第7章　推定

この章では, 点推定法における推定量の導出原理と推定量の精度評価についての考え方の説明を行う. 第 7.1 節では, モーメント法について説明する. 第 7.2 節では, 最尤法について説明する. 最尤推定量の漸近分布は, 統計的モデルが指数型分布族の場合について示している. 第 7.3 では, 推定量のひとつの性質である不偏推定の概念を説明し, それに関係する情報不等式を与える.

7.1　モーメント法

この方法で得られた推定量は最適性を持たないことがある. しかし求めるのが容易であるので, 陽には式が与えられない別の推定量の値を繰り返し計算で求めるときの初期値として用いることができる.

$\mathbb{X} \subset \mathbb{R}$ を空でない部分集合とする. $d, n \in \mathbb{N}$ とし, 統計的実験を $\bigl(\mathbb{X}^n, \bigl\{\mathsf{P}_{\boldsymbol{\theta}} = \mathsf{P}_{\boldsymbol{\theta}};\, \boldsymbol{\theta} \in \Theta \subset \mathbb{R}^d\bigr\}\bigr)$ を考え

$$X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{P}_{\boldsymbol{\theta}^*} \quad (\boldsymbol{\theta}^* \in \Theta)$$

とする. ただし, $\mathsf{P}_{\boldsymbol{\theta}}^{\otimes n}$ と $\mathsf{P}_{\boldsymbol{\theta}}$ は, それぞれ測度空間 $\bigl(\mathbb{X}^n, \mathcal{B}(\mathbb{X}^n)\bigr)$ と $\bigl(\mathbb{X}, \mathcal{B}(\mathbb{X})\bigr)$ 上の分布である.

$n \geq d$ とする. $1 \leq j \leq d$ に対して

$$\alpha_j := \alpha_j(\boldsymbol{\theta}) := \mathsf{E}\bigl[X_1^j\bigr], \quad \widehat{\alpha}_j := \frac{1}{n}\sum_{\ell=1}^{n} X_\ell^j$$

と定める.

定義 7.1. $n \geq d$ とし $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{P}_{\boldsymbol{\theta}^*}\ \bigl(\boldsymbol{\theta}^* \in \Theta\bigr)$ とする. このとき $\boldsymbol{\theta}^*$ の**モーメント法推定量** $\widehat{\boldsymbol{\theta}}_n$ とは次をみたす解 (存在すれば) である.

$$\alpha_1(\widehat{\boldsymbol{\theta}}_n) = \widehat{\alpha}_1,\ \alpha_2(\widehat{\boldsymbol{\theta}}_n) = \widehat{\alpha}_2, \ldots,\ \alpha_d(\widehat{\boldsymbol{\theta}}_n) = \widehat{\alpha}_d \tag{7.1}$$

注意 7.2. $n \geq 2$ とする. $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{N}\bigl(\mu^*, \bigl(\sigma^*\bigr)^2\bigr)$ とする. ただし $\boldsymbol{\theta}^* := (\mu^*, \sigma^*) \in \mathbb{R} \times (0, \infty) =: \Theta$ で (μ^*, σ^*) は未知とする. このとき

$$\alpha_1 = \mathsf{E}\bigl[X_1\bigr] = \mu^*, \quad \alpha_2 = \mathsf{E}\bigl[X_1^2\bigr] = \mathsf{Var}\bigl[X_1\bigr] + \bigl\{\mathsf{E}\bigl[X_1\bigr]\bigr\}^2 = (\sigma^*)^2 + (\mu^*)^2$$

である．したがって $\widehat{\boldsymbol{\theta}}_n = (\widehat{\mu}_n,\, \widehat{\sigma}_n)$ とすれば

$$\widehat{\mu}_n = \frac{1}{n}\sum_{\ell=1}^{n} X_\ell, \quad \widehat{\sigma}_n^2 + \{\widehat{\mu}_n\}^2 = \frac{1}{n}\sum_{\ell=1}^{n} X_\ell^2$$

を解くと $(\mu^*,\, \sigma^*)$ のモーメント法推定量は

$$\widehat{\mu}_n = \frac{1}{n}\sum_{\ell=1}^{n} X_\ell, \quad \widehat{\sigma}_n = \sqrt{\frac{1}{n}\sum_{\ell=1}^{n}(X_\ell - \widehat{\mu}_n)}$$

となる．　□

定理 7.3. $d,\, n \in \mathbb{N}$ は $n \geq d$ とする．

$$X_1,\, X_2,\, \ldots,\, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{P}_{\boldsymbol{\theta}^*}\,(\boldsymbol{\theta}^* \in \Theta \subset \mathbb{R}^d), \quad \mathsf{E}\big[\big|X_1\big|^{2d}\big] < \infty$$

とする．(7.1) で定義した $\alpha_j\,(j = 1,\, 2,\, \ldots,\, d)$ の逆写像 α_j^{-1} が存在して，$\boldsymbol{\theta}^*$ の近傍で全微分可能とする．このとき適当な条件のもとで以下が成立する．

(1) $\mathsf{Pr}_{\boldsymbol{\theta}^*}\big(\, \widehat{\boldsymbol{\theta}}_n$ は存在$\big) \overset{n\to\infty}{\longrightarrow} 0$ である．

(2) $\forall \epsilon > 0$ に対して

$$\mathsf{Pr}_{\boldsymbol{\theta}^*}\big(\big|\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*\big|_{2,\,d} \geq \epsilon\big) \overset{n\to\infty}{\longrightarrow} 0$$

である．ただし $\big|\cdot\big|_{2,\,d}$ は $\mathbb{R}^d$ の Euclid ノルムである．

(3) $\sqrt{n}\big(\widehat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}^*\big) \rightsquigarrow \mathsf{N}_d\big(\mathbf{0},\, \boldsymbol{\Sigma}\big)$ である．ただし

$$\begin{aligned}
\boldsymbol{\Sigma} &= \boldsymbol{G}\big(\boldsymbol{\theta}^*\big)\mathsf{E}_{\boldsymbol{\theta}^*}\big[\boldsymbol{Y}\boldsymbol{Y}^\top\big]\boldsymbol{G}^\top\big(\boldsymbol{\theta}^*\big),\\
\boldsymbol{G}\big(\boldsymbol{\theta}\big) &= \big(\overset{\bullet}{\boldsymbol{g}}_1\,(\boldsymbol{\theta}),\, \overset{\bullet}{\boldsymbol{g}}_2\,(\boldsymbol{\theta}),\, \ldots,\, \overset{\bullet}{\boldsymbol{g}}_d\,(\boldsymbol{\theta})\big),\\
\overset{\bullet}{\boldsymbol{g}}_j\,(\boldsymbol{\theta}) &= \left(\frac{\partial\alpha_j^{-1}}{\partial\theta_1}(\boldsymbol{\theta}),\, \frac{\partial\alpha_j^{-1}}{\partial\theta_2}(\boldsymbol{\theta}),\, \ldots,\, \frac{\partial\alpha_j^{-1}}{\partial\theta_d}(\boldsymbol{\theta})\right)^\top (j = 1,\, 2,\, \ldots,\, d),\\
\boldsymbol{Y} &= \big(X_1,\, X_1^2,\, \ldots,\, X_1^d\big)^\top
\end{aligned}$$

