

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

مهدی شمس؛ دانشگاه کاشان، گروه آمار

پذیرش ۹۵/۰۹/۱۵

دریافت ۹۳/۰۲/۰۶

چکیده

بر اساس یک مدل بیزی نرمال چندمتغیره با واریانس معلوم می‌توان یک فاصله اطمینان برای مؤلفه‌های میانگین که دارای توزیع نرمال‌اند پیدا کرد. این فاصله اطمینان را می‌توان به صورت مشروط به یک آماره کمکی نیز پیدا کرد که در هر دو حالت یک فاصله اطمینان بیز تجربی ناوردا با سطح اطمینان داده شده به دست آورد. در پایان می‌توان مسئله را در حالت کلی تر برای خانواده نمایی تعمیم داد.

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

مقدمه

روش بیز تجربی^۱ (EB) یکی از شاخه‌های مهم آمار به شمار می‌رود که به‌ویژه برای مسئله برآورد هم‌زمان چند پارامتر به کار می‌رود. در این روش پارامترها با روش‌های کلاسیک مثل روش گشتاوری یا روش حداکثر درست‌نمایی برآورد می‌شوند و در نتیجه توزیع پیشین برآورد شده به منظور استنباط به کار می‌رود [۶]. در روش EB دو مسیر تصادفی داده‌ها و پارامترها به هم مربوط می‌شود. برای این منظور فرض کنید داده X از خانواده 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 خانواده Π دارد [۱۱]. برای مثال، اگر دلیلی برای این اعتقاد که θ_i 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 فاصله اطمینان EB در ساده‌ترین حالت برای توزیع نرمال، برای ساختن فاصله اطمینان EB، شرطی کردن روی برخی آماره‌ها را توصیه کرد. دیدگاه دیگر را موریس [۱۱] ارائه کرد که در آن فاصله اطمینان بیز سلسله مراتبی^۲ (HB) برای پارامترها ساخته می‌شد و در آخرین مرحله پارامترهای توزیع پیشین برآورد می‌شد. با توجه به چوله‌بودن توزیع پسین، موریس [۱۱] صدک‌هایی را که برای ساختن فاصله اطمینان EB به کار گرفته می‌شود به صورت تقریبی محاسبه کرد. بعد از آن لایرد و لوئیس [۹] فاصله اطمینان EB بوت‌استرپ را پیشنهاد

*نویسنده مسئول mehdishams@kashanu.ac.ir

1. Empirical Bayes
2. Hierarchical Bayes

کردند و کارلین و گلفند [۲] به دنبال پیشنهاد افرون [۴]، درجه‌بندی کردن فاصله 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 فاصله اطمینان EB بر پایه 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 فاصله اطمینان EB مجانبی هیچ روشی پیشنهاد نداد. ایده 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.

مخاطره بیز برای برآوردهای بیز تجربی

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

مثال ۱. فرض کنید $X_1|\theta_1, \dots, X_k|\theta_k$ 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_i ها که از یک‌دیگر مستقل هستند $N(\mathbf{z}_i^T \mathbf{b}, A)$ 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 $\mathbf{Z}^T = (\mathbf{z}_1^T, \dots, \mathbf{z}_k^T)$ و $\text{rank}(\mathbf{Z}) = p$ ، آن‌گاه بر اساس درست‌نمایی و توزیع پیشین داده شده در بالا، توزیع پسین θ_i ها مستقل از هم با توزیع $N((1-B)X_i + B\mathbf{z}_i^T \mathbf{b}, V(1-B))$ است که در آن $B = V/(V+A)$ است $(0 \leq B \leq 1)$ را ضریب انقباضی^۱ گویند. از این رو میانگین‌های توزیع پسین یعنی $E(\theta_i | X_i)$ ، برآوردهای بیز هستند [۱۱]:

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

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

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

همچنین براساس توزیع کناری $\mathbf{X}$ ، $\hat{\mathbf{b}} \perp S^2$ و $\hat{\mathbf{b}}$ دارای توزیع $N(\mathbf{b}, (V+A)(\mathbf{Z}^T \mathbf{Z})^{-1})$ و S^2 دارای توزیع $\chi^2_{k-p}(V+A)$ است. بنا براین $\mathbf{b}$ به وسیله $\hat{\mathbf{b}}$ برآورد می‌شود. MLE برای B با $\min(kV/S^2, 1)$ داده می‌شود، در صورتی که با توجه به رابطه $E(V(k-p-2)/S^2) = B$ ، UMVUE برای B به صورت $\hat{B}_{js} = V(k-p-2)/S^2$ است [۱۱]. از این رو، یک برآوردهای EB عمومی برای θ_i بدین صورت است [۶]:

1. Shrinking Coefficient

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

هدف یافتن برآوردگر کمترین مربعات خطا یعنی $MSE(\hat{\boldsymbol{\theta}}) = E[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T]$ برای $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_k)^T$

است که در آن امید ریاضی روی توزیع توأم $\mathbf{X}$ و $\boldsymbol{\theta}$ گرفته شده است. با قرار دادن $\hat{\boldsymbol{\theta}}^{BA} = (1-B)\mathbf{X} + B\mathbf{Z}\mathbf{b}$ داریم:

