Bobrovsky-Mayer-Wolf-Zakai の下界の改良

日本大学・商学部　小池健一 (Ken-ichi Koike)
(College of Commerce, Nihon University)

広島大学・理学研究科　橋本真太郎 (Shintaro Hashimoto)
(Department of Mathematics, Hiroshima University)

1　はじめに

ベイズリスクの下界を与える Cramér-Rao 型の不等式には様々なものがあり，これまでに多くの研究がある（例えば，van Trees [12]，Borovkov and Sakhanenko [4], Borovkov [3]）．また，母数の次元を多次元化したものとして，Gill and Levit [5] などがよく知られている．しかし，これらの下界は sharp であるとはいえず，Bhattacharrya 型への改良が Koike [8] や Hashimoto and Koike [6] により与えられている．最近，Abu-Shanab and Veretennikov [1] や Koike [10] により，これらの下界の漸近的な比較が行われ，Borovkov-Sakhanenko の下界 [4] はあるクラスにおいて漸近的に最適であり，van Trees の下界 [12] を漸近的に優越することが示された．さらに，Koike [9] では指数型分布族に対して共役事前分布と Jeffreys の事前分布のもとで [12] や [3] の不等式の等号達成条件が導出されている．

一方，van Trees の下界の Chapman-Robbins 型への拡張である Bobrovsky-Mayor-Wolf-Zakai の不等式（Bobrovsky, Mayer-Wolf and Zakai [2]）は，ベイズリスクの下界に関する差分型の不等式として知られている．本論では，[2] による下界を Borovkov and Sakhanenko の下界の Chapman-Robbins 型への拡張を行うことにより改良する．さらに，共役事前分布のもとで，正規分布の平均母数の２乗の推定において下界の正確な比較を与える．また，ベルヌーイ・ロジットモデルの場合に，サンプルサイズが大きいときにラプラス近似を用いた漸近的な下界の比較を行うことにより，導出した下界が Bobrovsky-Mayor-Wolf-Zakai の下界 [2] よりも漸近的に優れていることを示す．

2 Bobrovsky-Mayer-Wolf-Zakai の下界の改良

$X_1, \ldots, X_n$ を互いに独立にいずれも（σ-有限測度 μ に関する）確率密度関数 $f_1(x|\theta)$ $(\theta \in \Theta)$ をもつ分布に従う確率変数列とする．ただし，Θ は（無限区間でも構わない）$\mathbb{R}$ 上の区間であり，その端点を a, b $(-\infty \leq a < b \leq \infty)$ とする．$X := (X_1, \ldots, X_n)$ の同時確率密度関数は $f(x|\theta) := \prod_{i=1}^n f_1(x_i|\theta)$ となる．ただし，$x = (x_1, \ldots, x_n)$ とする．また，$\lambda(\theta)$ を（ルベーグ測度に関する）θ の事前密度とする．Θ 上の関数 g に対して，g の台を $\mathrm{supp}(g)$ と表記する．いま，θ の微分可能な関数 $\varphi(\theta)$ の 2 乗損失関数 $L(\theta, a) = (a - \varphi(\theta))^2$ のもとでのベイズ推定問題を考える．(X, θ) の同時確率密度関数 $p(x, \theta)$ は，$p(x, \theta) = f(x|\theta)\lambda(\theta)$ により与えられる．以下では，$E(\cdot)$ とかいたら $p(x, \theta)$ のもとでの期待値をとると約束する．また，$E_\theta(\cdot)$ とかいたら $f(x|\theta)$ のもとでの期待値をとると約束する．$I(\theta)$ を単位標本あたりのフィッシャー情報量，すなわち $I(\theta) = -E_\theta\{\partial^2 \log f_1(X_1|\theta)/\partial\theta^2\} = E_\theta[\{\partial \log f_1(X_1|\theta)/\partial\theta\}^2]$ とする．

このとき，次の定理が成り立つ：

定理 1. $\varphi(\theta)$ の推定量 $\hat{\varphi}(X)$ と任意の 0 でない実数 h に対して，ベイズリスクに関する以下の不等式が成り立つ：

$$
\begin{aligned}
E\left\{(\hat{\varphi}(X) - \varphi(\theta))^2\right\} &\geq \frac{\{\mathrm{Cov}(G_h, \hat{\varphi}(X) - \varphi(\theta))\}^2}{E(G_h^2)} \\
&= \frac{\left\{E\left[\{\varphi(\theta) - \varphi(\theta - h)\}\frac{\varphi'(\theta)}{I(\theta)}\right]\right\}^2}{E\left[\left\{\frac{p(X,\theta+h)}{p(X,\theta)}\frac{\varphi'(\theta+h)}{I(\theta+h)} - \frac{\varphi'(\theta)}{I(\theta)}\right\}^2\right]}.
\end{aligned} \tag{1}
$$

