsilon_m}{1-B}\right)^{-5/2} \int_a^0 (a-x)^2 dx = \tfrac{(-a)^3}{3}\left(1-\tfrac{\varepsilon_m}{1-B}\right)^{-5/2}.$$

بنا براین:

$$\begin{aligned}
E\left\{\left[\int_{\frac{B-\hat{B}}{1-B}}^{0} \left(\tfrac{B-\hat{B}}{1-B}-x\right)^2 (1+x)^{-5/2} dx\right] I_{(B,B+\varepsilon_m)}(\hat{B})\right\} &\le E\left[\left(1-\tfrac{\varepsilon_m}{1-B}\right)^{-5/2}\left\{\int_{\frac{B-\hat{B}}{1-B}}^{0}\left(\tfrac{B-\hat{B}}{1-B}-x\right)^2 dx\right\} I_{(B,B+\varepsilon_m)}(\hat{B})\right] \\
&\le \tfrac{1}{3}\left(1-\tfrac{\varepsilon_m}{1-B}\right)^{-5/2} E\left[\left(\tfrac{\hat{B}-B}{1-B}\right)^3 I_{(B,B+\varepsilon_m)}(\hat{B})\right] \\
&\le \tfrac{1}{3}\left(1-\tfrac{\varepsilon_m}{1-B}\right)^{-5/2} \frac{E\left|\hat{B}-B\right|^3}{(1-B)^3} \\
&= O(m^{-3/2}).
\end{aligned} \qquad (۹)$$

علاوه بر آن:

$$\begin{aligned}
&E\left[\left\{\int_{\frac{B-\hat{B}}{1-B}}^{0} \left(\tfrac{B-\hat{B}}{1-B}-x\right)^2 (1+x)^{-5/2} dx\right\} I_{[B+\varepsilon_m,\infty)}(\hat{B})\right] \\
&E\left[\left\{\int_{\frac{B-\hat{B}}{1-B}}^{0} \left(1+x-\tfrac{1-\hat{B}}{1-B}\right)^2 (1+x)^{-5/2} dx\right\} I_{[B+\varepsilon_m,\infty)}(\hat{B})\right] \\
&\le E\left[\left\{\int_{\frac{B-\hat{B}}{1-B}}^{0} (1+x)^{-1/2} dx\right\} I_{[B+\varepsilon_m,\infty)}(\hat{B})\right] \\
&= E\left[2\left\{1-\sqrt{\tfrac{1-\hat{B}}{1-B}}\right\} I_{[B+\varepsilon_m,\infty)}(\hat{B})\right] \\
&\le 2P\left(\hat{B} > B+\varepsilon_m\right) \\
&\le 2P\left(\tfrac{(m-d)\sigma^2}{(m-1)S^2} > B+\varepsilon_m\right) \\
&= 2P\left[(m-1)BS^2/\sigma^2 < \tfrac{(m-d)B}{B+\varepsilon_m}\right].
\end{aligned}$$

اکنون با توجه به‌این‌که $(m-1)BS^2/\sigma^2$ دارای توزیع χ^2_{m-1} است، متغیرهای تصادفی مستقل و هم‌توزیع $Z_1,\ldots,Z_{m-1}$ با توزیع χ^2_1 وجود دارند به قسمی‌که $\left(=(m-1)\overline{Z}_{m-1}\right)$ $(m-1)BS^2/\sigma^2 \overset{d}{=} \sum_{i=1}^{m-1} Z_i$ و بنا براین:

$$
\begin{aligned}
E\left[\left\{\int_{\frac{B-\hat{B}}{1-B}}^{0} \left(\tfrac{B-\hat{B}}{1-B}-x\right)^2 (1+x)^{-5/2}dx\right\} I_{[B+\varepsilon_m,\infty)}(\hat{B})\right]
&= 2P\left(\sum_{i=1}^{m-1} Z_i < \tfrac{B(m-d)}{B+\varepsilon_m}\right)\\
&= 2P\left(\overline{Z}_{m-1}-1 < -\frac{\varepsilon_m + B\,\frac{d-1}{m-1}}{B+\varepsilon_m}\right)\\
&\le 2P\left(\left|\overline{Z}_{m-1}-1\right| > \frac{\varepsilon_m + B\,\frac{d-1}{m-1}}{B+\varepsilon_m}\right) \quad ;(d>1)\\
&\le 2P\left(\left|\overline{Z}_{m-1}-1\right| > \tfrac{\varepsilon_m}{B+\varepsilon_m}\right)\\
&\le 2\left(\tfrac{\varepsilon_m}{B+\varepsilon_m}\right)^{-2r} E\left|\overline{Z}_{m-1}-1\right|^{2r}\\
&= O\left(\varepsilon_m^{-2r} m^{-r}\right)\\
&= O\left(m^{-r(1-2\alpha)}\right)\\
&= O\left(m^{-3/2}\right)
\end{aligned}
$$