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

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

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

هم‌چنین:

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

و بالاخره داریم:

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

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

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

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

و در پی آن

$$\begin{aligned} E[(\hat{\theta}^{BA} - \theta)(\hat{\theta}^{BA} - \theta)^T] &= E\left[\left((\hat{B}(S^2) - B)(X - Z\hat{b}) - BZ(\hat{b} - b)\right)\left((\hat{B}(S^2) - B)(X - Z\hat{b}) - BZ(\hat{b} - b)\right)^T\right] \\ &= E\left[(\hat{B}(S^2) - B)^2(X - Z\hat{b})(X - Z\hat{b})^T\right] + 0 - 0 + B^2 E(Z(\hat{b} - b)(\hat{b} - b)^T Z^T) \\ &= E\left[(\hat{B}(S^2) - B)^2(X - Z\hat{b})(X - Z\hat{b})^T\right] + B^2 Z \text{Var}(\hat{b}) Z^T \\ &= E\left[(\hat{B}(S^2) - B)^2(X - Z\hat{b})(X - Z\hat{b})^T\right] + B^2 ZVB^{-1}(Z^T Z)^{-1} Z^T \\ &= VBZ(Z^T Z)^{-1} Z^T. \end{aligned}$$

با در نظر گرفتن محاسبات بالا ماتریس MSE بدین صورت در می‌آید:

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

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

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

دارای توزیع χ^2_1 است. همچنین چون $(V + A)^{-1}S^2$ دارای توزیع χ^2_{k-p} است، از این رو، توزیع $U = (X - Z\hat{b})(X - Z\hat{b})^T / S^2$ به پارامتر مجهول بستگی ندارد و بنا براین یک آماره کمکی است که با توجه به قضیهٔ باسو از S^2 (که تابعی از آماره بسنده کامل $(\hat{b}, S^2)$ است) مستقل است، یعنی $U \perp S^2$. بنا براین:

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

که $E[U]$ با توجه فرمول گشتاور نسبت‌ها بدین صورت محاسبه می‌شود:

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

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

$$\begin{aligned} E[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] &= V(1 - B)I_k + VBZ(Z^T Z)^{-1} Z^T \\ &\quad + E\left[(\hat{B}(S^2) - B)^2 S^2\right](k - p)^{-1}(I_k - Z(Z^T Z)^{-1} Z^T). \end{aligned}$$

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 $E[(\hat{B}(S^2) - B)^2 S^2]$ در بالا بدین صورت ساده می‌شود که در آن S^2 دارای توزیع $VB^{-1}\chi^2_{k-p}$ است:

1. Lindley
2. James – Stein Shrinker

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

از این رو، ماتریس MSE بدین صورت است:

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

نکته ۲. تحت یک چارچوب مشابه، هیل^۱ [۷] مثال‌هایی برای پیدا کردن $E[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T | S^2]$ مطرح شده است به‌طوری‌که $S^2 = \|\mathbf{X} - Z\hat{\mathbf{b}}\|^2$ یک آماره کمکی بر پایه توزیع کناری X است. تنها تغییری که در MSE به وجود می‌آید این است که در عبارت $(k-p)^{-1}(\mathbf{I}_k - Z(Z^T Z)^{-1}Z^T)$ امید غیرشرطی روی S^2 جای خود را به خود S^2 می‌دهد، یعنی:

$$\begin{aligned}
 E((\hat{\theta} - \theta)(\hat{\theta} - \theta)^T | S^2) &= V(1-B)\mathbf{I}_k + VBZ(Z^T Z)^{-1}Z^T \\
 &\quad + (\hat{B}(S^2) - B)^2 S^2 (k-p)^{-1} (\mathbf{I}_k - Z(Z^T Z)^{-1}Z^T).
 \end{aligned}$$

اکنون مخاطره بیز برآوردگرهای EB نسبت به توابع زیان $L_1(\theta, \mathbf{a}) = \|\theta - \mathbf{a}\|^2$ و $L_2(\theta, \mathbf{a}) = (\mathbf{a} - \theta)^T \mathbf{Q}(\mathbf{a} - \theta)$ 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 $\mathbf{X}|\theta$ دارای توزیع $N(\theta, \mathbf{I}_p)$ و θ دارای توزیع $N(\mu, A\mathbf{I}_p)$ باشد، به‌طوری‌که $p \geq 3$ ، μ معلوم و A نامعلوم است. در این صورت کناری $\mathbf{X}$ دارای توزیع $N(\mu, B^{-1}\mathbf{I}_p)$ و پسین $\theta|\mathbf{X} = \mathbf{x}$ دارای توزیع $N((1-B)\mathbf{x} + B\mu, (1-B)\mathbf{I}_p)$ است که در آن $B = (A+1)^{-1}$ $0 \leq B \leq 1$ ضریب انقباضی است. بنا براین برآوردگر بیز (میانگین پسین) به‌صورت $E(\theta|\mathbf{X} = \mathbf{x}) = (1-B)\mathbf{x} + B\mu$ 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 $\mathbf{X}$ ، $\|\mathbf{X} - \mu\|^2$ یک آماره بسنده کامل و دارای توزیع $B^{-1}\chi_p^2$ است. همچنین برای $p \geq 3$