が成立する．

Proof. 証明は加筆予定である．　□

7.2 最尤法

$\mathbb{X}$ を $\mathbb{R}$ の空でない部分集合とし，$d,\,n \in \mathbb{N}$ とする．統計的実験 $\left(\mathbb{X}^n,\,\left\{\mathsf{p}(x|\,\boldsymbol{\theta});\,\boldsymbol{\theta} \in \Theta\right\}\right)$ を考える．ただし $\mathsf{p}(x|\,\boldsymbol{\theta})$ は p.d.f.（または p.m.f.）で $\Theta \subset \mathbb{R}^d$ である．$\boldsymbol{\theta}^* \in \Theta$ とし

$$X_1,\,X_2,\,\ldots,\,X_n \overset{\text{i.i.d.}}{\sim} \mathsf{p}(x|\,\boldsymbol{\theta}^*)$$

とする．

定義 7.4. $X_1 = x_1,\,X_2 = x_2,\,\ldots,\,X_n = x_n$ を観測したとき $\boldsymbol{\theta}$ の **尤（ゆう）度関数** $\mathsf{lik}_n(\boldsymbol{\theta}|\,\boldsymbol{x})$ を

$$\mathsf{lik}_n\big(\boldsymbol{\theta}|\,\boldsymbol{x}\big) = \prod_{j=1}^{n} \mathsf{p}(x_j|\,\boldsymbol{\theta})$$

で定義し，**対数尤度** $\ell_n(\boldsymbol{\theta}|\,\boldsymbol{x})$ を

$$\ell_n(\boldsymbol{\theta}|\,\boldsymbol{x}) = \log \mathsf{lik}_n(\boldsymbol{\theta}|\,\boldsymbol{x})$$

で定義する．ただし，$\boldsymbol{x} = (x_1,\,x_2,\,\ldots,\,x_n)$ である．

注意 7.5. 尤度関数は

$$\mathsf{lik}_n(\,\cdot\,|\,\boldsymbol{x}) :\ \Theta \ni \boldsymbol{\theta} \mapsto \mathsf{lik}_n(\boldsymbol{\theta}|\,\boldsymbol{x}) \in [0,\,\infty)$$

である． □

記法: 関数 $g(x)$ の最大値を取る点を表す集合を

$$\arg\max_{x\in\mathbb{R}} g(x)$$

と書く．たとえば $g(x) = -(x-1)^2$ のとき

$$\arg\max_{x\in\mathbb{R}} g(x) = \{1\}$$

となる．$g(x) = \sin x$ のとき

$$\arg\max_{0\le x\le 4\pi} g(x) = \{\pi/2,\,5\pi/2\}$$

となる．

定義 7.6. $X_1 = x_1,\,X_2 = x_2,\,\ldots,\,X_n = x_n$ を観測したとき $\boldsymbol{\theta}^*$ の**最尤推定値** (maximum likelihood estimate) を $\mathsf{lik}_n(\boldsymbol{\theta}|\,\boldsymbol{x})$ を最大にする値 $\widehat{\boldsymbol{\theta}}_n(\boldsymbol{x})$ で定義する．すなわち

$$\widehat{\boldsymbol{\theta}}_n(\boldsymbol{x}) \in \arg\max_{\boldsymbol{\theta}\in\Theta} \mathsf{lik}_n(\boldsymbol{\theta}|\,\boldsymbol{x})$$

である．$\boldsymbol{x} = (x_1,\,x_2,\,\ldots,\,x_n)$ に $\boldsymbol{X} = (X_1,\,X_2,\,\ldots,\,X_n)$ を代入したもの $\widehat{\boldsymbol{\theta}}_n(\boldsymbol{X})$ を $\boldsymbol{\theta}^*$ の**最尤推定量** (maximum likelihood estimator=m.l.e.) という．

例 7.7. $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{Ber}(\theta^*)$ とする．ただし $(0, 1) =: \Theta \ni \theta^*$ は未知とする．すなわち

$$\mathsf{p}(x|\,\theta^*) = \begin{cases} (\theta^*)^x\bigl(1-\theta^*\bigr)^{1-x} & (x = 0, 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

である．$X_j = x_j\,(j = 1, 2, \ldots, n)$ を観測したとき

$$\mathsf{lik}_n\bigl(\theta|\,\boldsymbol{x}\bigr) = \prod_{j=1}^{n} \mathsf{p}(x_j|\,\theta) = \prod_{j=1}^{n} \theta^{x_j}(1-\theta)^{1-x_j} = \theta^{t_n}(1-\theta)^{n-t_n}$$

となる．ただし, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ と $t_n = \sum_{j=1}^{n} x_j$ である．よって対数尤度は

$$\ell_n(\theta|\,\boldsymbol{x}) = t_n \log\theta + (n - t_n)\log(1-\theta) \qquad (0 < \theta < 1)$$

となる．このことから, $0 < t_n < n$ のとき

$$\frac{t_n}{n} \in \arg\max_{\theta\in(0,\,1)} \ell_n(\theta|\,\boldsymbol{x})$$

がわかる．したがって, $0 < t_n < n$ のとき θ^* の最尤推定値は $\widehat{\theta}_n = \dfrac{\sum_{j=1}^{n} x_j}{n}$ となり, $t_n = 0$ または $t_n = n$ のとき, 最尤推定値は存在しない． □

注意 7.8. 定義 7.6 は [3] の流儀に従った．一方, 母数空間 Θ の閉包 $\mathrm{cl}(\Theta)$ を考え, $\mathrm{cl}(\Theta)$ に尤度関数 lik_n の定義域を拡張して

$$\widehat{\boldsymbol{\theta}}_n(\boldsymbol{x}) \in \arg\min_{\boldsymbol{\theta}\in\mathrm{cl}(\Theta)} \mathsf{lik}_n(\boldsymbol{\theta}|\,\boldsymbol{x})$$