以下に証明の概略を述べる．

Proof. $G_h = \frac{1}{h}\left(\frac{p(x,\theta+h)}{p(x,\theta)}\frac{\varphi'(\theta+h)}{I(\theta+h)} - \frac{\varphi'(\theta)}{I(\theta)}\right)$ とおく．まず，

$$
\begin{aligned}
\lim_{h \to 0} G_h =& \frac{1}{p(x,\theta)} \lim_{h \to 0} \frac{1}{h}\left\{p(x, \theta+h)\frac{\varphi'(\theta+h)}{I(\theta+h)} - p(x,\theta)\frac{\varphi'(\theta)}{I(\theta)}\right\} \\
=& \frac{1}{p(x,\theta)} \frac{\partial}{\partial\theta}\left\{p(x,\theta)\frac{\varphi'(\theta)}{I(\theta)}\right\} \quad (=: G_0)
\end{aligned}
$$

であることに注意する．

このとき，

$$E(G_h)=0,\quad E\{G_h\varphi(\theta)\}=\frac{1}{h}E\left[\{\varphi(\theta-h)-\varphi(\theta)\}\frac{\varphi'(\theta)}{I(\theta)}\right]$$

が成り立つ．フビニの定理により，$E\{G_h\hat{\varphi}(X)\}=0$ が成り立つので，

$$\mathrm{Cov}(G_h,\hat{\varphi}(X)-\varphi(\theta))=\frac{1}{h}E\left[\{\varphi(\theta)-\varphi(\theta-h)\}\frac{\varphi'(\theta)}{I(\theta)}\right]$$

となる．分散共分散不等式（variance-covariance inequality）により，

$$\begin{aligned}E\left\{(\hat{\varphi}(X)-\varphi(\theta))^2\right\}&\geq\frac{\{\mathrm{Cov}(G_h,\hat{\varphi}(X)-\varphi(\theta))\}^2}{E(G_h^2)}\\&=\frac{\left\{E\left[\{\varphi(\theta)-\varphi(\theta-h)\}\frac{\varphi'(\theta)}{I(\theta)}\right]\right\}^2}{E\left[\left\{\frac{p(X,\theta+h)}{p(X,\theta)}\frac{\varphi'(\theta+h)}{I(\theta+h)}-\frac{\varphi'(\theta)}{I(\theta)}\right\}^2\right]}\end{aligned}$$

となり，定理の不等式を得る． □

Borovkov-Sakhanenko の下界 [4] は分散共分散不等式から

$$E\left\{(\hat{\varphi}(X)-\varphi(\theta))^2\right\}\geq\frac{\{\mathrm{Cov}(G_0,\hat{\varphi}(X)-\varphi(\theta))\}^2}{E(G_0^2)}$$

により与えられる．$\lim_{h\to 0}G_h=G_0$ であることにより，絶対値が十分小さな $h\neq 0$ に対する下界 (1) の値は，いくつかの正則条件のもとで Borovkov-Sakhanenko の下界の値に十分近いものになる．

同様にして，Bobrovsky-Mayer-Wolf-Zakai の下界 [2] は分散共分散不等式から

$$E\left\{(\hat{\varphi}(X)-\varphi(\theta))^2\right\}\geq\frac{\{\mathrm{Cov}(B_h,\hat{\varphi}(X)-\varphi(\theta))\}^2}{E(B_h^2)}=\frac{[E\{\varphi(\theta)-\varphi(\theta-h)\}]^2}{E\left[\left\{\frac{p(X,\theta+h)}{p(X,\theta)}\right\}^2\right]-1}\qquad(2)$$

により与えられる．ただし，$B_h=\frac{1}{h}\left(\frac{p(x,\theta+h)}{p(x,\theta)}-1\right)$ とする．また，$\lim_{h\to 0}B_h=B_0=\frac{\frac{\partial}{\partial\theta}p(x,\theta)}{p(x,\theta)}$ であることと， B_0 を用いた分散共分散不等式から van Trees の下界 [12] が導かれるので，絶対値が十分小さな $h\neq 0$ に対する Bobrovsky-Mayor-Wolf-Zakai の下界はいくつかの正則条件のもとで van Trees の下界の値に十分近くなる．

