

ベイズリスクに対するバツタチャリヤ型情報不等式

筑波大学・数理物質科学研究科* 橋本 真太郎

Shintaro Hashimoto

Graduate School of Pure and Applied Sciences

University of Tsukuba

筑波大学・数理物質系† 小池 健一

Ken-ichi Koike

Faculty of Pure and Applied Sciences

University of Tsukuba

1 序章

不偏推定量の下界を与える情報不等式としてクラメール・ラオの不等式があり, その下界を達成するのは指数型に限るということはよく知られている. しかし, 指数型分布族であってもクラメール・ラオの下界を達成しないことは多い. そこで, これを精密化した情報不等式としてバツタチャリヤの不等式がある.

一方で, ベイズリスクの下界を与える不等式として Borovkov and Sakhanenko (1980) による不等式がある. これはクラメール・ラオ型の下界であり, Koike(2006) はこの Borovkov and Sakhanenko (1980) の不等式をバツタチャリヤ型に精密化し, さらに下界の二次漸近評価式を与えた. なお, Brown and Gajek (1990) や Sato and Akahira (1996) は推定量のバイアス (偏り) を考慮したクラメール・ラオ型の下界について考察している. そこで, 今回は Brown and Gajek (1990) による下界をバツタチャリヤ型に精密化し, その二次漸近評価式を導出することで従来の下界に対する明示的な差を与える. また, 応用として局所ミニマックスリスクの下界の二次漸近近似式について述べる.

2 準備

$X_1, \dots, X_n$ を互いに独立に、いずれも (σ 有限測度 μ に関する) 密度関数 $f_0(x, \theta)$ ($\theta \in \Theta$) に従う確率変数列とする。ただし、 Θ は端点 a, b ($-\infty \leq a < b \leq \infty$) をもつ実数の区間とする。また、 $X := (X_1, \dots, X_n)$ の同時密度関数を $f(x, \theta) := \prod_{i=1}^n f_0(x_i, \theta)$ とおく。ただし、 $f(x, \theta)$ は θ に関して二回微分可能であるとする。 θ の関数 $g(\theta)$ に対して、二乗損失 $L(\theta, d) = (d - g(\theta))^2$ のもとでのベイズ推定問題を考える。ここで、 θ の (ルベーグ測度に関する) 事前密度を $q(\theta)$ とし、任意の推定量 δ に対し、 $R(\theta, \delta) = E_\theta\{L(\theta, \delta)\}$ を δ の期待損失 (リスク)、 $B(q, \delta) = \int R(\theta, \delta)q(\theta)d\theta$ とおき、 $B(q) = \inf_\delta B(q, \delta)$ を事前密度 q のもとでのベイズリスクとする。以後、簡単のために θ に関する微分を $\partial f / \partial \theta = f'$, $\partial^2 f / \partial \theta^2 = f''$ などと表記することにする。

さらに、次の条件 (A1) と (A2) を仮定する。

(A1) 二階微分可能な関数 $g(\theta)$ の推定量 δ に対し、 $e(\theta) = E_\theta(\delta)$ とおくと、以下の (二次の) バッタチャリヤの不等式が成り立つ：

$$\text{Var}_\theta(\delta) \geq (e'(\theta), e''(\theta)) V_n^{-1} (e'(\theta), e''(\theta))^T, \quad (2.1)$$

ただし

$$V_n = \begin{pmatrix} I_n & K_n \\ K_n & J_n \end{pmatrix}$$

とし、 $I_n(\theta) = E_\theta\{(f'/f)^2\}$, $J_n(\theta) = E_\theta\{(f''/f)^2\}$, $K_n(\theta) = E_\theta(f'f''/f^2)$ である ($n \geq 1$)。また、 $I_1(\theta)$ は連続微分可能である。(2.1) が成り立つための十分条件は Zacks (1971) を参照のこと。

(A2) θ の事前密度 $q(\theta)$ は二回連続微分可能で、 $q(a) = q(b) = 0$ である。

また次の事実を与えておく。

補題 2.1. (Koike (2006)) 条件 (A1) と (A2) のもとで、以下の関係が成り立つ

$$I_n = nI_1, \quad J_n = nJ_1 + n(n-1)I_1^2, \quad K_n = nK_1. \quad (2.2)$$

3 先行研究について

この節では、次節の準備として、ベイズリスクの下界を与える不等式についての代表的な先行研究について証明なしで紹介する。まず、Borovkov and Sakhanenko (1980) は標