と最尤推定値を定義する流儀もある．[29, 41] を参照のこと．例 **??** を踏まえるとこちらの定義のが数学的は扱いやすくなるようにみえる．

例 7.9. $n \geq 2\,(n \in \mathbb{N})$ とする．$\boldsymbol{\theta}^* = (\mu^*, \sigma^*) \in \mathbb{R}\times(0, \infty)$ とし

$$X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{N}\bigl(\mu^*, (\sigma^*)^2\bigr)$$

とする．$X_j = x_j\,(j = 1, 2, \ldots, n)$ を観測したとき, 尤度関数は

$$\begin{aligned}\mathsf{lik}_n(\mu, \sigma|\,\boldsymbol{x}) &= \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\sigma}\exp\Biggl\{-\frac{(x_j-\mu)^2}{\sigma^2}\Biggr\} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\sigma^n}\exp\Biggl\{-\frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_j-\mu)^2\Biggr\} \\ &= \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\sigma^n}\exp\Biggl\{-\frac{ns_n^2}{2\sigma^2}\Biggr\}\exp\Biggl\{-\frac{n(\overline{x}_n-\mu)^2}{2\sigma^2}\Biggr\} \qquad (7.2)\end{aligned}$$

となる. ただし

$$\overline{x}_n = \frac{1}{n}\sum_{j=1}^{n} x_j, \qquad s_n^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \overline{x}_n)^2, \qquad \sum_{j=1}^{n}(x_j - \overline{x}_n)^2 \neq 0$$

である. (7.2) の最後の等号は

$$\sum_{j=1}^{n}(x_j - \mu)^2 = ns_n^2 + n(\overline{x}_n - \mu)^2 \tag{7.3}$$

からわかる. 対数尤度は

$$\ell_n(\mu,\, \sigma|\, \boldsymbol{x}) = -n\log\sigma - \frac{ns_n^2}{2\sigma^2} - \frac{n(\overline{x}_n - \mu)^2}{2\sigma^2} + (\text{定数項})$$

となる. よって

$$\begin{cases} \dfrac{\partial \ell_n}{\partial \mu}(\mu,\, \sigma|\, \boldsymbol{x}) = -\dfrac{n(\overline{x}_n - \mu)}{2\sigma^2} = 0 \\ \dfrac{\partial \ell_n}{\partial \sigma}(\mu,\, \sigma|\, \boldsymbol{x}) = -\dfrac{n}{\sigma} + \dfrac{ns_n^2}{\sigma^3} + \dfrac{n(\overline{x}_n - \mu)^2}{\sigma^3} = 0 \end{cases}$$

を解くと

$$\mu = \overline{x}_n, \qquad \sigma = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\big(x_j - \overline{x}_n\big)^2}$$

となる. $\ell_n(\mu,\, \sigma|\, \boldsymbol{x})$ の Hessian を求める.

$$\boldsymbol{H} := \begin{pmatrix} \dfrac{\partial \ell_n^2}{\partial \mu^2}\big(\overline{x}_n,\, s_n|\, \boldsymbol{x}\big) & \dfrac{\partial \ell_n^2}{\partial \mu\partial\sigma}\big(\overline{x}_n,\, s_n|\, \boldsymbol{x}\big) \\ \dfrac{\partial \ell_n^2}{\partial\partial\sigma\mu}\big(\overline{x}_n,\, s_n|\, \boldsymbol{x}\big) & \dfrac{\partial \ell_n^2}{\partial \sigma^2}\big(\overline{x}_n,\, s_n|\, \boldsymbol{x}\big) \end{pmatrix} = \begin{pmatrix} -\dfrac{1}{2s_n^2} & 0 \\ 0 & -\dfrac{2n}{s_n^2} \end{pmatrix} \tag{7.4}$$

より, $-\boldsymbol{H}$ は正定値行列となる. したがって, 関数

$$\Theta \ni (\mu,\, \sigma) \mapsto \ell_n(\mu,\, \sigma|\, \boldsymbol{x})$$

は $(\mu,\, \sigma) = (\overline{x}_n,\, s_n)$ で最大となる. 以上の議論から $(\mu^*,\, \sigma^*)$ の最尤推定量は

$$\widehat{\mu}_n = \frac{1}{n}\sum_{j=1}^{n} X_j =: \overline{X}_n, \quad \widehat{s}_n = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\big(X_j - \overline{X}_n\big)^2}$$

となる. □

注意 7.10. $g: \mathbb{R} \times [0, \infty) \longrightarrow [0, \infty)$ は 1 対 1 とする．すると $\widehat{\boldsymbol{\theta}}_n$ が $\boldsymbol{\theta}^*$ の最尤推定量ならば, $g\bigl(\widehat{\boldsymbol{\theta}}_n\bigr)$ は $g(\boldsymbol{\theta}^*)$ の最尤推定量であることがわかる．すなわち, 最尤推定量は母数の変換に関して不変である．このことに注意すれば, 例 $refex:7-8$ の結果から $(\widehat{\mu}_n,\, s_n^2)$ は $(\mu^*,\, (\sigma^*)^2)$ の最尤推定量になることがわかる． □

問 7.1. (7.3) と (7.4) を確認せよ．

定理 7.11.

$$X_1,\, X_2,\, \ldots,\, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{p}(x|\,\eta) = h(x)\exp\{\eta^* T(x) - \kappa(\eta^*)\}\,(\eta^* \in \mathcal{E}^o)$$

とする．ただし $\mathcal{E}^o$ は自然母数空間 $\mathcal{E} \subset \mathbb{R}$ の内部である．さらに $\ddot{\kappa}\,(\eta^*) > 0\,(\eta^* \in \mathcal{E}^o)$ を仮定する．このとき η^* の最尤推定量 $\widehat{\eta}_n$ は十分大きな n に対して確率 1 で一意的に存在[1]して

$$\widehat{\eta}_n \xrightarrow{\text{a.s.}} \eta^*,$$
$$\sqrt{n}\bigl(\widehat{\eta}_n - \eta^*\bigr) \rightsquigarrow \mathsf{N}\biggl(0,\, \frac{1}{\ddot{\kappa}\,(\eta^*)}\biggr)$$