一方で，Abu-Shanab and Veretennikov [1] や Koike [10] は，Borovkov-Sakhanenko

の下界 [4] はあるクラスにおいて漸近的に最適であり，van Trees の下界 [12] を漸近的に優越すること，すなわち，以下の不等式が成り立つことを示した：

$$\lim_{n\to\infty} n \frac{\left\{E\left(\dfrac{\psi'^2}{I}\right)\right\}^2}{nE\left(\dfrac{\psi'^2}{I}\right)+E\left[\left\{\dfrac{(\psi'\lambda/I)'}{\lambda}\right\}^2\right]} \geq \lim_{n\to\infty} n\frac{\{E(\psi')\}^2}{nE(I)+E\{\left(\frac{\lambda'}{\lambda}\right)^2\}}.$$

ゆえに，下界 (1) は絶対値が十分小さな $h\neq 0$ に対して，Bobrovsky-Mayer-Wolf-Zakai の下界を漸近的に優越するということになる．

例 1 (正規分布)．$X_1,\ldots,X_n$ 互いに独立にいずれも正規分布 $\mathrm{N}(\theta,1)$ $(\theta\in\Theta=\mathbb{R})$ に従う確率変数列とする．θ の事前分布を $\mathrm{N}(m,\tau^2)$ とする．ただし，m と $\tau^2>0$ は既知の定数とする．また，$X=(X_1,\ldots,X_n)$, $x=(x_1,\ldots,x_n)$ とおく．このモデルにおける単位標本あたりのフィッシャー情報量 $I(\theta)$ は 1 である．$\varphi(\theta)=\theta^2$ の推定問題を考える[*1]．

この設定のもとで，Bobrovsky-Mayer-Wolf-Zakai の下界 (2) は

$$\frac{\{h(2m-h)\}^2}{\exp\left\{h^2\left(n+\frac{1}{\tau^2}\right)\right\}-1}\ (=:\mathrm{BMZ}_h)$$

により与えられる．一方，定理 1 による下界 (1) は

$$\frac{\left[h\left\{2(m^2+\tau^2)-mh\right\}\right]^2}{\{\tau^2+(m-h)^2\}\exp\left\{h^2\left(n+\frac{1}{\tau^2}\right)\right\}-(m^2+\tau^2-2hm)}\ (=:\mathrm{KH}_h)$$

で与えられる．

下界の極限，つまり $\mathrm{BMZ}_0=\lim\limits_{h\to 0}\mathrm{BMZ}_h$ と $\mathrm{KH}_0=\lim\limits_{h\to 0}\mathrm{KH}_h$ の比較を行う．まず，

$$\mathrm{BMZ}_0=\frac{4m^2}{n+\frac{1}{\tau^2}},\quad \mathrm{KH}_0=\frac{4(m^2+\tau^2)}{\frac{1}{m^2+\tau^2}+n+\frac{1}{\tau^2}}$$

より，

$$\frac{4}{\mathrm{BMZ}_0}-\frac{4}{\mathrm{KH}_0}=\frac{1}{m^2(m^2+\tau^2)}\left(\frac{m^2}{m^2+1}-n\tau^2-1\right)$$

[*1] $\varphi(\theta)=\theta$ に対しては，下界 (1) は下界 (2) に一致することに注意．

$$
\begin{aligned}
&< \frac{1}{m^2(m^2+\tau^2)}\left(1 - n\tau^2 - 1\right) \\
&= \frac{-n\tau^2}{m^2(m^2+\tau^2)} < 0
\end{aligned}
$$

が成り立つので，$4/\mathrm{BMZ}_0 > 4/\mathrm{KH}_0$ となる．したがって，任意の $n \in \mathbb{N}$ に対して，$\mathrm{BMZ}_0 < \mathrm{KH}_0$ が成り立つ．

3　ラプラス近似による漸近比較

前節の例 1 では，正確な下界の比較を行った．ここでは，Koike [9] の Example 2 でも論じられているベルヌーイ・ロジットモデル[*2]を考え，サンプルサイズ n が大きいときの下界の漸近比較をすることにより定理 1 による下界 (1) が Bobrovsky-Mayer-Wolf-Zakai の下界 (2) よりも漸近的に良いことを示す．