برآوردگر EB از θ بدین صورت به دست می‌آید:

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

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

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

که در آن که E امید ریاضی روی توزیع توأم $\mathbf{X}$ و θ است و $\hat{\theta}_B = (1-B)\mathbf{X} + B\boldsymbol{\mu}$ از این‌که $\|\mathbf{X} - \boldsymbol{\mu}\|^2$ آماره 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:

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

و در پی آن

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

از طرف دیگر از این‌که

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

نتیجه می‌شود که

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

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

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

می توان نتیجه گرفت که $E[L_1(\theta, \hat{\theta}_{EB})] \leq E[L_1(\theta, X)]$. این نامساوی بیان گر این است که مخاطره برآوردگر جیمز- اشتاین تحت SEL از مخاطره X کوچک تر است.

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

$$L_2(\theta, a) = (\theta - a)^T Q(\theta - a) = tr[(\theta - a)^T Q(\theta - a)] = tr[Q(\theta - a)^T (\theta - a)] = tr[QL_1(\theta, a)]$$

داریم:

$$\begin{aligned} E[L_2(\theta, \hat{\theta}_{EB})] &= E[tr(QL_1(\theta, \hat{\theta}_{EB}))] \\ &= E[tr(Q)L_1(\theta, \hat{\theta}_{EB})] \\ &= tr(Q)E[L_1(\theta, \hat{\theta}_{EB})] \\ &= tr(Q)[I_p - B(p-2)p^{-1}I_p] \\ &= tr(Q) - B(p-2)p^{-1}tr(Q). \end{aligned}$$

مثال ۳. اگر در مثال ۲ فرض کنید عناصر بردار میانگین μ همگی یکسان و برابر μ باشند، یعنی θ دارای توزیع $N(\mu 1_p, A I_p)$ باشد به طوری که $\mu \in R$ و $A > 0$ هر دو نامعلوم اند. توزیع کناری X به صورت $N(\mu 1_p, B^{-1} I_p)$ است و از این رو، $\left(\bar{X}, \sum_{i=1}^p (X_i - \bar{X})^2\right)$ یک آماره بسنده کامل است و بنا براین $UMVUE$ برای μ و B به ترتیب برابر $\bar{X}$ و $\sum_{i=1}^p (X_i - \bar{X})^2 / (p-3)$ است. با جای گذاری این برآوردگرها در $E(\theta | X = x) = (1-B)x + B\mu$ ، برآوردگر EB برای θ بدین صورت به دست می آید:

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

این تعدیل برآوردگر جیمز- اشتاین را لپندلی [۶] مطرح کرده است [۱۰]. برآوردگر $\hat{\theta}_{EB}$ از این رو اهمیت دارد که برای $p \geq 4$ بر X غلبه می کند، یعنی مخاطره بیز آن (تحت تابع زیان L_1) از مخاطره X کم تر است. برای اثبات این حقیقت مخاطره بیز $\hat{\theta}_{EB}$ تحت دو تابع زیان L_1 و L_2 بدین صورت به دست می آید:

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

به طوری که $J_p = 1_p 1_p^T$ و همچنین برای $p \geq 4$ و تحت تابع زیان L_1 ، $\hat{\theta}_{EB}$ بر X غلبه می کند، یعنی [۶]:

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

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

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

با توجه به این که:

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

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

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

دوباره با استفاده از قضیهٔ باسو داریم:

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

و نیز با توجه به این که $E[(\mathbf{X} - \bar{X} \mathbf{1}_p)(\mathbf{X} - \bar{X} \mathbf{1}_p)^T] = B^{-1}(\mathbf{I}_p - p^{-1} \mathbf{J}_p)$ و چون $\sum_{i=1}^p (X_i - \bar{X})^2$ دارای

توزیع χ^2_{p-1} است، از این رو، برای هر $p \geq 4$ داریم:

$$\begin{aligned} E[(\hat{\theta}_{EB} - \hat{\theta}_B)(\hat{\theta}_{EB} - \hat{\theta}_B)^T] &= B^2 B^{-1}(\mathbf{I}_p - p^{-1} \mathbf{J}_p) - 2B(p-3)(p-1)^{-1}(\mathbf{I}_p - p^{-1} \mathbf{J}_p) \\ &\quad + (p-3)^2 B(p-3)^{-1}(p-1)^{-1}(\mathbf{I}_p - p^{-1} \mathbf{J}_p) + Bp^{-1} \mathbf{J}_p \\ &= B \mathbf{I}_p B(p-3)(p-1)^{-1}(\mathbf{I}_p - p^{-1} \mathbf{J}_p). \end{aligned}$$

بنا براین مخاطرهٔ بیز $\hat{\theta}_{EB}$ بدین صورت ساده می‌شود:

$$E[L_1(\theta, \hat{\theta}_{EB})] = E[L_1(\theta, \hat{\theta}_B)] + B \mathbf{I}_p - B(p-3)(p-1)^{-1}(\mathbf{I}_p - p^{-1} \mathbf{J}_p).$$