が成立する．

Proof. $X_1 = x_1,\, X_2,\, \ldots,\, X_n = x_n$ を観測したときの尤度関数 lik_n は

$$\mathsf{lik}_n(\eta|\,\boldsymbol{x}) = \biggl\{\prod_{j=1}^n h(x_j)\biggr\}\exp\Bigl[\eta\sum_{j=1}^n T(x_j) - n\kappa(\eta)\Bigr]$$

となる．このことから対数尤度関数は

$$\ell_n(\eta|\,\boldsymbol{x}) = \log\bigl(\mathsf{lik}_n(\eta|\,\boldsymbol{x})\bigr) = n\{\eta\overline{T}_n - \kappa(\eta)\} + (\text{constant})$$

となる．ただし $\overline{T}_n = n^{-1}\sum_{j=1}^n T(x_j)$ である．したがって

$$\dot{\ell}_n\,(\eta|\,\boldsymbol{x}) = \frac{d\ell_n}{d\eta}(\eta|\,\boldsymbol{x}) = n\bigl(\overline{T}_n - \dot{\kappa}\bigr) = 0 \Leftrightarrow \overline{T}_n = \dot{\kappa}\,(\eta) = \mathsf{E}[T(X_1)]$$

となる．最後の等号は (6.4) からわかる．

[1] これは不明瞭な表現である．意味するところは, $A_n = \{\widehat{\eta}_n$ は存在 $\}$ としたとき

$$\lim_{n\to\infty}\mathsf{Pr}\biggl(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k\biggr) = 1$$

が成立することである．

以後, 対数尤度関数の $x_j\,(j=1,2,\ldots,n)$ のところに X_j を代入して, 確率変数にしたものを考える. $T(X_1)$ は有限の期待値をもつので, 大数の法則 (定理 4.17) から

$$\overline{T}_n = \frac{1}{n}\sum_{j=1}^{n} T(X_j) \xrightarrow{\text{a,s.}} \mathsf{E}[T(X_1)] = \dot{\kappa}(\eta^*) \quad (n\to\infty) \tag{7.5}$$

がわかる. η^* の適当な近傍を $U(\eta^*)$ と書き, $\dot{\kappa}(U(\eta^*)) := \{\dot{\kappa}(\eta);\, \eta \in U(\eta^*)\}$ とおく. (7.5) から

$$\mathsf{Pr}\big(\lim_{n\to\infty} \overline{T}_n \in \dot{\kappa}(U(\eta^*))\big) = 1$$

となる. さらに, $\ddot{\kappa}(\eta^*) > 0$ なので, $\ddot{\kappa}(\eta) > 0\,(\eta \in U(\eta^*))$ となる. したがって, 関数 $\mathcal{E}^o \ni \eta \mapsto \dot{\kappa}(\eta)$ は近傍 $U(\eta^*)$ 上で狭義単調増加関数である. このことより

$$\overline{T}_n = \dot{\kappa}(\eta)$$

をみたす $\eta \in U(\eta^*) \subset \mathcal{E}^o$ が存在する. したがって, 近傍 $\kappa(U(\eta^*))$ 上で $\dot{\kappa}$ の逆写像 $\dot{\kappa}^{-1}$ が存在する. このことを踏まえて

$$\widehat{\eta}_n = \dot{\kappa}^{-1}\big(\overline{T}_n\big)$$

とおく. さらに

$$\overline{T}_n \xrightarrow{\text{a.s.}} \dot{\kappa}(\eta^*) \quad \text{かつ} \quad \dot{\kappa}^{-1}(t) \text{ は近傍 } \dot{\kappa}\big(U(\eta^*)\big) \text{ 上で連続}$$

なので定理 4.14(6) から

$$\widehat{\eta}_n = \dot{\kappa}^{-1}\big(\overline{T}_n\big) \xrightarrow{\text{a.s.}} \dot{\kappa}^{-1}\big(\dot{\kappa}(\eta^*)\big) = \eta^*$$

がわかる. 一方, (6.4) から $\mathsf{E}[T(X_1)] = \dot{\kappa}(\eta^*)$, $\mathsf{Var}[T(X_1)] = \dot{\kappa}(\eta^*)$ となることに注意して, 中心極限定理 (定理 4.22) を用いると

$$\sqrt{n}\big(\overline{T}_n - \dot{\kappa}(\eta^*)\big) \rightsquigarrow \mathsf{N}\big(0,\, \ddot{\kappa}(\eta^*)\big)$$

がわかる. 以上のことを踏まえて, $\widehat{\eta}_n = \dot{\kappa}^{-1}\big(\overline{T}_n\big)$ に対して, デルタ法 (定理 4.27) を適用すると

$$\sqrt{n}\big(\widehat{\eta}_n - \eta^*\big) \rightsquigarrow \mathsf{N}\left(0,\, \frac{1}{\ddot{\kappa}(\eta^*)}\right)$$

がわかる. □

7.3 不偏推定と情報不等式

$\mathbb{X} \subset \mathbb{R}$ を空でない部分集合とする. $(\Omega, \mathcal{A}, \mathsf{Pr})$ を確率空間とし, $X_1, X_2, \ldots, X_n$ をこの空間上の独立同一分布に従う $\mathbb{X}$ 値確率変数列 (ランダム標本) とする. また $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^\top$ と記し, $\boldsymbol{X}$ の値域 (標本空間) を $\mathbb{X}^n \subset \mathbb{R}^n$ と表記する. さらに

$$\mathsf{P} = \mathsf{Pr} \circ X_1^{-1}, \quad \mathsf{P}^{\otimes n} = \mathsf{Pr} \circ \boldsymbol{X}^{-1}$$

とおく. すると $\mathsf{P}^{\otimes n} = \underbrace{\mathsf{P} \times \mathsf{P} \times \cdots \times \mathsf{P}}_{n\text{ 個}}$ となっている.

$\mathcal{P} = \{\mathsf{P}_\theta;\, \theta \in \Theta \subset \mathbb{R}\}$ を $\mathbb{X}$ 上の正則母数モデルとする. 以下では, $\Theta \subset \mathbb{R}$ とし, P_θ に関する期待値と分散を $\mathsf{E}_\theta[\cdot]$, $\mathsf{Var}_\theta[\cdot\,]$ と表記する.

定義 7.12. $\boldsymbol{X} := (X_1, X_2, \ldots, X_n)^\top \sim \mathsf{P}_\theta^{\otimes n}\,(\theta \in \Theta \subset \mathbb{R})$ とする. 統計量 $T(\boldsymbol{X})$ が $\forall\theta \in \Theta$ に対して

$$\mathsf{E}_\theta\big[T(\boldsymbol{X})\big] = \theta$$

をみたすとき $T(\boldsymbol{X})$ は θ の**不偏推定量** (unbiased estimator) という.