که در تساوی آخر $r=3/2(1-2\alpha)$، و بالاخره از رابطهٔ بالا و رابطهٔ (۹) داریم:

$$
\begin{aligned}
E(I_2) &= E\left[\left\{\int_{\frac{B-\hat{B}}{1-B}}^{0} \left(\tfrac{B-\hat{B}}{1-B}-x\right)^2 (1+x)^{-5/2}dx\right\} I_{(B,B+\varepsilon_m)}(\hat{B})\right]\\
&\quad + E\left[\left\{\int_{\frac{B-\hat{B}}{1-B}}^{0} \left(\tfrac{B-\hat{B}}{1-B}-x\right)^2 (1+x)^{-5/2}dx\right\} I_{[B+\varepsilon_m,\infty)}(\hat{B})\right]\\
&\le O(m^{-3/2}) + O(m^{-3/2})\\
&= O(m^{-3/2})
\end{aligned}
$$

که با استفاده از رابطهٔ بالا و روابط (۷) و (۸) نتیجه می‌شود:

$$
E\left[\int_0^{\frac{B-\hat{B}}{1-B}} \left(\tfrac{B-\hat{B}}{1-B}-x\right)^2 (1+x)^{-5/2}dx\right] = E(I_1)+E(I_2) \le O(m^{-3/2}).
$$

با توجه به‌این‌که $0<\alpha<\frac{1}{2}$، با جای‌گذاری عبارت بالا و روابط (۴) و (۵) در رابطه (۳)، $E(Z(t))$ بدین‌صورت ساده می‌شود:

$$
E(Z(t)) = t\left[\frac{(d-3)B}{2m(1-B)} - \frac{B^2}{4m(1-B)^2}\right] + O(m^{-3/2}). \qquad (۱۰)
$$

از این‌که $(\overline{Y},S^2)$ آماره بسندهٔ کامل برای (μ,τ^2) و $A=\left(Y_1-\overline{Y}\right)/S$ یک آمارهٔ کمکی است، با استفاده از قضیهٔ باسو استقلال دودویی $\overline{Y}$ و S و A نتیجه می‌شود و چون $E(A^2)=1$، می‌توان $E\left(Z^2(t)\right)$ را نیز بدین‌صورت نوشت:

$$
E\left(Z^2(t)\right) = \frac{1}{1-B}\left[t^2 E\left\{1-\hat{B}+1-B-2\sqrt{1-B}\sqrt{1-\hat{B}}\right\} + Bm^{-1} + E\frac{(B-\hat{B})^2 S^2}{\sigma^2}\right]. \qquad (۱۱)
$$

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

اما برای هر $u>0$ ثابت و دل‌خواه:

$$E(1-\hat{B})=E(1-\tilde{B}_d)+E(\tilde{B}_d-\hat{B})=1-B+\frac{d-3}{m-3}B+O(m^{-u})$$

و هم‌چنین:

$$E\left(\sqrt{1-\hat{B}}\right)=\sqrt{1-B}+\frac{(d-3)B}{2m\sqrt{1-B}}-\frac{B^2}{4m\sqrt{(1-B)^3}}+O(m^{-3/2})$$

و

$$E\left[(B-\hat{B})^2S^2\right]=E\left[\left\{(B-\tilde{B}_d)^2+(\tilde{B}_d-\hat{B})^2+2(B-\tilde{B}_d)(\tilde{B}_d-\hat{B})\right\}S^2\right]=2B\sigma^2/m+O(m^{-3/2})$$

پس با قرار دادن سه عبارت بالا در رابطهٔ (۱۱)، $E\left(Z^2(t)\right)$ به‌صورت (۱۲) ساده می‌شود:

$$E\left(Z^2(t)\right)=\frac{t^2B^2}{2m(1-B)^2}+\frac{3B}{m(1-B)}+O(m^{-3/2}). \qquad (۱۲)$$

هم‌چنین از این که $\left|(x^2-1)\phi(x)\right|<2\phi(\sqrt{3})$ و $\phi''(x)=(x^2-1)\phi(x)$ داریم:

$$\begin{aligned}E\left|\int_t^{t+Z(t)}\phi''(x)(t+Z(t)-x)^2dx\right|&\le E\left\{\left|Z(t)\right|^2\left|\int_t^{t+Z(t)}2\phi(\sqrt{3})dx\right|\right\}\\&=2E\left|Z(t)\right|^3\phi(\sqrt{3}) \qquad (۱۳)\\&=O(m^{-3/2}).\end{aligned}$$

بنا براین با استفاده از روابط (۱)، (۲)، (۱۰)، (۱۲) و (۱۳) داریم:

$$\begin{aligned}P\left[\theta_1\le(1-\hat{B})Y_1+\hat{B}\bar{Y}+t\sigma\sqrt{1-\hat{B}}\right]&=E\left[\Phi\left(t+Z(t)\right)\right]\\&=\Phi(t)+\frac{t\phi(t)}{m}\left[\frac{(d-3)B}{2(1-B)}-\frac{(1+t^2)B^2}{4(1-B)^2}-\frac{3B}{2(1-B)}\right]+O(m^{-3/2}).\end{aligned}$$

با کم کردن این عبارت از عبارت مشابه به‌ازای $-t$ حکم قضیه نتیجه می‌شود.