از طرف دیگر چنان که در مثال ۲ به‌دست آوردیم $E[L_1(\theta, \hat{\theta}_B)] = (1-B) \mathbf{I}_p$ که با جای‌گذاری آن در مخاطرهٔ

بیز $\hat{\theta}_{EB}$ داریم:

$$\begin{aligned} E[L_1(\theta, \hat{\theta}_{EB})] &= (1-B) \mathbf{I}_p + B \mathbf{I}_p - B(p-3)(p-1)^{-1}(\mathbf{I}_p - p^{-1} \mathbf{J}_p) \\ &= \mathbf{I}_p - B(p-3)(p-1)^{-1}(\mathbf{I}_p - p^{-1} \mathbf{J}_p). \end{aligned}$$

مشابه مثال ۲ با توجه به این که $L_2(\theta, \mathbf{a}) = \text{tr}(Q L_1(\theta, \mathbf{a}))$ ، مخاطرهٔ بیز $\hat{\theta}_{EB}$ 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 $E[L_1(\theta, \mathbf{X})] = \mathbf{I}_p$ داریم:

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

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

فواصل اطمینان بر پایهٔ احتمال‌های غیرشرطی

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

که $B = \frac{\sigma^2}{\sigma^2 + \tau^2}$ و همچنین متغیر تصادفی کناری $\mathbf{Y}$ دارای توزیع $N(\boldsymbol{\mu}, \frac{\sigma^2}{B} \mathbf{I}_m)$ است. حال اگر $\boldsymbol{\mu}$ نامعلوم باشد، آن گاه به طور معمول $\mathbf{Y}$ را به صورت 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 پایه این توزیع $(\bar{Y}, S^2)$ یک آماره 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 $\boldsymbol{\theta}$ به صورت $\hat{\boldsymbol{\theta}}^{EB} = (1 - \hat{B}(S))\bar{\mathbf{Y}} + \hat{B}(S)\bar{\mathbf{Y}}\mathbf{1}_m$ 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 به صورت $\hat{B}(S) = \min\left(\frac{(m-3)\sigma^2}{(m-1)S^2}, 1\right)$ است. هدف تعیین یک فاصله اطمینان EB برای یکی از مؤلفه‌های $\boldsymbol{\theta}$ مثل θ_1 است. اگر μ و τ^2 معلوم باشند، $(1-B)\bar{Y}_1 + B\mu \pm z_{\alpha/2}\sigma(1-B)^{1/2}$ 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 فاصله اطمینان برای θ_1 به صورت $(1 - \hat{B}(S))\bar{Y}_1 + \hat{B}(S)\bar{Y} \pm z_{\alpha/2}\sigma(1 - \hat{B}(S))^{1/2}$ است که در آن $0 < \hat{B}(S) < 1$ که این فاصله در $\bar{Y}$ تباهیده می‌شود اگر $\hat{B}(S) = 1$. موریس [۱۱] فاصله اطمینان بالا را با تقریب برآوردگرهای HB برای میانگین و واریانس $\boldsymbol{\theta}$ بنا ساخت. لایرد و لوئیس [۹] فاصله 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 فاصله اطمینان EB بر اساس اصلاح اریبی ارائه می‌شود. برای این منظور یک درجه‌بندی برای فاصله اطمینان EB به منظور بهتر کردن احتمالات پوشش غیرشرطی مجانبی معرفی می‌شود.

قضیه ۱. اگر Φ و ϕ به ترتیب تابع توزیع و تابع چگالی احتمال توزیع نرمال استاندارد باشند و برای $d > 1$ که به

m بستگی ندارد، $\hat{B}_d(S) = \min\left\{\frac{(m-d)\sigma^2}{(m-1)S^2}, \frac{m-d}{m-1}\right\}$ ، آن گاه برای m های بزرگ:

$$P\left[\theta_1 \in (1 - \hat{B}_d(S))\bar{Y}_1 + \hat{B}_d(S)\bar{Y} \pm t\sigma\sqrt{1 - \hat{B}_d(S)}\right] = 2\Phi(t) - 1 - \frac{t\phi(t)}{m}\left[\frac{(6-d)B}{1-B} + \frac{(1+t^2)B^2}{2(1-B)^2}\right] + O(m^{-3/2}).$$

اثبات. با در نظر گرفتن آماره کمکی $A = (\bar{Y}_1 - \bar{Y})/S$ 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:

$$P\left[\theta_1 \leq (1 - \hat{B})\bar{Y}_1 + \hat{B}\bar{Y} + t\sigma\sqrt{1 - \hat{B}}\right] = E\left\{P\left[\frac{\theta_1 - ((1-B)\bar{Y}_1 + B\mu)}{\sigma\sqrt{1-B}} \leq t + Z(t) | Y_1 = y_1\right]\right\} = E[\Phi(t + Z(t))] \quad (۱)$$

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

$$\Phi(t+Z(t)) = \Phi(t) + Z(t)\phi(t) + \frac{1}{2}Z^2(t)\phi'(t) + \frac{1}{2}\int_t^{t+Z(t)}(t+Z(t)-x)^2\phi''(x)dx. \quad (2)$$

از طرف دیگر با به کارگیری بسط تیلور داریم:

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

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

$$\begin{aligned} E(B-\hat{B}) &= E(B-\tilde{B}_d) + E(\tilde{B}_d-\hat{B}) \\ &= B - B(m-d)E\left(\frac{\sigma^2}{(m-1)BS^2}\right) + E(\tilde{B}_d-\hat{B}) \\ &= \frac{d-3}{m}B + E(\tilde{B}_d-\hat{B}) + O(m^{-2}) \end{aligned} \quad (4)$$

و به طور مشابه:

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

حال برای هر $1 \leq r$ ثابت و دل خواه داریم:

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

عبارت بالا برای هر $0 < u$ ثابت و دل خواه برقرار است. بنا براین برای هر $1 \leq r$ با انتخاب u به گونه ای که $r \leq u$ داریم، $\max\{m^{-u/2}, m^{-r/2}\} = m^{-r/2}$ و از این رو،

$$E|B-\hat{B}|^r \leq 2^{r-1}\left[E|B-\tilde{B}_d|^r + E|\tilde{B}_d-\hat{B}|^r\right] \leq 2^{r-1}\left[O(m^{-u/2}) + O(m^{-r/2})\right] = O(m^{-r/2}) \quad (6)$$

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

$$\left| \int_0^{\frac{B-\hat{B}}{1-B}} \left(\frac{B-\hat{B}}{1-B} - x \right)^2 (1+x)^{-5/2} dx \right| = I_1 + I_2 \quad (7)$$

که در آن

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

و

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

اما با توجه به این که برای هر $a > 0$,

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

بدین صورت داریم:

$$E(I_1) \leq E \left[\left(\frac{B-\hat{B}}{1-B} \right)^2 \frac{B-\hat{B}}{1-B} I_{(-\infty, B]}(\hat{B}) \right] \leq \frac{E|B-\hat{B}|^3}{(1-B)^3} = O(m^{-3/2}). \quad (8)$$

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:

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

بنا براین:

$$\begin{aligned} E \left\{ \left[\int_{\frac{B-\hat{B}}{1-B}}^0 \left(\frac{B-\hat{B}}{1-B} - x \right)^2 (1+x)^{-5/2} dx \right] I_{(B, B+\varepsilon_m)}(\hat{B}) \right\} &\leq E \left[\left(1 - \frac{\varepsilon_m}{1-B} \right)^{-5/2} \left\{ \int_{\frac{B-\hat{B}}{1-B}}^0 \left(\frac{B-\hat{B}}{1-B} - x \right)^2 dx \right\} I_{(B, B+\varepsilon_m)}(\hat{B}) \right] \\ &\leq \frac{1}{3} \left(1 - \frac{\varepsilon_m}{1-B} \right)^{-5/2} E \left[\left(\frac{B-\hat{B}}{1-B} \right)^3 I_{(B, B+\varepsilon_m)}(\hat{B}) \right] \\ &\leq \frac{1}{3} \left(1 - \frac{\varepsilon_m}{1-B} \right)^{-5/2} \frac{E|B-\hat{B}|^3}{(1-B)^3} \\ &= O(m^{-3/2}). \end{aligned} \quad (9)$$

علاوه بر آن:

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

اکنون با توجه به این که $(m-1)BS^2/\sigma^2$ دارای توزیع χ_{m-1}^2 است، متغیرهای تصادفی مستقل و هم‌توزیع $Z_1, \dots, Z_{m-1}$ با توزیع χ_1^2 وجود دارند به قسمی که $\sum_{i=1}^{m-1} Z_i = (m-1)\bar{Z}_{m-1}$ و بنا براین:

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

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

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

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

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

با توجه به این که $0 < \alpha < \frac{1}{2}$ ، با جای گذاری عبارت بالا و روابط (۴) و (۵) در رابطه (۳)، $E(Z(t))$ بدین صورت ساده می‌شود:

$$E(Z(t)) = t \left[\frac{(d-3)B}{2m(1-B)} - \frac{B^2}{4m(1-B)^2} \right] + O(m^{-3/2}). \quad (10)$$

از این که $(\bar{Y}, S^2)$ آماره بسنده کامل برای (μ, τ^2) و $A = (Y_1 - \bar{Y})/S$ یک آماره کمکی است، با استفاده از قضیهٔ باسو استقلال دودویی $\bar{Y}$ و S و A نتیجه می‌شود و چون $E(A^2) = 1$ ، می‌توان $E(Z^2(t))$ را نیز بدین صورت نوشت:

$$E(Z^2(t)) = \frac{1}{1-B} \left[t^2 E \left\{ 1 - \hat{B} + 1 - B - 2\sqrt{1-B}\sqrt{1-\hat{B}} \right\} + Bm^{-1} + E \frac{(B-\hat{B})^2 S^2}{\sigma^2} \right]. \quad (11)$$

اما برای هر $u > 0$ ثابت و دل خواه:

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

و همچنین:

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

و

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

پس با قرار دادن سه عبارت بالا در رابطه (۱۱)، $E(Z^2(t))$ به صورت (۱۲) ساده می شود:

$$E(Z^2(t)) = \frac{t^2 B^2}{2m(1-B)^2} + \frac{3B}{m(1-B)} + O(m^{-3/2}). \quad (12)$$

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

$$\begin{aligned} E\left|\int_t^{t+Z(t)} \phi''(x)(t+Z(t)-x)^2 dx\right| &\leq E\left\{|Z(t)|^2 \int_t^{t+Z(t)} 2\phi(\sqrt{3}) dx\right\} \\ &= 2E|Z(t)|^3 \phi(\sqrt{3}) \\ &= O(m^{-3/2}). \end{aligned} \quad (13)$$

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

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

با کم کردن این عبارت از عبارت مشابه به ازای $-t$ حکم قضیه نتیجه می شود.