例 7.13. $n \geq 2$ とし, $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{N}(\theta, 1)\,(\theta \in \mathbb{R})$ とする. このとき $\overline{X}_n = \dfrac{1}{n}\sum_{j=1}^n X_j$ は θ の不偏推定量である. □

定理 7.14 (Rao-Blackwell の定理)**.** $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)^\top \sim \mathsf{P}_\theta^{\otimes n}\,(\theta \in \Theta \subset \mathbb{R})$ で $T = T(\boldsymbol{X})$ は θ の十分統計量とする. $\widehat{\theta}_n(\boldsymbol{X})$ は θ の任意の不偏推定量で, 有限の分散 $\mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]$ をもつものとする. 推定量 $\widetilde{\theta}_n(T)$ を

$$\widetilde{\theta}_n(T) := \mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})|\,T\big]$$

により定めたとき, 次の (1), (2) が成立する.

(1) $\widetilde{\theta}_n(T)$ は θ の不偏推定量である.

(2) $\mathsf{Var}_\theta\big[\widetilde{\theta}_n(T)\big] \leq \mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]$ $(\forall\theta \in \Theta)$ が成立する.

Proof. (1) T は十分統計量なので, $\mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})|\,T\big]$ は θ に依存しないので $\widetilde{\theta}_n(T)$ は推定量となる. また定理 2.28(2) より

$$\mathsf{E}_\theta\big[\widetilde{\theta}_n(T)\big] = \mathsf{E}_\theta\big[\mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})|\,T\big]\big] = \mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big] = \theta \qquad (\theta \in \Theta)$$

となり $\widetilde{\theta}_n(T)$ は θ の不偏推定量であることが示せた.

(2) 定理 2.28(4) より

$$\begin{aligned}\mathsf{E}_\theta\big[\big(\widehat{\theta}_n(\boldsymbol{X})-\widetilde{\theta}_n(T)\big)\widetilde{\theta}_n(T)\big|T\big] &= \widetilde{\theta}_n(T)\mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})-\widetilde{\theta}_n(T)|T\big]\\ &= \widetilde{\theta}_n(T)\bigg(\mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})|T\big]-\widetilde{\theta}_n(T)\bigg)\\ &= \widetilde{\theta}_n(T)\big(\widetilde{\theta}_n(T)-\widetilde{\theta}_n(T)\big)=0\end{aligned}$$

となり

$$\mathsf{E}_\theta\big[\big(\widehat{\theta}_n(\boldsymbol{X})-\widetilde{\theta}_n(T)\big)\widetilde{\theta}_n(T)\big]=\mathsf{E}_\theta\big[\mathsf{E}_\theta\big[\big(\widehat{\theta}_n(\boldsymbol{X})-\widetilde{\theta}_n(T)\big)\widetilde{\theta}_n(T)\big|T\big]\big]=0 \tag{7.6}$$

を得る．一方

$$\begin{aligned}\mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big] &= \mathsf{E}_\theta\big[\big\{\widehat{\theta}_n(\boldsymbol{X})-\mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]\big\}^2\big]\\ &= \mathsf{E}_\theta\big[\big\{\widehat{\theta}_n(\boldsymbol{X})-\theta\big\}^2\big]\\ &= \mathsf{E}_\theta\big[\big\{\big(\widehat{\theta}_n(\boldsymbol{X})-\widetilde{\theta}_n(T)\big)+\big(\widetilde{\theta}_n(T)-\theta\big)\big\}^2\big]\\ &= \mathsf{E}_\theta\big[\big\{\widehat{\theta}_n(\boldsymbol{X})-\widetilde{\theta}_n(T)\big\}^2\big]+\mathsf{E}_\theta\big[\big\{\widetilde{\theta}_n(T)-\theta\big\}^2\big]\\ &\quad+2\mathsf{E}_\theta\big[\big(\widehat{\theta}_n(\boldsymbol{X})-\widetilde{\theta}_n(T)\big)\big(\widetilde{\theta}_n(T)-\theta\big)\big]\end{aligned}$$

である．しかし, (7.6) および $\widehat{\theta}_n(\boldsymbol{X})$ と $\widetilde{\theta}_n(T)$ は θ の不偏推定量であることに注意すると

$$\begin{aligned}&\mathsf{E}_\theta\big[\big(\widehat{\theta}_n(\boldsymbol{X})-\widetilde{\theta}_n(T)\big)\big(\widetilde{\theta}_n(T)-\theta\big)\big]\\ &\quad=\underbrace{\mathsf{E}_\theta\big[\big(\widehat{\theta}_n(\boldsymbol{X})-\widetilde{\theta}_n(T)\big)\widetilde{\theta}_n(T)\big]}_{=0\quad\because(7.6)}-\theta\underbrace{\big\{\mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]-\mathsf{E}_\theta\big[\widetilde{\theta}_n(T)\big]\big\}}_{=0}\\ &\quad=0\end{aligned}$$

がわかる．以上から

$$\begin{aligned}\mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big] &= \mathsf{E}_\theta\big[\big\{\widehat{\theta}_n(\boldsymbol{X})-\widetilde{\theta}_n(T)\big\}^2\big]+\mathsf{Var}_\theta\big[\widetilde{\theta}_n(T)\big]\\ &\geq \mathsf{Var}_\theta\big[\widetilde{\theta}_n(T)\big]\qquad(\theta\in\Theta)\end{aligned}$$

を得る． □

定義 7.15. $\boldsymbol{X}=(X_1, X_2, \ldots, X_n)^\top \sim \mathsf{P}_\theta^{\otimes n}$ $(\theta\in\Theta\subset\mathbb{R})$ とする．θ の任意の不偏推定量 $\widetilde{\theta}_n(\boldsymbol{X})$ に対して

$$\mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]\leq\mathsf{Var}_\theta\big[\widetilde{\theta}_n(\boldsymbol{X})\big]\qquad(\theta\in\Theta)$$

をみたす θ の不偏推定量 $\widehat{\theta}_n(\boldsymbol{X})$ が存在するとき, $\widehat{\theta}_n(\boldsymbol{X})$ を θ の**一様最小分散不偏推定量** (uniformly minimum variance unbiased estimator: UMVUE) という．

注意 7.16. 存在すれば, 一意的である. 演習問題 7.7 を参照のこと.