مثال ۴. مدل رگرسیون بیزی را که در آن $\mathbf{Y}|\boldsymbol{\theta}$ دارای توزیع $N(\boldsymbol{\theta},\sigma^2I_m)$ و $\boldsymbol{\theta}$ دارای توزیع $N(\mathbf{X}\boldsymbol{\beta},\tau^2I_m)$ باشد در نظر بگیرید که در آن $\mathbf{X}$ یک ماتریس $m\times p$ با رتبه $p(<m)$ و $\boldsymbol{\beta}$ بردار $p\times1$ از ضرایب رگرسیونی است. فرض کنید σ^2 معلوم است. در حالتی که $\boldsymbol{\beta}$ و τ^2 معلوم هستند، توزیع پسین $\boldsymbol{\theta}|\mathbf{Y}$ به‌صورت $N((1-B)\mathbf{Y}+B\mathbf{X}\boldsymbol{\beta},\sigma^2(1-B))$ خواهد بود که $B=\frac{\sigma^2}{\sigma^2+\tau^2}$. اما در حالتی که $\boldsymbol{\beta}$ و τ^2 مجهول هستند، برآوردگرهای $\boldsymbol{\beta}$ و B به‌ترتیب برابر با $\hat{\boldsymbol{\beta}}=\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}$ و $\hat{B}(S)=\hat{B}_d(S)=\min\left\{\frac{(m-d)\sigma^2}{(m-p)S^2},\frac{m-d}{m-p}\right\}$ هستند که در آن $S^2=\left\|\mathbf{Y}-\mathbf{X}\hat{\boldsymbol{\beta}}\right\|^2/(m-p)$ و $d>p$. بنا براین برآوردگر EB برای $\boldsymbol{\theta}$ برابر است با $(1-\hat{B})\mathbf{Y}+\hat{B}\mathbf{X}\hat{\boldsymbol{\beta}}$. فاصلهٔ اطمینان معمولی برای θ_1 به‌صورت $(1-\hat{B})Y_1+\hat{B}\mathbf{x}_1^T\hat{\boldsymbol{\beta}}\pm z_{\alpha/2}\sigma\sqrt{1-\hat{B}}$ است که در آن $\mathbf{x}_1^T$ اولین بردار ستونی $\mathbf{X}$ است. مطابق قضیهٔ ۱ با جای‌گزاری $\mathbf{x}_1^T(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$ به‌جای $\bar{Y}-\mu$ و تعریف $\hat{B}_{js}(S)=\frac{(m-p-2)\sigma^2}{(m-p)S^2}$ و هم‌چنین تحت شرط $\max_{1\le i\le m}\mathbf{x}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{x}_i=O(m^{-1})$ نتیجه می‌شود که:

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

$$P\left[\theta_1 \in \left(1-\hat{B}(S)\right)Y_1 + \hat{B}(S)\mathbf{x}_1^T\hat{\boldsymbol{\beta}} \pm t\sigma\sqrt{1-\hat{B}(S)}\right]$$

$$=2\Phi(t)-1-t\phi(t)\left[\frac{(1+t^2)B^2}{2m(1-B)^2}+\frac{B}{1-B}\left\{\mathbf{x}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{x}_i+\frac{5-d}{m}\right\}\right]+O(m^{-3/2}).$$

فواصل اطمینان EB برپایهٔ احتمال‌های شرطی

هیل [۷]، در حالتی که B معلوم است، که تنها پارامتر مجهول در توزیع پیشین $\boldsymbol{\mu}$ است فواصل اطمینان EB را محاسبه کرد. کارلین و گلفند [۲] در حالتی که τ^2 معلوم است روی آماره Y_1 شرطی کردند در صورتی که هیل [۷] در این حالت روی آماره کمکی $Y_1-\bar{Y}$ شرطی کرد. در این بخش در حالتی که τ^2 مجهول است، فاصله اطمینان EB بر اساس احتمال شرطی به شرط آماره کمکی $A=\left(Y_1-\bar{Y}\right)/S$ بررسی می‌شوند و قضیه زیر رهیافتی برای رسیدن به این منظور است.