مثال ۴. مدل رگرسیون بیزی را که در آن $\mathbf{Y}|\boldsymbol{\theta}$ دارای توزیع $N(\boldsymbol{\theta}, \sigma^2 I_m)$ و $\boldsymbol{\theta}$ دارای توزیع $N(\mathbf{X}\boldsymbol{\beta}, \tau^2 I_m)$ باشد در نظر بگیرید که در آن $\mathbf{X}$ یک ماتریس $m \times p$ با رتبه $p (< m)$ و بردار $p \times 1$ از ضرایب رگرسیونی است. فرض کنید σ^2 معلوم است. در حالتی که $\boldsymbol{\beta}$ و τ^2 معلوم هستند، توزیع پسین $\boldsymbol{\theta}|\mathbf{Y}$ به صورت $N((1-B)\mathbf{Y} + B\mathbf{X}\boldsymbol{\beta}, \sigma^2(1-B))$ 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 به ترتیب برابر با $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$ و $\hat{B} = \hat{B}_d(S) = \min\left\{\frac{(m-d)\sigma^2}{(m-p)S^2}, \frac{m-d}{m-p}\right\}$ هستند که در آن $S^2 = \|\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}\|^2 / (m-p)$ و $d > p$. بنا براین برآوردگر EB برای $\boldsymbol{\theta}$ برابر است با $(1 - \hat{B})\mathbf{Y} + \hat{B}\mathbf{X}\hat{\boldsymbol{\beta}}$. فاصله اطمینان معمولی برای θ_1 به صورت $(1 - \hat{B})Y_1 + \hat{B}x_1^T \hat{\boldsymbol{\beta}} \pm z_{\alpha/2} \sigma \sqrt{1 - \hat{B}}$ است که در آن x_1^T اولین بردار ستونی $\mathbf{X}$ است. مطابق قضیه ۱ با جای گذاری $\mathbf{x}_1^T (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$ به جای $\bar{Y} - \mu$ و تعریف $\hat{B}_{js}(S) = \frac{(m-p-2)\sigma^2}{(m-p)S^2}$ و همچنین تحت شرط $\max_{1 \leq i \leq m} \mathbf{x}_i^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_i = O(m^{-1})$ نتیجه می شود که:

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

$$= 2\Phi(t) - 1 - t\phi(t) \left[\frac{(1+t^2)B^2}{2m(1-B)^2} + \frac{B}{1-B} \left\{ \mathbf{x}_1^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{x}_1 + \frac{5-d}{m} \right\} \right] + O(m^{-3/2}).$$

فواصل اطمینان EB بر پایه احتمال‌های شرطی

هیل [۷]، در حالتی که B معلوم است، که تنها پارامتر مجهول در توزیع پیشین μ 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 $A = (Y_1 - \bar{Y})/S$ بررسی می‌شوند و قضیه زیر رهیافتی برای رسیدن به این منظور است.

قضیه ۲. باتوجه به شرایط قضیه ۱ داریم:

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

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

$$P\left[\theta_1 \leq (1 - \hat{B})Y_1 + \hat{B}\bar{Y} + t\sigma\sqrt{1 - \hat{B}} | A\right] = \Phi(t) + \phi(t)E[Z(t)|A] - \frac{1}{2}t\phi(t)E(Z^2(t)|A)$$

$$+ \frac{1}{2}E\left[\int_t^{t+Z(t)} (t + Z(t) - x)^2 \phi''(x) dx | A\right]. \quad (14)$$

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

$$E(Z(t)|A) = tE\left[\frac{B - \hat{B}}{2(1-B)} - \frac{(B - \hat{B})^2}{8(1-B)^2} + \frac{3}{16} \int_0^{\frac{B - \hat{B}}{1-B}} \left\{ \frac{B - \hat{B}}{1-B} - x \right\}^2 (1+x)^{-5/2} dx\right] + E\left[\frac{(B - \hat{B})S}{\sigma\sqrt{1-B}}\right]A$$

$$= t\left[\frac{(d-3)B}{2m(1-B)} - \frac{B^2}{4m(1-B)^2} + O(m^{-3/2})\right] + E\left[\frac{(B - \hat{B})S}{\sigma\sqrt{1-B}}\right]A \quad (15)$$

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

$$E(Z^2(t)|A) = \frac{t^2 B^2}{2m(1-B)^2} + \frac{B + 2BA^2}{m(1-B)} + \frac{t}{\sigma(1-B)} E[(B - \hat{B})^2 S]A + O_p(m^{-3/2}) \quad (16)$$