可測関数 $g : \mathbb{R} \to \mathbb{R}$ は連続微分可能とする. $\boldsymbol{X} \sim \mathsf{P}_\theta^{\otimes n}$ に基づき $g(\theta)$ の推定問題を考える.

定理 7.17. $\{\mathsf{P}_\theta; \theta \in \Theta \subset R\}$ は正則母数モデルとする. $\boldsymbol{X} \sim \mathsf{P}_\theta^{\otimes n}$ $(\theta \in \Theta \subset \mathbb{R})$ とし $\boldsymbol{X}$ は同時 p.d.f. または p.m.f. $\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)$ をもつとする. $\widehat{\theta}_n = \widehat{\theta}_n(\boldsymbol{X})$ を $g(\theta)$ の任意の不偏推定量とし

$$
\begin{aligned}
A(\widetilde{\theta}, \theta) &:= \mathsf{Var}_\theta\left[\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\theta)}\right] \qquad (\widetilde{\theta}, \theta \in \Theta) \qquad (7.7)\\
A(\widetilde{\theta}, \theta) &> 0 \qquad (\widetilde{\theta} \in \Theta, \widetilde{\theta} \neq \theta)
\end{aligned}
$$

とおく. このとき

$$
\mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big] \geq \sup_{\widetilde{\theta} \in \Theta} \frac{\{g(\widetilde{\theta}) - g(\theta)\}^2}{A(\widetilde{\theta}, \theta)}
$$

が成り立つ.

Proof. $\boldsymbol{X}$ は連続型確率変数の場合の証明を与える. $\boldsymbol{X}$ の値域を $\mathbb{X}^n$ と書く. さらに, $\widetilde{\mathbb{X}}^n = \{\boldsymbol{x} \in \mathbb{X}^n; \mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta) > 0\}$ とおく. また, 離散型の場合は, 積分記号を和の記号に変更すればよい. 推定量 $\widehat{\theta}_n$ は $g(\theta)$ に対して不偏なので

$$
\mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big] = \int_{\mathbb{X}^n} \widehat{\theta}_n(\boldsymbol{x}) \mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)\,\mathrm{d}\boldsymbol{x} = g(\theta) \qquad (\theta \in \Theta)
$$

となる. よって

$$
\begin{aligned}
\mathsf{E}_\theta\left[\widehat{\theta}_n(\boldsymbol{X})\left\{\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\theta)} - 1\right\}\right] &= \int_{\widetilde{\mathbb{X}}^n} \widehat{\theta}_n(\boldsymbol{x})\left\{\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)} - 1\right\}\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)\,\mathrm{d}\boldsymbol{x}\\
&= \int_{\mathbb{X}^n} \widehat{\theta}_n(\boldsymbol{x}) \mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\widetilde{\theta})\,d\boldsymbol{x} - \int_{\mathbb{X}^n} \widehat{\theta}_n(\boldsymbol{x}) \mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)\,\mathrm{d}\boldsymbol{x}\\
&= g(\widetilde{\theta}) - g(\theta) \qquad (7.8)
\end{aligned}
$$

となる. 一方

$$
\mathsf{E}_\theta\left[\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\theta)}\right] = \int_{\widetilde{\mathbb{X}}^n} \frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)} \mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)\,\mathrm{d}\boldsymbol{x} = \int_{\mathbb{X}^n} \mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\widetilde{\theta})\,\mathrm{d}\boldsymbol{x} = 1 \quad (7.9)
$$

と (7.8) に注意すると

$$
\begin{aligned}
g(\widetilde{\theta}) - g(\theta) &= \mathsf{E}_\theta\left[\widehat{\theta}_n(\boldsymbol{X})\left\{\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\theta)} - 1\right\}\right] \\
&= \mathsf{E}_\theta\left[\left\{\widehat{\theta}_n(\boldsymbol{X}) - \mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big] + \mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]\right\}\left\{\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\theta)} - 1\right\}\right] \\
&= \mathsf{E}_\theta\left[\left\{\widehat{\theta}_n(\boldsymbol{X}) - \mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]\right\}\left\{\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\theta)} - 1\right\}\right] \\
&\qquad + \mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]\,\underbrace{\mathsf{E}_\theta\left[\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\theta)} - 1\right]}_{=0} \\
&= \mathsf{E}_\theta\left[\big\{\widehat{\theta}_n(\boldsymbol{X}) - \mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]\big\}\left\{\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\theta)} - 1\right\}\right]
\end{aligned}
$$

を得る．Cauchy-Schwarz の不等式 (定理 3.10) を上の式の最右辺に適用すると

$$
\begin{aligned}
\big|g(\widetilde{\theta}) - g(\theta)\big| &= \left|\mathsf{E}_\theta\left[\widehat{\theta}_n(\boldsymbol{X})\left\{\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)} - 1\right\}\right]\right| \\
&\le \sqrt{\big\{\mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X}) - \mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]\big\}^2\big]}\sqrt{\mathsf{E}_\theta\left[\left\{\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)} - 1\right\}^2\right]} \\
&= \sqrt{\mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]}\sqrt{\mathsf{Var}_\theta\left[\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)}\right]} \qquad (7.10)
\end{aligned}
$$

を得る．最後の等号は (7.9) を用いた．したがって

$$
\mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]\mathsf{Var}_\theta\left[\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)}\right] \ge \{g(\widetilde{\theta}) - g(\theta)\}^2 \qquad (7.11)
$$

を得る．したがって (7.7) と (7.11) から $\forall\theta \in \Theta$ に対して

$$
\mathsf{Var}_\theta\big[T(\boldsymbol{X})\big] \ge \frac{\{g(\widetilde{\theta}) - g(\theta)\}^2}{A\big(\widetilde{\theta},\, \theta\big)}
$$

となる．上式の $\widetilde{\theta}$ は任意だったので，上式の左辺において $\widetilde{\theta}$ に関して sup を取ると定理の主張はわかる． □

定理 7.18. (Cramér-Rao の不等式) 次の条件を仮定する．

(1) $\{\mathsf{P}_\theta;\, \theta \in \Theta \subset R\}$ は正則母数モデルとする．

(2) $g : \mathbb{R} \longrightarrow \mathbb{R}$ は連続微分可能でその導関数は $\overset{\bullet}{g}(\theta) \neq 0\,(\theta \in \Theta)$ をみたす．さらに

$$
\lim_{\widetilde{\theta}\to\theta} \frac{A\big(\widetilde{\theta},\, \theta\big)}{(\widetilde{\theta} - \theta)^2} = J(\theta) > 0
$$

が存在する. ただし, $A\big(\widetilde{\theta},\theta\big)$ は (7.7) で与えたものである.