قضیه ۲. باتوجه به شرایط قضیهٔ ۱ داریم:

$$.P\left[\theta_1 \in \left(1-\hat{B}(S)\right)Y_1+\hat{B}(S)\bar{Y} \pm t\sigma\sqrt{1-\hat{B}(S)}\,\middle|A\right]=2\Phi(t)-1-\frac{t\phi(t)}{m}\left[\frac{(2A^2+4-d)B}{1-B}+\frac{(1+t^2)B^2}{2(1-B)^2}\right]+O_p(m^{-3/2})$$

اثبات. برای نشان دادن این مطلب با توجه به این‌که $A=O_p(1)$، با روندی مشابه با قضیهٔ ۱ داریم:

$$\begin{aligned}P\left[\theta_1 \le (1-\hat{B})Y_1+\hat{B}\bar{Y}+t\sigma\sqrt{1-\hat{B}}\,\middle|A\right]&=\Phi(t)+\phi(t)E\left[Z(t)\middle|A\right]-\tfrac{1}{2}t\phi(t)E\left(Z^2(t)\right)A\right]\\&+\tfrac{1}{2}E\left[\int_t^{t+z(t)}\left(t+Z(t)-x\right)^2\phi''(x)dx\middle|A\right].\end{aligned} \tag{۱۴}$$

حال با توجه به قضیه باسو از استقلال $(\bar{Y},S)$ و A مشابه (۳) و (۱۰) داریم:

$$\begin{aligned}E\left(Z(t)\middle|A\right)&=tE\left[\frac{B-\hat{B}}{2(1-B)}-\frac{(B-\hat{B})^2}{8(1-B)^2}+\frac{3}{16}\int_0^{\frac{B-\hat{B}}{1-B}}\left\{\frac{B-\hat{B}}{1-B}-x\right\}^2(1+x)^{-5/2}dx\right]+E\left[\frac{(B-\hat{B})S}{\sigma\sqrt{1-B}}\right]A\\&=t\left[\frac{(d-3)B}{2m(1-B)}-\frac{B^2}{4m(1-B)^2}+O(m^{-3/2})\right]+E\left[\frac{(B-\hat{B})S}{\sigma\sqrt{1-B}}\right]A\end{aligned} \tag{۱۵}$$

و هم‌چنین با استفاده از استقلال دوبه‌دوی $\bar{Y}$ و S و A (قضیه باسو) و هم‌چنین مشابه رابطهٔ (۱۲) داریم:

$$E\left(Z^2(t)\middle|A\right)=\frac{t^2B^2}{2m(1-B)^2}+\frac{B+2BA^2}{m(1-B)}+\frac{t}{\sigma(1-B)}E\left[(B-\hat{B})^2S\right]A+O_p(m^{-3/2}) \tag{۱۶}$$

و سرانجام مثل قبل داریم:

$$\left|E\left[\int_t^{t+Z(t)}\left(t+Z(t)-x\right)\phi''(x)dx\middle|A\right]\right|\le 2\phi(\sqrt{3})E\left[\left|Z(t)\right|^3\middle|A\right] \tag{۱۷}$$

و نیز برای هر $t>0$، $E\left[\left|Z(t)\right|^3\middle|A\right]\le 9(1-B)^{-3/2}E\left[t^3\sqrt{1-\hat{B}}-\sqrt{1-B}\right]^3+\frac{B^3\left|\bar{Y}-\mu\right|^3}{\sigma^3}+\frac{(B-\hat{B})^3S^3|A|^3}{\sigma^3|A|}$ و هم‌چنین $E\left|\bar{Y}-\mu\right|^3=O_p(m^{-3/2})$ و نیز با کمک نامساوی (۶) برای $r=3$ و جای‌گذاری $(B-\hat{B})S$ به‌جای $(B-\hat{B})$ می‌توان نتیجه گرفت، $E\left[(B-\hat{B})^3S^3\right]\le 4E\left[\left|B-\tilde{B}_d\right|^3S^3+\left|\tilde{B}_d-\hat{B}\right|^3S^3\right]=O(m^{-3/2})$. با استفاده از بسط تیلور مشابه (۳) داریم:

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

$$E\left|\sqrt{1-\hat{B}}-\sqrt{1-B}\right|^3=E\left|\frac{B-\hat{B}}{2(1-B)}-\frac{(B-\hat{B})^2}{8(1-B)^2}+\frac{3}{16}\int_0^{\frac{B-\hat{B}}{1-B}}\left(\frac{B-\hat{B}}{1-B}-x\right)^2(1+x)^{-5/2}dx\right|^3=O(m^{-3/2}).$$

بنا براین با توجه به چهار رابطهٔ بالا به‌راحتی نتیجه می‌شود که:

$$E\left[\left|Z(t)\right|^3\left|A\right.\right]=9t^3(1-B)^{-3/2}O(m^{-3/2})+\frac{B^3}{\sigma^3}O_p(m^{-3/2})+\frac{|A|^2}{\sigma^2}O(m^{-3/2})=O_p(m^{-3/2}) \qquad (۱۸)$$