例 2 (ベルヌーイ・ロジットモデル)**.** $X_1, \ldots, X_n$ を互いに独立にいずれもベルヌーイ分布 Ber $\left(\frac{e^\theta}{1+e^\theta}\right)$ $(\theta \in \mathbb{R})$ に従う確率変数列とする．θ を与えたもとでの X_i の確率密度関数は

$$
f(x_i|\theta) = \left(\frac{e^\theta}{1+e^\theta}\right)^{x_i}\left(\frac{1}{1+e^\theta}\right)^{1-x_i} = e^{\theta x_i}\frac{1}{1+e^\theta} \quad (x_i = 0, 1;\ \theta \in \mathbb{R};\ i = 1, \ldots, n)
$$

により与えられ，尤度比は次で与えられる：

$$
\frac{f(x_i|\theta+h)}{f(x_i|\theta)} = e^{(\theta+h)x_i}\frac{1}{1+e^{\theta+h}}e^{-\theta x_i}(1+e^\theta) = e^{hx_i}\frac{1+e^\theta}{1+e^{\theta+h}} \quad (h \in \mathbb{R}).
$$

また，θ の事前密度を共役事前分布である，第 2 種ベータ分布（例えば，Johnson, Kotz and Balakrishnan [7]）とすると事前密度関数は

$$
\lambda(\theta) = \frac{1}{\mathrm{B}(c_1, c_2 - c_1)}e^{c_1\theta}(1+e^\theta)^{-c_2} \quad (\theta \in \mathbb{R};\ c_2 > c_1 > 0)
$$

[*2] 例 1 と同様，正確な下界を比較することは興味深いが，事前分布の超母数 $c_1, c_2 > 0$ の定め方に注意が必要であることと，計算が複雑になるためここでは漸近的な比較のみ示す．

となる．ただし，$\mathrm{B}(\cdot,\cdot)$ はベータ関数とする．このとき，事前密度の比は次で与えられる：

$$\frac{\lambda(\theta+h)}{\lambda(\theta)}=e^{(\theta+h)c_1}(1+e^{\theta+h})^{-c_2}e^{-\theta c_1}(1+e^{\theta})^{c_2}=e^{c_1h}(1+e^{\theta+h})^{-c_2}(1+e^{\theta})^{c_2}.$$

いま，$X=(X_1,\ldots,X_n)$, $x=(x_1,\ldots,x_n)$ とかくことにすると (X,θ) の同時確率密度関数の比は

$$\begin{aligned}P:=\frac{p(x,\theta+h)}{p(x,\theta)}&=\left\{\prod_{i=1}^{n}\frac{f(x_i|\theta+h)}{f(x_i|\theta)}\right\}\frac{\lambda(\theta+h)}{\lambda(\theta)}\\&=e^{h\sum_{i=1}^{n}x_i}(1+e^{\theta})^{n+c_2}(1+e^{\theta+h})^{-n-c_2}e^{c_1h}\end{aligned}$$

により与えられ，

$$E(P^2)=E\left[\left\{E_{X|\theta}\left(e^{2hX_1}\right)\right\}^n(1+e^{\theta})^{2n+2c_2}(1+e^{\theta+h})^{-2n-2c_2}e^{2c_1h}\right]$$

が成り立つ．また，$E_{X|\theta}(e^{2hX_1})=(1+e^{\theta})^{-1}(1+e^{\theta+2h})$ であるから，

$$\begin{aligned}E(P^2)=&\frac{e^{2c_1h}}{\mathrm{B}(c_1,c_2-c_1)}\\&\times\int_{-\infty}^{\infty}(1+e^{\theta})^{c_2}(1+e^{\theta+h})^{-2c_2}e^{c_1\theta}\left\{(1+e^{\theta})(1+e^{\theta+h})^{-2}(1+e^{\theta+2h})\right\}^n\mathrm{d}\theta.\end{aligned}$$

と計算できる．ここで，次の積分 I に関するラプラス近似 (例えば，Small [11]) を考える：

$$I=\int_{-\infty}^{\infty}(1+e^{\theta})^{c_2}(1+e^{\theta+h})^{-2c_2}e^{c_1\theta}\left\{(1+e^{\theta})(1+e^{\theta+h})^{-2}(1+e^{\theta+2h})\right\}^n\mathrm{d}\theta.$$