本と母数の積測度に対して、コーシー・シュワルツの不等式を適用することで次のようなベイズリスクに対するクラメール・ラオ型の不等式を示した。

定理 3.1. (Borovkov and Sakhanenko (1980)) 標本数 $n = 1$ のとき、条件 (A1) と (A2)のもとで以下の不等式が成り立つ

$$B(q, \delta) \geq \frac{C^2}{C + D}, \quad (3.1)$$

ただし

$$C := \int_{\Theta} \frac{g'^2 q}{I_1} d\theta, \quad D := \int_{\Theta} \frac{1}{q} \left(\frac{g' q}{I_1} \right)'^2 d\theta.$$

標本数が一般の $n \geq 1$ なら、(2.2) を用いることにより、(3.1) の漸近近似は

$$B(q, \delta) \geq Cn^{-1} + (-D)n^{-2} + O(n^{-3}) \quad (n \rightarrow \infty) \quad (3.2)$$

となる。

一方で、Brown and Gajek (1990) は母数が与えられた時の標本の確率測度に対して、コーシー・シュワルツの不等式（つまり、この場合クラメール・ラオの不等式）を適用し、それを母数で積分することで推定量の偏りを考慮したベイズリスクに対するクラメール・ラオ型の不等式を与えた。

定理 3.2. (Brown and Gajek (1990)) 標本数 $n = 1$ のとき、条件 (A1) と (A2) のもとで以下の不等式が成り立つ

$$B(g, \delta) \geq \frac{C^2}{C + D} + Q_1(b) + Q_2(b) + Q_3(b), \quad (3.3)$$

ただし

$$\begin{aligned} b(\theta) &= E_{\theta}\{\delta(\mathbf{X})\} - g(\theta), \\ \gamma(b) &= \int_{\Theta} b \left(\frac{g' q}{I_1} \right)' d\theta = - \int_{\Theta} \frac{g' b' q}{I_1} d\theta, \\ Q_1(b) &:= \frac{C + D}{CD} \left(\gamma - \frac{CD}{C + D} \right)^2 \geq 0, \\ Q_2(b) &:= \int_{\Theta} \left(b' + \frac{g' \gamma}{C} \right)^2 \frac{q}{I_1} d\theta \geq 0, \\ Q_3(b) &:= \int_{\Theta} \left\{ b - \frac{\gamma}{D} \frac{1}{q} \left(\frac{g' q}{I_1} \right)' \right\}^2 q d\theta \geq 0. \end{aligned} \quad (3.4)$$

Borovkov and Sakhanenko (1980) の下界の精密化として, Koike (2006) はベイズリスクに対するバツチャリャ型の情報不等式を与えた.

定理 3.3. (Koike (2006)) 標本数 $n = 1$ のとき, 条件 (A1) と (A2) のもとで以下の不等式が成り立つ

$$B(q, \delta) \geq \left(\int_{\Theta} \frac{g'^2 q}{I_1} d\theta, - \int_{\Theta} \frac{g' g'' q}{I_1} d\theta \right) W^{-1} \left(\int_{\Theta} \frac{g'^2 q}{I_1} d\theta, - \int_{\Theta} \frac{g' g'' q}{I_1} d\theta \right)^T, \quad (3.5)$$

ただし, $W = \{E(S_i S_j)\}_{i,j=1,2}$ は 2×2 の行列であり

$$S_i = \frac{1}{f_0 q} \frac{\partial^i}{\partial \theta^i} \left(\frac{g' q f_0}{I_1} \right) \quad (i, j = 1, 2).$$

系 3.1. (Koike (2006)) 標本数が一般の $n \geq 1$ のとき, 条件 (A1) と (A2) のもとで (3.5) の漸近近似は次で与えられる

$$B(q, \delta) \geq C n^{-1} + (-D + E) n^{-2} + O(n^{-3}) \quad (n \rightarrow \infty) \quad (3.6)$$

ただし

$$E := \frac{\left[\int_{\Theta} \left\{ \frac{g'^2}{I_1^2} K_1 - \frac{g' g''}{I_1} \right\} q d\theta \right]^2}{2 \int_{\Theta} g'^2 q d\theta} \geq 0.$$

下界 (3.6) は下界 (3.2) を n^{-2} のオーダーで改良していることがわかる.