و سرانجام مثل قبل داریم:

$$\left| E\left[\int_t^{t+Z(t)} (t + Z(t) - x) \phi''(x) dx | A\right] \right| \leq 2\phi(\sqrt{3})E[|Z(t)|^3 | A] \quad (17)$$

و نیز برای هر $t > 0$ ، $E[|Z(t)|^3 | A] \leq 9(1-B)^{-3/2} E\left[t^3 \sqrt{1 - \hat{B}} - \sqrt{1 - B}\right]^3 + \frac{B^3 |\bar{Y} - \mu|^3}{\sigma^3} + \frac{(B - \hat{B})^3 S^3 |A|^3}{\sigma^3 |A|}$

همچنین $E|\bar{Y} - \mu|^3 = O_p(m^{-3/2})$ و نیز با کمک نامساوی (۶) برای $r = 3$ و جای‌گذاری $(B - \hat{B})S$

به جای $(B - \hat{B})$ می‌توان نتیجه گرفت، $E[(B - \hat{B})^3 S^3] \leq 4E[|B - \tilde{B}_d|^3 S^3 + |\tilde{B}_d - \hat{B}|^3 S^3] = O(m^{-3/2})$ با

استفاده از بسط تیلور مشابه (۳) داریم:

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

بنا بر این با توجه به چهار رابطه بالا به راحتی نتیجه می شود که:

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

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

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

که با تفاضل این عبارت از عبارت مشابه به ازای $-t$ و استفاده از رابطه (۱۵) مثل قضیه قبلی نتیجه مطلوب حاصل می شود.

مثال ۵. مدل رگرسیون بیزی مثال ۴ را در نظر بگیرید. تحت شرایط گفته شده در این مثال و با تعریف آماره کمکی

$$A = (Y_1 - \mathbf{x}_1^T \hat{\boldsymbol{\beta}}) / S$$

نتیجه می شود که:

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

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 می شود. فرض کنید $X_{ij} | \theta_i$ ها که $i = 1, \dots, k$ $j = 1, \dots, n_i$ متغیرهای تصادفی مستقل از خانواده نمایی با تابع چگالی $f(x | \theta_i) = \exp\{\theta_i x - Q(\theta_i) + R(x)\}$ باشند. اگر $\pi(\theta_i) \propto \exp\{\lambda[m_i \theta_i - Q(\theta_i)]\}$ و توزیع های پیشین مزدوج مستقل با تابع چگالی $\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$ باشند، آن گاه توزیع های پسین بدین صورت است:

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

میانگین پسین با استفاده از تساوی $\mu_i = Q'(\theta_i)$ به صورت $\hat{\mu}_{i,B} = E(\mu_i | \bar{X}_i) = [n_i \bar{X}_i + \lambda m_i] / [n_i + \lambda]$ محاسبه می شود. اگر $m_i = h(\mathbf{x}_i^T \mathbf{b})$ ، $i = 1, \dots, k$ از دیدگاه EB باید پارامترهای $\mathbf{b}$ و λ برآورد شوند. در این حالت $E(\bar{X}_i) = m_i$ و $\text{var}(\bar{X}_i) = \text{var}(\mu_i) + E(M(\mu_i)) = \frac{\lambda + n_i}{\lambda - a_2} M(m_i)$ که در آن $\sigma^2 = M(\mu) = a_0 + a_1 \mu + a_2 \mu^2$ و a_i ها همزمان صفر نیستند. اگر $\hat{\mathbf{b}}$ و $\hat{\lambda}$ برآوردگرهای به ترتیب $\mathbf{b}$ و λ باشند، $\hat{\mu}_{i,EB} = [n_i \bar{X}_i + \hat{\lambda} \hat{m}_i] / [n_i + \hat{\lambda}]$ برآوردگرهای EB برای μ_i هستند که $\hat{m}_i = h(\mathbf{x}_i^T \hat{\mathbf{b}})$ و برای k بزرگ و هر $\eta \geq 2$:

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

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

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

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 رابطه (۱۸) می‌توان مشاهده کرد که $E |\hat{m}_i - m_i|^{2r} = O(k^{-r})$ با به‌کارگیری نامساوی‌های شوارتز و مارکف:

(۱۹)

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

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

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

و سپس با استفاده از (۱۹) و

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

نتیجه می‌شود که

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

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

$$\left(1 - \frac{a_2 z_{\alpha/2}^2}{n_i} \right)^{-1} \left[\hat{\mu}_{i,EB} + \frac{a_1 z_{\alpha/2}}{2n_i} \pm \frac{z_{\alpha/2}}{n_i^{1/2}} \sqrt{M(\hat{\mu}_{i,EB}) + \frac{dz_{\alpha/2}^2}{4n_i}} \right].$$

شبیه‌سازی و نتیجه‌گیری

برای به‌دست آوردن فاصله اطمینان بیز تجربی ناوردا $100(1-\alpha)\%$ 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 قضیه ۱ داریم:

$$P \left[\theta_1 \in (1-\hat{B})Y_1 + \hat{B}\bar{Y} \pm t(\hat{B}, \alpha) \sigma \sqrt{(1-\hat{B})} \right] = 1 - \alpha + O(m^{-3/2}). \quad (20)$$