و در پی آن با استفاده از روابط (۱۴) تا (۱۸) داریم:

$$P\left[\theta_1\le(1-\hat{B})Y_1+\hat{B}\bar{Y}+t\sigma\sqrt{1-\hat{B}}\right]=E\left[\Phi\left(t+Z(t)\right)\right]$$
$$=\Phi(t)+\frac{t\phi(t)}{m}\left[\frac{(d-3)B}{2(1-B)}-\frac{(1+t^2)B^2}{4(1-B)^2}-\frac{3B}{2(1-B)}\right]+O(m^{-3/2}).$$

که با تفاضل این عبارت از عبارت مشابه به‌ازای $-t$ و استفاده از رابطهٔ (۱۵) مثل قضیه قبلی نتیجه مطلوب حاصل می‌شود.

مثال ۵. مدل رگرسیون بیزی مثال ۴ را در نظر بگیرید. تحت شرایط گفته شده در این مثال و با تعریف آماره کمکی $A=\left(Y_1-\mathbf{x}_1^T\hat{\boldsymbol{\beta}}\right)/S$ نتیجه می‌شود که:

$$P\left[\theta_1\in\left(1-\hat{B}(S)\right)Y_1+\hat{B}(S)\mathbf{x}_1^T\hat{\boldsymbol{\beta}}\pm t\sigma\sqrt{1-\hat{B}(S)}\left|A\right.\right]$$
$$=2\Phi(t)-1-t\phi(t)\left[\frac{(1+t^2)B^2}{2m(1-B)^2}+\frac{(2A^2+3-d)B}{m(1-B)}+\frac{B\mathbf{x}_i^T\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{x}_i}{1-B}\right]+O(m^{-3/2}).$$

فواصل اطمینان EB برای خانواده نمایی

در این بخش حالتی که مشاهدات و توزیع پیشین نرمال هستند به حالتی که مشاهدات و توزیع پیشین هر دو متعلق به یک خانواده نمایی عمومی هستند تعمیم داده می‌شود. فرض کنید $X_{ij}\left|\theta_i\right.$ها که $j=1,\ldots,n_i$ $i=1,\ldots,k$ متغیرهای تصادفی مستقل از خانواده نمایی با تابع چگالی $f\left(x\left|\theta_i\right.\right)=\exp\left\{\theta_i x-Q(\theta_i)+R(x)\right\}$ باشند. اگر $\bar{X}_i=\frac{1}{n_i}\sum_{j=1}^{n_i}X_{ij}$ و توزیع‌های پیشین مزدوج مستقل با تابع چگالی $\pi(\theta_i)\propto\exp\left\{\lambda\left[m_i\theta_i-Q(\theta_i)\right]\right\}$ باشند، آن‌گاه توزیع‌های پسین بدین‌صورت است:

$$\pi(\theta_i\left|\bar{X}_i\right.)\propto\exp\left\{\theta_i\left[n_i\bar{X}_i+\lambda m_i\right]-\left[n_i+\lambda\right]Q(\theta_i)\right\}.$$

میانگین پسین با استفاده از تساوی $\mu_i=Q'(\theta_i)$ به‌صورت $\hat{\mu}_{i,B}=E(\mu_i\left|\bar{X}_i\right.)=\left[n_i\bar{X}_i+\lambda m_i\right]/\left[n_i+\lambda\right]$ محاسبه می‌شود. اگر $m_i=h(\mathbf{x}_i^T\mathbf{b})$، $i=1,\ldots,k$ از دیدگاه EB باید پارامترهای $\mathbf{b}$ و λ براورد شوند. در این حالت $E\left(\bar{X}_i\right)=m_i$ و $\mathrm{var}\left(\bar{X}_i\right)=\mathrm{var}\left(\mu_i\right)+E\left(M(\mu_i)\right)=\frac{\lambda+n_i}{\lambda-a_2}M(m_i)$ که در آن $\sigma^2=M(\mu)=a_0+a_1\mu+a_2\mu^2$ و a_iها هم‌زمان صفر نیستند. اگر $\hat{\mathbf{b}}$ و $\hat{\lambda}$ براوردگرهای به‌ترتیب $\mathbf{b}$ و λ باشند، $\hat{\mu}_{i,EB}=\left[n_i\bar{X}_i+\hat{\lambda}\hat{m}_i\right]/\left[n_i+\hat{\lambda}\right]$ براوردگرهای EB برای μ_i هستند که $\hat{m}_i=h(\mathbf{x}_i^T\hat{\mathbf{b}})$ و برای k بزرگ و هر $\eta\ge 2$:

$$E\left(\left\|\hat{\mathbf{b}}-\mathbf{b}\right\|^\eta\right)=O_p(k^{-\eta/2}),\ E\left(\left\|\hat{\lambda}-\lambda\right\|^\eta\right)=O_p(k^{-\eta/2}). \qquad (۱۸)$$