4 主結果

以後, 次の条件を仮定する.

(A3) ベイズ推定量 δ のバイアスは次のように展開される.

$$b(\theta) = E_{\theta}\{\delta(X)\} - g(\theta) = \frac{b_1(\theta)}{n} + O(n^{-2}) \quad (n \rightarrow \infty).$$

(A3) を仮定することにより, Brown and Gajek (1990) の下界の漸近近似は

$$B(q, \delta) \geq C n^{-1} + (R_1 + R_2) n^{-2} + O(n^{-3}) \quad (n \rightarrow \infty) \quad (4.1)$$

となる. ただし,

$$R_1 := 2 \int_{\Theta} \frac{b'_1 g' q}{I_1} d\theta, \quad R_2 := \int_{\Theta} b_1^2 q d\theta.$$

Brown and Gajek (1990) を精密化するバツチャリャ型の不等式として次を得る.

定理 4.1. 標本数 $n = 1$ のとき, 条件 (A1) と (A2) のもとで以下の不等式が成り立つ

$$B(q, \delta) \geq \int_{\Theta} (g' + b', g'' + b'') V_1^{-1} (g' + b', g'' + b'')^T q d\theta + \int_{\Theta} b^2 q d\theta. \quad (4.2)$$

系 4.1. 標本数 $n \geq 1$ のとき, 条件 (A1) から (A3) のもとで (4.2) の漸近近似は

$$B(q, \delta) \geq Cn^{-1} + (R_1 + R_2 + R_3) n^{-2} + O(n^{-3}) \quad (n \rightarrow \infty) \quad (4.3)$$

となる. ただし,

$$R_3 := \int_{\Theta} \frac{(g' K_1 - g'' I_1)^2}{I_1^4} q d\theta \geq 0.$$

(3.3) と (4.2) の差を明示的に与えるのは容易ではないが, 漸近下界では (4.1) と (4.3) の差を明示的に与えることができ, 実際に次のようになる

$$R_3 n^{-2} + O(n^{-3}) \geq 0 \quad (n \rightarrow \infty).$$

注意 4.1. Brown and Gajek (1990) の定理 2.7 と同様にして, (4.3) を b_1 に関して最小化すると次のようなバイアスに依存しない下界を得ることができる

$$B(q, \delta) \geq Cn^{-1} + (-D + R_3) n^{-2} + O(n^{-3}) \quad (n \rightarrow \infty).$$

特に $K_1 = 0$ とすると, 以下の系を得る.

系 4.2. 標本数 $n = 1$ のとき, 条件 (A1) と (A2) のもとで $K_1 = 0$ なら以下の不等式が成り立つ

$$B(g, \delta) \geq \frac{C^2}{C + D} + Q_1(b) + Q_2(b) + Q_3(b) + Q_4(b), \quad (4.4)$$

ただし,

$$Q_4(b) := \int_{\Theta} \frac{(g'' + b'')^2}{J_1} q d\theta \geq 0.$$

例 4.1. X を正規分布 $N(\theta, 1)$ に従う確率変数とする. θ の事前分布が正規分布 $N(\mu, \sigma^2)$ のとき, 二乗損失のもとでの $g(\theta) = \theta^2$ のベイズ推定問題を考える. X を与えたときの θ の事後分布は $N\left(\frac{X + (\mu/\sigma^2)}{1 + (1/\sigma^2)}, \frac{1}{1 + (1/\sigma^2)}\right)$ であることより, θ^2 のベイズ推定量は

$$\delta(X) = E(\theta^2 | X) = \frac{1}{1 + (1/\sigma^2)} + \left\{ \frac{X + (\mu/\sigma^2)}{1 + (1/\sigma^2)} \right\}^2$$

となる。さらに、簡単な計算より推定量のバイアスとベイズリスクは次のようになる

$$b(\theta) = \frac{\mu^2 + 2\sigma^4 + 2\sigma^2(\mu - \theta)\theta - \theta^2}{(1 + \sigma^2)^2},$$

$$B(q, \delta) = \frac{2\sigma^2(2\sigma^4 + 2\mu^2\sigma^2 + 2\mu^2 + \sigma^2)}{(1 + \sigma^2)^2}.$$