(3) $\forall\theta\in\Theta$ に対して十分小さな $\epsilon>0$ をとると任意の $\widetilde{\theta}\in\big\{\psi\in\Theta;\,\big|\psi-\theta\big|<\epsilon\big\}$ に対して, ある関数 $G(\boldsymbol{x}|\theta)$ が存在して

$$\left|\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\widetilde{\theta})-\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)}{(\widetilde{\theta}-\theta)\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)}\right|<G(\boldsymbol{x}|\,\theta)\quad \text{かつ}\quad \mathsf{E}_\theta\big[G^2(\boldsymbol{X}|\,\theta)\big]<\infty$$

をみたす.

このとき $g(\theta)$ の任意の不偏推定量 $T(\boldsymbol{X})$ に対して

$$\mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]\geq\frac{\{\overset{\bullet}{g}(\theta)\}^2}{\mathcal{F}_{\boldsymbol{X}}(\theta)}\qquad(\theta\in\Theta)\tag{7.12}$$

が成り立つ. ただし

$$\mathcal{F}_{\boldsymbol{X}}(\theta)=\mathsf{E}_\theta\left[\left\{\frac{\partial}{\partial\theta}\log\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{X}|\,\theta)\right\}^2\right]$$

である.

注意 7.19. (7.12) を Cramér-Rao の不等式といい, その右辺を Cramér-Rao の下限という. □

<u>**定理** 7.18 **の証明**</u>: $\boldsymbol{X}$ の値域を $\mathbb{X}^n$ とし, $\widetilde{\mathbb{X}}^n=\big\{\boldsymbol{x}\in\mathbb{X}^n;\,\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)>0\big\}$ とおく. 定理 7.17 と仮定 (2) から

$$\mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big]\geq\lim_{\widetilde{\theta}\to\theta}\frac{\left\{\dfrac{g(\widetilde{\theta})-g(\theta)}{\widetilde{\theta}-\theta}\right\}^2}{\dfrac{A\big(\widetilde{\theta},\theta\big)}{(\widetilde{\theta}-\theta)^2}}=\frac{\{\overset{\bullet}{g}(\theta)\}^2}{J(\theta)}\tag{7.13}$$

を得る. 次に仮定 (3) に注意して Lebesgue の優収束定理を用いると

$$\begin{aligned}
J(\theta)&=\lim_{\widetilde{\theta}\to\theta}\frac{A\big(\widetilde{\theta},\theta\big)}{\big(\widetilde{\theta}-\theta\big)^2}\\
&=\lim_{\widetilde{\theta}\to\theta}\int_{\widetilde{\mathbb{X}}^n}\left\{\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\widetilde{\theta})}{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)}-1\right\}^2\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)}{\big(\widetilde{\theta}-\theta\big)^2}\,\mathrm{d}\boldsymbol{x}\\
&=\int_{\widetilde{\mathbb{X}}^n}\left\{\lim_{\widetilde{\theta}\to\theta}\frac{\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\widetilde{\theta})-\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)}{\big(\widetilde{\theta}-\theta\big)\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)}\right\}^2\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)\,\mathrm{d}\boldsymbol{x}\\
&=\int_{\widetilde{\mathbb{X}}^n}\left\{\frac{\partial}{\partial\theta}\log\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)\right\}^2\mathsf{p}^{\boldsymbol{X}}(\boldsymbol{x}|\,\theta)\,\mathrm{d}\boldsymbol{x}=\mathcal{F}_{\boldsymbol{X}}(\theta)
\end{aligned}$$

となる．これと (7.13) を合わせると

$$\mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big] \geq \frac{\{\overset{\bullet}{g}(\theta)\}^2}{\mathcal{F}_{\boldsymbol{X}}(\theta)} \tag{7.14}$$

を得る．

□

注意 7.20. 指数型分布族が定理 7.18 の正則条件 (2) と (3) をみたしているかを確認する．

以下, 加筆をすること．

□

定義 7.21. $\boldsymbol{X} \sim \mathsf{P}_\theta^{\otimes n}$ $(\theta \in \Theta \subset \mathbb{R})$ とする．θ の不偏推定量 $\widehat{\theta}_n(\boldsymbol{X})$ は R 有効[2]であるとは

$$\mathsf{Var}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big] = \frac{1}{\mathcal{F}_{\boldsymbol{X}}(\theta)} \qquad (\theta \in \Theta)$$

をみたすときをいう．

定理 7.22. $\boldsymbol{X} = (X_1, X_2, \ldots, X_n) \sim \mathsf{P}_\theta^{\otimes n}$ $(\theta \in \Theta \subset \mathbb{R})$ とする．θ の不偏推定量 $\widehat{\theta}_n$ が R 有効であるための必要十分条件は $\mathbb{R}$ 上の $\mathbb{R}$ 値関数 $T(x)$, $A(\theta)$, $\kappa(\theta)$, $g(x)$ があって

$$\widehat{\theta}_n(\boldsymbol{X}) = \frac{1}{n}\sum_{j=1}^{n} T(X_j) \tag{7.15}$$

$$\log \mathsf{p}(x|\,\theta) = A(\theta)T(x) - \kappa(\theta) + g(x), \tag{7.16}$$

$$\int_{\mathbb{X}} T(x)\mathsf{p}(x|\,\theta)\,\mathrm{d}x = \theta \tag{7.17}$$

をみたすときである．ただし X_1 の p.d.f.(または p.m.f.) と値域をそれぞれ $\mathsf{p}(x|\,\theta)$ と $\mathbb{X}(\subset \mathbb{R})$ と書いた．

Proof. 連続型分布に対する証明を与える．離散型分布に対しては, 積分記号を和の記号に替えればよい．表現 (7.15) – (7.17) が成立したとする．$\ell_n(\theta|\,\boldsymbol{x}) = \sum_{j=1}^n \log \mathsf{p}(x_j|\theta)$ $\big(\boldsymbol{x} = (x_1, x_2, \ldots, x_n)\big)$ とおく．すると

$$\overset{\bullet}{\ell}_n(\theta|\,\boldsymbol{x}) = \frac{\partial}{\partial\theta}\ell_n(\theta|\,\boldsymbol{x}) = \overset{\bullet}{A}(\theta)\sum_{j=1}^{n} T(x_j) - n\,\overset{\bullet}{\kappa}(\theta),$$

$$\overset{\bullet}{A}(\theta) = \frac{\partial A}{\partial\theta}(\theta), \quad \overset{\bullet}{\kappa}(\theta) = \frac{\partial B}{\partial\theta}(\theta)$$