ラプラス近似を正当化するいくつかの条件は容易に確かめられる．積分 I のラプラス近似により十分大きい n に対して

$$I\sim\exp\{nk(-h)\}\,g_1(-h)\sqrt{\frac{2\pi}{-nk''(-h)}}$$

となる．ただし，$k(\theta)=\log\left\{(1+e^{\theta})(1+e^{\theta+h})^{-2}(1+e^{\theta+2h})\right\}$, $g_1(\theta)=(1+e^{\theta})^{c_2}(1+e^{\theta+h})^{-2c_2}e^{c_1\theta}$ とする．従って，Bobrovsky-Mayer-Wolf-Zakai の下界は十分大きい n に

対して

$$\frac{h^2}{\frac{e^{2c_1h}}{\mathrm{B}(c_1,c_2-c_1)}J_n(-h)} \tag{3}$$

と近似できる．ただし，

$$J_n(-h)=\exp\{nk(-h)\}\,g_1(-h)\sqrt{\frac{2\pi}{-nk''(-h)}}$$

とする．上と同様にして，定理 1 の下界の主要項は十分大きい n に対して

$$\frac{\left\{h\frac{\mathrm{B}(c_1-1,c_2-c_1-1)}{\mathrm{B}(c_1,c_2-c_1)}\right\}^2}{\frac{e^{2(c_1-1)h}}{\mathrm{B}(c_1,c_2-c_1)}J_n(-h)} \tag{4}$$

と近似される．(3) と (4) の比を考えることにより

$$\begin{aligned}
\frac{(h^2)\Big/\left\{\frac{e^{2c_1h}}{\mathrm{B}(c_1,c_2-c_1)}J_n(-h)\right\}}{\left\{h\frac{\mathrm{B}(c_1-1,c_2-c_1-1)}{\mathrm{B}(c_1,c_2-c_1)}\right\}^2\Big/\left\{\frac{e^{2(c_1-1)h}}{\mathrm{B}(c_1,c_2-c_1)}J_n(-h)\right\}} &=\left\{\frac{\mathrm{B}(c_1,c_2-c_1)}{\mathrm{B}(c_1-1,c_2-c_1-1)}\right\}^2 e^{-2h}\\
&=\left\{\frac{(c_1-1)(c_2-c_1-1)}{(c_2-2)(c_2-1)}\right\}^2 e^{-2h}\\
&<\left\{\frac{(c_2-1)(c_2-2)}{(c_2-2)(c_2-1)}\right\}^2 e^{-2h}\\
&=e^{-2h}
\end{aligned}$$

となる．絶対値が十分小さな h に対して $e^{-2h}\approx 1$ なる近似が成り立つため，十分小さな h に対して (4) は (3) よりも漸近的に良い下界となることがわかる．

参考文献

[1] Abu-Shanab, R. and Veretennikov, A.Yu. (2015). On asymptotic Borovkov-Sakhanenko inequality with unbounded parameter set. *Theor. Probability and Math. Statist.*, **90**, 1–12.

[2] Bobrovsky, B.Z., Mayer-Wolf, E. and Zakai, M. (1987). Some classes of Grobal Cramér-Rao bounds. *Ann. Statist.*, **15**, 1421–1438.

[3] Borovkov, A. A. (1998). *Mathematical Statistics.* Gordon and Breach, Amsterdam.

[4] Borovkov, A. A. and Sakhanenko, A. U. (1980). On estimates of the expected quadratic risk (in Russian). *Probab. Math. Statist.*, **1**, 185–195.

[5] Gill, R.D. and Levit, B.Y. (1995). Applications of the van Trees inequality: A Bayesian Cramér-Rao bound. *Bernoulli*, **1**, 59–79.

[6] Hashimoto, S. and Koike, K. (2015). Bhattacharyya type information inequality for the Bayes risk, *Commun. Statist.–Theory Meth.*, **44**, 5213–5224.

[7] Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions, volume 2 (2nd edition).* Wiley, New York.

[8] Koike, K. (2006). An integral Bhattacharyya type bound for the Bayes risk. *Commun. Statist.–Theory Meth.* **35**, 2185–2195.

[9] Koike, K. (2019). Attainments of the Bayesian information bounds. *Commun. Statist.–Theory Meth.* **48**.

[10] Koike, K. (2020). Asymptotic comparison of some Bayesian information bounds. *Commun. Statist.–Theory Meth.* **49**.

[11] Small, C.G. (2010). *Expansions and Asymptotics for Statistics.* CRC Press, New York.

[12] van Trees, H.L. (1968). *Detection, Estimation, and Modulation Theory, Part I*, Wiley, New York.