一方, $g'(\theta) = 2\theta$, $g''(\theta) = 2$, $I_1(\theta) = 1$, $K_1(\theta) = 0$ より定理 4.1 の下界は

$$(4.2) = \frac{2\sigma^2(2\sigma^4 + 2\mu^2\sigma^2 + 2\mu^2 + \sigma^2)}{(1 + \sigma^2)^2}.$$

この例では, 下界 (4.2) はベイズリスクを達成する.

次に, Koike (2006) による下界との比較のために, 標本数が $n \geq 1$ の場合について考える. $X_1, \dots, X_n$ を与えたときの θ の事後分布は $N\left(\frac{\sum_{i=1}^n X_i + (\mu/\sigma^2)}{n + (1/\sigma^2)}, \frac{1}{n + (1/\sigma^2)}\right)$ となるので, θ^2 のベイズ推定量は次のようになる

$$\delta(\mathbf{X}) = E(\theta^2 | \mathbf{X}) = \frac{1}{n + (1/\sigma^2)} + \left\{ \frac{\sum_{i=1}^n X_i + (\mu/\sigma^2)}{n + (1/\sigma^2)} \right\}^2.$$

バイアスとベイズリスクはそれぞれ次のようになる

$$b(\theta) = \frac{2(\mu\theta + \sigma^2 - \theta^2)}{\sigma^2}n^{-1} + o(n^{-1}),$$

$$B(q, \delta) = 4(\mu^2 + \sigma^2)n^{-1} + \frac{-2(2\mu^2 + 3\sigma^2)}{\sigma^2}n^{-2} + O(n^{-3}) \quad (n \rightarrow \infty).$$

よって, 系 4.1 より

$$(4.3) = 4(\mu^2 + \sigma^2)n^{-1} + \left\{ -4\left(2 + \frac{\mu^2}{\sigma^2}\right) + 2 \right\}n^{-2} + O(n^{-3})$$

as $n \rightarrow \infty$. このとき, ベイズリスクと (4.3) の差は $O(n^{-3})$ ($n \rightarrow \infty$) である. 一方で, 下界 (3.6) は

$$4(\mu^2 + \sigma^2)n^{-1} + \left\{ -4\left(2 + \frac{\mu^2}{\sigma^2}\right) + \frac{2\mu^2}{\mu^2 + \sigma^2} \right\}n^{-2} + O(n^{-3}), \quad (4.5)$$

となり, (4.3) と (4.5) の差は

$$\left(2 - \frac{2\mu^2}{\mu^2 + \sigma^2}\right)n^{-2} + O(n^{-3}) \quad (n \rightarrow \infty), \quad (4.6)$$

となり, $2 - \frac{2\mu^2}{\mu^2 + \sigma^2} \geq 0$ であることから, この例において (4.3) は (3.6) より n^{-2} のオーダーでよいということがわかる.

例 4.2. X をポアソン分布 $Po(\lambda)$ に従う確率変数とする. λ の事前分布がガンマ分布 $Ga(a, b)$ ($a, b > 0$) のとき, 二乗損失のもとでの $g(\theta) = e^{-\lambda}$ ($X = 0$ となる確率) のベイズ推定問題を考える, X を与えたときの λ の事後分布は $Ga\left(a + X, \frac{b}{1+b}\right)$ であることより, $e^{-\lambda}$ のベイズ推定量は次のようになる

$$\delta(X) = E(e^{-\lambda} | X) = \left(1 + \frac{b}{1+b}\right)^{-a-X}.$$

このようなベイズ推定量に対してベイズリスクと下界 (4.2) は解析的に導出するのは容易ではないので, 数値結果を以下に示す.

事前分布	(3.1)	(3.3)	(4.2)	ヘイズリスク
$Ga(1, 1)$	2.614×10^{-2}	4.635×10^{-2}	4.753×10^{-2}	4.761×10^{-2}
$Ga(1, 2)$	1.592×10^{-2}	3.971×10^{-2}	5.269×10^{-2}	5.285×10^{-2}
$Ga(2, 2)$	3.312×10^{-3}	1.403×10^{-2}	1.490×10^{-2}	1.507×10^{-2}
$Ga(4, 2)$	2.258×10^{-5}	9.054×10^{-4}	9.611×10^{-4}	9.785×10^{-4}

この例では, 下界 (4.2) はベイズリスクを達成しない.