برای اثبات ناوردایی این فاصله اطمینان تحت گروه مکانی

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

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

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

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

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

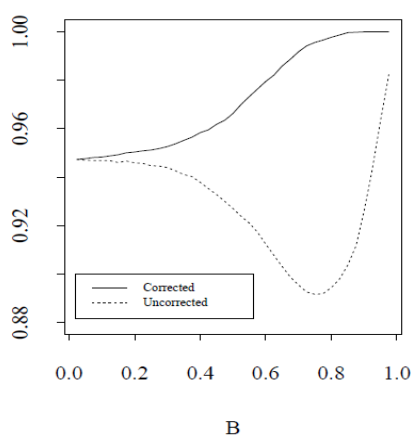
در قضیه ۲ یک فاصله اطمینان بیز تجربی ناوردای شرطی $100(1-\alpha)\%$ برای θ_1 به صورت:

$$P \left[\theta_1 \in (1-\hat{B})Y_1 + \hat{B}\bar{Y} \pm t(\hat{B}, A, \alpha) \sigma \sqrt{(1-\hat{B})} \mid A \right] = 1 - \alpha + O_p(m^{-3/2}) \quad (21)$$

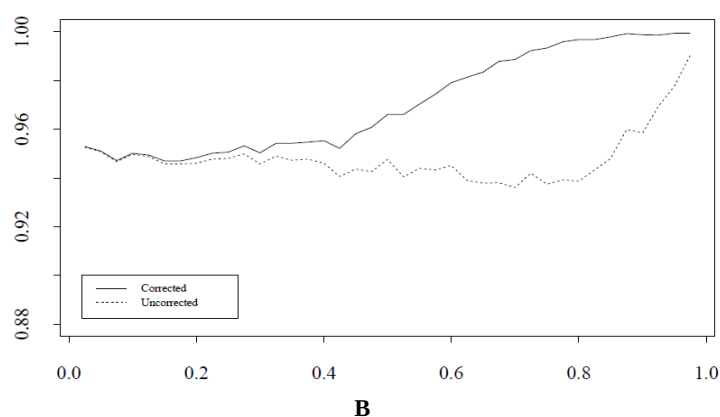
به دست می آید که به طور مشابه فاصله اطمینان بالا تحت گروه مکانی G ناورداست. همچنین مشاهده می شود که فواصل اطمینان تصحیح نشده در قضیه های ۱ و ۲ خاصیت ناوردایی تحت گروه مکانی را ندارند ولی با توجه به روابط (۲۰) و (۲۱) فواصل اطمینان تصحیح شده تحت گروه مکانی ناوردا هستند.

برای مشاهده نتایج قضیه های ۱ و ۲ با استفاده از شبیه سازی در حالت $\mu = 0$ و $\sigma^2 = 1$ و برای مقادیر مختلف B از ۰/۰۲۵ تا ۰/۹۷۵، تعداد ۱۰۰۰۰ زوج داده های مستقل (Y_i, θ_i) برای $i = 1, \dots, 30$ از توزیع نرمال شبیه سازی می شود. برای قضیه ۲، ۱۰۰۰۰ نمونه شبیه سازی شده را به شرط آماری کمکی A با مقدار ۱ تولید می شود. در شکل ۱ و ۲ به ترتیب فاصله اطمینان بیز تجربی شرطی و غیرشرطی $100(1-\alpha)\%$ برای θ_1 ، در حالت تصحیح نشده یعنی استفاده از قضیه های ۱ و ۲ (نمودار Uncorrected با خطوط نقطه چین) و تصحیح شده یعنی روابط (۲۱) و (۲۱) (نمودار Corrected با خطوط توپر) با $t = z_{\alpha/2}$ شبیه سازی شده است.

A=1.0



شکل ۱. فاصله اطمینان بیز تجربی شرطی برای θ_1



شکل ۲. فاصله اطمینان بیز تجربی غیرشرطی برای θ_1

منابع

- Berger J. O., "Statistical Decision Theory and Bayesian Analysis", 2nd Edition, Verlag, New York, Springer (1985).
- Carlin B., Gelfand A., "Approaches for empirical Bayes confidence intervals", J. Amer. Statist. Assoc. 85 (1990) 105-114.
- 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, New York (1975) 47-55.
- Efron B., "Better bootstrap confidence intervals", J. Amer. Statist. Assoc. 82 (1987) 171-185.
- Efron B., Morris C., "Stein's Estimation Rule and its Competitors Empirical Bayes Approach", J. Amer. Statist. Assoc., 68 (1973) 117-130.
- 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.
- Hill J. R., "A General Framework for Model-Based Statistics", Biometrika, 77 (1990) 115-126.
- 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.
- Laird N. M., Louis T. A., "Empirical Bayes confidence intervals based on bootstrap samples", J. Amer. Statist. Assoc. 82 (1987) 739-757.
- Lindley D. V., "Discussion of the paper by Stein", J. Roy. Statist. Soc. Ser. B, 24 (1962) 265-296.
- Morris C. N., "Parametric Empirical Bayes Inference: Theory and Applications", J. Amer. Statist. Assoc., 78 (1983) 47-65.