[DOR: 20.1001.1.25882546.1396.3.1.7.5]
[DOI: 10.29252/mmr.3.1.57]

برای ساختن یک فاصلهٔ اطمینان EB برای μ_i، فرض می‌شود $n_i = O(k^{\frac{1}{3}-\delta})$ که $0<\delta<\frac{1}{3}$. در این صورت:

$$\begin{aligned}\left|\hat{\mu}_{i,EB}-\hat{\mu}_{i,B}\right| &= \left|\tfrac{n_i\bar{X}_i+\hat{\lambda}\hat{m}_i}{n_i+\hat{\lambda}}-\tfrac{n_i\bar{X}_i+\lambda m_i}{n_i+\lambda}\right| \\ &\le \left|n_i\bar{X}_i\left(\tfrac{1}{n_i+\hat{\lambda}}-\tfrac{1}{n_i+\lambda}\right)\right|+\left(1-\tfrac{n_i}{n_i+\hat{\lambda}}\right)\left|\hat{m}_i-m_i\right|+\left|\tfrac{n_i}{n_i+\hat{\lambda}}-\tfrac{n_i}{n_i+\lambda}\right|\left|m_i\right| \\ &\le \left|\bar{X}_i\right|\lambda^{-1}\left|\hat{\lambda}-\lambda\right|+\left|\hat{m}_i-m_i\right|+\left|m_i\right|\lambda^{-1}\left|\hat{\lambda}-\lambda\right|.\end{aligned}$$

با استفاده از بسط تیلور $h(\mathbf{x}_i^T\hat{\mathbf{b}})=h(\mathbf{x}_i^T\mathbf{b})+\mathbf{x}_i^T(\hat{\mathbf{b}}-\mathbf{b})h'(\mathbf{x}_i^T\mathbf{b})+\int_{\mathbf{x}_i^T\mathbf{b}}^{\mathbf{x}_i^T\hat{\mathbf{b}}}(\mathbf{x}_i^T\hat{\mathbf{b}}-\zeta)h''(\zeta)d\zeta$ و با کمک نامساوی شوارتز و رابطهٔ (۱۸) می‌توان مشاهده کرد که $E\left|\hat{m}_i-m_i\right|^{2r}=O(k^{-r})$. با به‌کارگیری نامساوی‌های شوارتز و مارکف:

(۱۹)

$$\begin{aligned}&P\left(n_i^{1/2}\left|\hat{\mu}_{i,EB}-\hat{\mu}_{i,B}\right|>Cn_i^{-1}(\log n_i)^{-1/2}\right)\\ &\le P\left(\left|\bar{X}_i\right|\lambda^{-1}\left|\hat{\lambda}-\lambda\right|>\tfrac{C}{3}n_i^{-3/2}(\log n_i)^{-1/2}\right)+P\left(\left|\hat{m}_i-m_i\right|>\tfrac{C}{3}n_i^{-3/2}(\log n_i)^{-3/2}\right)\\ &+P\left(\left|m_i\right|\lambda^{-1}\left|\hat{\lambda}-\lambda\right|>\tfrac{C}{3}n_i^{-3/2}(\log n_i)^{-1/2}\right)\\ &\le\left(\tfrac{C}{3}\right)^{-2r}n_i^{3r}(\log n_i)^r\left[\lambda^{2r}\sqrt{E\left|\bar{X}_i\right|^{4r}}\sqrt{E\left|\hat{\lambda}-\lambda\right|^{4r}}+E\left|\hat{m}_i-m_i\right|^{2r}+\left|m_i\right|^{2r}\lambda^{-2r}E\left|\hat{\lambda}-\lambda\right|^{2r}\right]\\ &=O_p[k^{r-3\delta r-r}(\log k)^r]=o_p(k^{-s})\end{aligned}$$

که نابرابری بالا برای هر $r>s/3\delta$ و s به‌اندازهٔ کافی بزرگ برقرار است. اکنون با استفاده از نابرابری مثلثی $\left|\hat{\mu}_{i,EB}-\hat{\mu}_i\right|\le\left|\hat{\mu}_{i,B}-\hat{\mu}_i\right|+\left|\hat{\mu}_{i,EB}-\hat{\mu}_{i,B}\right|$ به‌راحتی نتیجه گرفته می‌شود که:

$$\begin{aligned}&P\left(n_i^{1/2}\left|\hat{\mu}_{i,EB}-\mu_i\right|/\sqrt{M(\mu_i)}\le z_{\alpha/2}\right)\\ &\le P\left(n_i^{1/2}\left|\hat{\mu}_{i,B}-\mu_i\right|/\sqrt{M(\mu_i)}\le z_{\alpha/2}\right)+P\left(n_i^{1/2}\left|\hat{\mu}_{i,EB}-\hat{\mu}_{i,B}\right|>Cn_i^{-1}(\log n_i)^{-1/2}\right)\end{aligned}$$