5 局所ミニマックスリスクの下界

系 4.1 で与えた情報不等式の応用として, θ のミニマックス推定に対する漸近効率について考える. 条件 (A1)–(A3) のもとで次の局所ミニマックスリスクの二次漸近下界を得る.

定理 5.1. $\varepsilon > 0$ と $a > 0$ があり, $\theta \in (\theta_0 - \varepsilon, \theta_0 + \varepsilon)$ について

$$0 < a \leq |K_1(\theta)| \quad (5.1)$$

を満たしているとする. $(\theta_0 - \varepsilon, \theta_0 + \varepsilon) \subset \Theta$ であるとき, 条件 (A1)–(A3) のもとで, θ_0 におけるミニマックスリスクは, $n \rightarrow \infty$ のとき

$$\sup_{\theta \in (\theta_0 - \varepsilon, \theta_0 + \varepsilon)} E_{\theta} \left\{ (\hat{\theta} - \theta)^2 \right\} \geq \frac{1}{I^*} n^{-1} - \frac{\pi^2}{\varepsilon^2 I_*^2} n^{-2} + \frac{a^2}{2I_*^4} n^{-2} + O(n^{-3}) \quad (5.2)$$

となる. ただし, $\hat{\theta}$ は θ の推定量であり, $I^* = \sup_{\theta} I_1(\theta)$, $I_* = \inf_{\theta} I_1(\theta)$ とする.

注意 5.1. Borovkov (1998) もミニマックスリスクに関して次の下界を与えている

$$\sup_{\theta \in (\theta_0 - \varepsilon, \theta_0 + \varepsilon)} E_{\theta} \left\{ (\hat{\theta} - \theta)^2 \right\} \geq \frac{1}{I^*} n^{-1} - \frac{\pi^2}{\varepsilon^2 I_*^2} n^{-2} + O(n^{-3}) \quad (n \rightarrow \infty). \quad (5.3)$$

下界 (5.2) は (5.3) を n^{-2} のオーダーで改良していることがわかる。

参考文献

- [1] Borovkov, A. A. (1998). *Mathematical Statistics*. Amsterdam: Gordon and Breach.
- [2] Borovkov, A. A., Sakhanenko, A. I. (1980). Estimates for averaged quadratic risk (in Russian). *Probability and Mathematical Statistics* **1**: 185–195.
- [3] Blight, B. J. N., Rao, P. V. (1974). The convergence of Bhattacharyya bounds. *Biometrika* **61**: 137–142.
- [4] Brown, L. D., Gajek, L. (1990). Information inequalities for the Bayes risk. *Annals of Statistics* **18**: 1578–1594.
- [5] Ghosh, J. K. (1994). *Higher Order Asymptotics*. NSF–CBMS Regional Conference Series in Probability and Statistics, Vol.4, Inst. of Math. Statist., Hayward.
- [6] Ghosh, J. K., Sathe, Y. S. (1987). Convergence of Bhattacharyya bounds—revisited. *Sankhyā Ser. A* **49**: 37–42.
- [7] Khan, R. A. (1984). On UMVU estimators and Bhattacharyya bounds in exponential distributions. *Journal of Statistical Planning and Inference* **9**: 199–206.
- [8] Koike, K. (2006). An integral Bhattacharyya type bound for the Bayes risk. *Communications in Statistics— Theory and Methods* **35**: 2185–2195.
- [9] Lehmann, E. L., Casella, G. (1998). *Theory of Point Estimation, 2nd Ed.* New York: Springer.
- [10] Sato, M., Akahira, M. (1996). An information inequality for the Bayes risk. *Annals of Statistics* **24**: 2288–2295.
- [11] Takada, Y. (1999). Lower bounds on the Bayes risk for statistical prediction problems. *Communications in Statistics— Theory and Methods* **28**: 693–703.
- [12] Zacks, S. (1971). *The Theory of Statistical Inference*. New York: Wiley.