و سپس با استفاده از (۱۹) و

$$P\left(n_i^{1/2}\left|\hat{\mu}_{i,B}-\mu_i\right|/\sqrt{M(\mu_i)}\le z_{\alpha/2}\right)=1-\alpha+O_p(n_i^{1/2})=1-\alpha+O_p(k^{-\frac{1}{3}+\delta}),$$

نتیجه می‌شود که

$$P\left(n_i^{1/2}\left|\hat{\mu}_{i,EB}-\mu_i\right|/\sqrt{M(\mu_i)}\le z_{\alpha/2}\right)\le 1-\alpha+O_p(k^{-\frac{1}{3}+\delta})$$

و از این رو فاصلهٔ اطمینان EB برای μ_i برای $i=1,\dots,k$ برابر است با:

$$\left(1-\tfrac{a_2z_{\alpha/2}^2}{n_i}\right)^{-1}\left[\hat{\mu}_{i,EB}+\tfrac{a_1z_{\alpha/2}^2}{2n_i}\pm\tfrac{z_{\alpha/2}}{n_i^{1/2}}\sqrt{M\left(\hat{\mu}_{i,EB}\right)+\tfrac{dz_{\alpha/2}^2}{4n_i}}\right].$$

شبیه‌سازی و نتیجه‌گیری

برای به‌دست آوردن فاصلهٔ اطمینان بیز تجربی ناوردای $100(1-\alpha)\%$ برای مؤلفه‌های میانگین که دارای توزیع نرمال‌اند مثل θ_1 در یک مدل بیزی نرمال چندمتغیره با واریانس معلوم با قراردادن $\hat{B}\equiv\hat{B}_d(S)$ و اختیار

$$t \equiv t(\hat{B},\alpha)=z_{\alpha/2}\left[1+\frac{(1+z_{\alpha/2}^2)\hat{B}^2}{4m(1-\hat{B})^2}+\frac{(6-d)\hat{B}}{2m(1-\hat{B})}\right]$$

در قضیهٔ ۱ داریم:

$$P\left[\theta_1 \in (1-\hat{B})Y_1+\hat{B}\overline{Y} \pm t(\hat{B},\alpha)\sigma\sqrt{(1-\hat{B})}\right]=1-\alpha+O(m^{-3/2}). \qquad (۲۰)$$

برای اثبات ناوردایی این فاصلهٔ اطمینان تحت گروه مکانی

$$G=\{g_c : g_c(y_1,...,y_n)=(y_1+c,...,y_n+c), c\in R\}$$

مشاهده می‌شود که اگر $C(\mathbf{Y})=(1-\hat{B})Y_1+\hat{B}\overline{Y} \pm t(\hat{B},\alpha)\sigma\sqrt{(1-\hat{B})}$:

$$C[g_c(\mathbf{Y})]=(1-\hat{B})Y_1+c-c\hat{B}+\hat{B}\overline{Y}+c\hat{B} \pm t(\hat{B},\alpha)\sigma\sqrt{(1-\hat{B})}=\tilde{g}_c C[\mathbf{Y}]$$

که در آن $\tilde{g}_c[(a,b)]=(a+c,b+c)$.

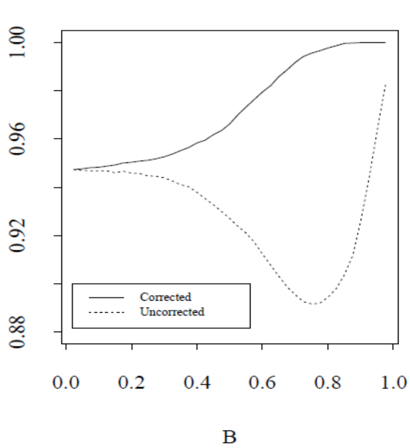
به‌طور مشابه با قرار دادن $\hat{B}=\hat{B}_d(S)$ و اختیار کردن $t \equiv t(\hat{B},A,\alpha)=z_{\alpha/2}\left[l+\frac{(1+z_{\alpha/2}^2)\hat{B}^2}{4m(1-\hat{B})^2}+\frac{(2A^2+4-d)\hat{B}}{2m(1-\hat{B})}\right]$

در قضیهٔ ۲ یک فاصله اطمینان بیز تجربی ناوردای شرطی $100(1-\alpha)\%$ برای θ_1 به‌صورت:

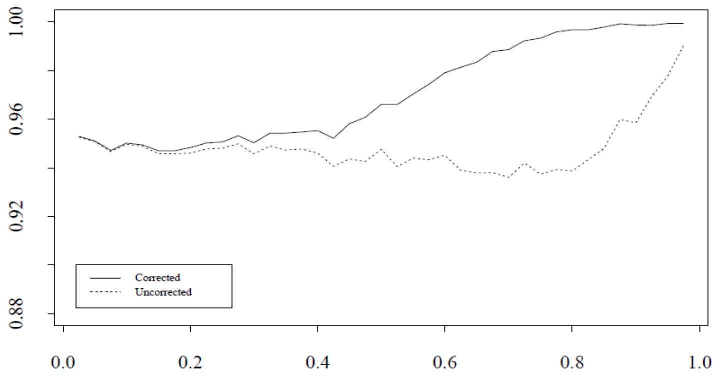
$$P\left[\theta_1 \in \left(1-\hat{B}\right)Y_1+\hat{B}\overline{Y} \pm t(\hat{B},A,\alpha)\sigma\sqrt{1-\hat{B}}\right)\Big| A\right]=1-\alpha+O_p(m^{-3/2}) \qquad (۲۱)$$

به‌دست می‌آید که به‌طور مشابه فاصلهٔ اطمینان بالا تحت گروه مکانی G ناورداست. هم‌چنین مشاهده می‌شود که فواصل اطمینان تصحیح‌نشده در قضیه‌های ۱ و ۲ خاصیت ناوردایی تحت گروه مکانی را ندارند ولی با توجه به روابط (۲۰) و (۲۱) فواصل اطمینان تصحیح‌شده تحت گروه مکانی ناوردا هستند.

برای مشاهده نتایج قضیه‌های ۱ و ۲ با استفاده از شبیه‌سازی در حالت $\mu=0$ و $\sigma^2=1$ و برای مقادیر مختلف B از 0/025 تا 0/975، تعداد ۱۰۰۰۰ زوج داده‌های مستقل (Y_i,θ_i) برای $i=1,...,30$ از توزیع نرمال شبیه‌سازی می‌شود. برای قضیهٔ ۲، ۱۰۰۰۰ نمونه شبیه‌سازی شده را به شرط آمارهٔ کمکی A با مقدار ۱ تولید می‌شود. در شکل ۱ و ۲ به‌ترتیب فاصله اطمینان بیز تجربی شرطی و غیرشرطی $100(1-\alpha)\%$ برای θ_1، در حالت تصحیح‌نشده یعنی استفاده از قضیه‌های ۱ و ۲ (نمودار Uncorrected با خطوط نقطه‌چین) و تصحیح شده یعنی روابط (۲۱) و (۲۱) (نمودار Corrected با خطوط توپر) با $t=z_{\alpha/2}$ شبیه‌سازی شده‌است.



شکل۱. فاصلهٔ اطمینان بیز تجربی شرطی برای θ_1



شکل۲. فاصله اطمینان بیز تجربی غیرشرطی برای θ_1

منابع

1. Berger J. O., "Statistical Decision Theory and Bayesian Analysis", 2nd Edition, Verlag, New York, Springer (1985).
2. Carlin B., Gelfand A., "Approaches for empirical Bayes confidence intervals", J. Amer. Statist, Assoc. 85 (1990) 105-114.
3. Cox D. R., "Prediction intervals and empirical Bayes confidence intervals", In Perspectives in probability and statistics, papers in honor of M.S. Bartlett (ed. J. Gani) Academic Press, NewYork (1975) 47-55.
4. Efron B., "Better bootstrap confidence intervals", J. Amer. Statist, Assoc. 82 (1987) 171-185.
5. Efron B., Morris C., "Stein’s Estimation Rule and it’s Competitorsn Empirical Bayes Approach", J. Amer. Statist. Assoc., 68 (1973) 117-130.
6. Ghosh M., "Hierarchical and Empirical Bayes Multivariate Estimation, Current Issues in Statistical Inference: Essays in Honor of D. Basu", Ghosh, M. and Pathak, P.K. eds., Institute of Mathematical Statistics(IMS) Lecture Notes and Monograph, 17 (1992) 151-177.
7. Hill J. R., "A General Framework for Model-Based Statistics", Biometrika, 77 (1990) 115-126.
8. James W., Stein C. "Estimation with Quadratic Loss, In Proceedings of the Fourth Berkeley Symposium Mathematical Statistics and Probability", (ed. Neyman, J.), University of California Press, 1 (1961) 361-379.
9. Laird N. M., Louis T. A., "Empirical Bayes confidence intervals based on bootstrap samples", J. Amer. Statist. Assoc. 82 (1987) 739-757.
10. Lindley D. V., "Discussion of the paper by Stein", J. Roy. Statist. Soc. Ser. B, 24 (1962) 265-296.
11. Morris C. N., "Parametric Empirical Bayes Inference: Theory and Applications", J. Amer. Statist. Assoc., 78 (1983) 47-65.