[2]Rao 有効の意味であろう．

となる．よって

$$\mathsf{E}_\theta\big[\overset{\bullet}{\ell}_n\,(\theta|\,\boldsymbol{X})\big] = 0 \Leftrightarrow \overset{\bullet}{A}\,(\theta)\mathsf{E}_\theta\big[T(X_1)\big] = \overset{\bullet}{\kappa}\,(\theta)$$

と (7.17) とあわせると

$$\theta\,\overset{\bullet}{A}\,(\theta) = \overset{\bullet}{\kappa}\,(\theta) \tag{7.18}$$

となる．さらに

$$0 = \int_{\mathbb{X}^n} \{T(x) - \theta\}\mathsf{p}(x|\,\theta)\,\mathrm{d}x = \int_{\mathbb{X}^n} \{T(x) - \theta\}\exp\{\log\mathsf{p}(x|\,\theta)\}\,\mathrm{d}x$$

を (7.16) と 7.18) に注意して, θ に関して微分すると

$$\begin{aligned}
0 &= -\int_{\mathbb{X}} \mathsf{p}_1(x|\,\theta)\,\mathrm{d}x + \int_{\mathbb{X}} \{T(x) - \theta\}\big\{\overset{\bullet}{A}\,(\theta)T(x) - \underbrace{\overset{\bullet}{\kappa}\,(\theta)}_{=\theta\overset{\bullet}{A}(\theta)}\big\}\mathsf{p}(x|\,\theta)\,\mathrm{d}x \\
&= -\int_{\mathbb{X}} \mathsf{p}(x|\,\theta)\,\mathrm{d}x + \int_{\mathbb{X}} \{T(x) - \theta\}\big\{\overset{\bullet}{A}\,(\theta)T(x) - \theta\,\overset{\bullet}{A}\,(\theta)\big\}\mathsf{p}(x|\,\theta)\,\mathrm{d}x
\end{aligned}$$

となることがわかる．よって, $\int_{\mathbb{X}} \mathsf{p}(x|\,\theta)\,\mathrm{d}x = 1$ に注意して, 上式を整理すると

$$1 = \overset{\bullet}{A}\,(\theta)\mathsf{E}_\theta\big[\{T(X_1) - \theta\}^2\big]$$

を得る．さらに $\mathsf{E}_\theta[T(X_1)] = \theta$ に注意すれば

$$\mathsf{Var}_\theta\big[T(X_1)\big] = \frac{1}{\overset{\bullet}{A}\,(\theta)} \tag{7.19}$$

となる．したがって

$$\begin{aligned}
\mathcal{F}_{\boldsymbol{X}}(\theta) &= \mathsf{Var}_\theta\big[\overset{\bullet}{\ell}_n\,(\theta|\,\boldsymbol{X})\big] && (7.20)\\
&= \mathsf{Var}_\theta\Big[\sum_{j=1}^n \{\overset{\bullet}{A}\,(\theta)T(X_j) - \overset{\bullet}{\kappa}\,(\theta)\}\Big] \\
&= \mathsf{Var}_\theta\Big[\sum_{j=1}^n \overset{\bullet}{A}\,(\theta)\big\{T(X_j)\big\}\Big] && (\because \text{分散は平行移動に関して不変}) \\
&= \sum_{j=1}^n \{\overset{\bullet}{A}\,(\theta)\}^2\mathsf{Var}\big[T(X_j)\big] && (\because X_1,\, X_2,\, \ldots,\, X_n \text{ は独立}) \\
&= n\{\overset{\bullet}{A}\,(\theta)\}^2\mathsf{Var}\big[T(X_1)\big] && (7.21)\\
&= n\,\overset{\bullet}{A}\,(\theta) && (\because (7.19))
\end{aligned}$$

を得る．また, $\widehat{\theta}_n(\boldsymbol{X}) = \dfrac{1}{n}\sum_{j=1}^n T(X_j)$ なので

$$\mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big] = \theta$$

となり $\widehat{\theta}_n$ は θ の不偏推定量である．再度, (7.19) に注意すると

$$\mathsf{Var}_\theta\big[\widehat{\theta}_n\big] = \mathsf{Var}_\theta\left[\frac{1}{n}\sum_{j=1}^n T(X_j)\right] = \frac{1}{n^2}\sum_{j=1}^n \mathsf{Var}_\theta\big[T(X_j)\big] = \frac{1}{n\,\overset{\bullet}{A}(\theta)} = \frac{1}{\mathcal{F}_{\boldsymbol{X}}(\theta)}$$

となり $\widehat{\theta}_n$ は R 有効推定量となる．

次に逆を示す．P_θ の同時 p.d.f. $\prod_{j=1}^n \mathsf{p}(x_j|\,\theta)$ に対して, その対数尤度関数を $\ell_n(\theta|\,\boldsymbol{x}) = \log\Big(\prod_{j=1}^n \mathsf{p}(x_j|\,\theta)\Big)$ と書くことにする．すると

$$\int_{\mathbb{X}^n} \exp\{\ell_n(\theta|\,\boldsymbol{X})\}\,d\boldsymbol{x} = 1$$

である．この式を θ に関して微分すると

$$0 = \int_{\mathbb{X}^n} \overset{\bullet}{\ell}_n(\theta|\,\boldsymbol{x})\exp\{\ell_n(\theta|\,\boldsymbol{x})\}\,\mathrm{d}\boldsymbol{x} = \mathsf{E}_\theta\big[\overset{\bullet}{\ell}_n(\boldsymbol{x}|\,\boldsymbol{X})\big] \tag{7.22}$$

となる．同様に $\mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\big] = \theta$ より

$$\int_{\mathbb{X}^n} \widehat{\theta}_n(\boldsymbol{x})\exp\{\ell_n(\theta|\,\boldsymbol{x})\}\,\mathrm{d}\boldsymbol{x} = \theta$$

となる．この式を θ に関して微分すると

$$1 = \int_{\mathbb{X}^n} \widehat{\theta}_n(\boldsymbol{x})\,\overset{\bullet}{\ell}_n(\theta|\,\boldsymbol{x})\exp\{\ell_n(\theta|\,\boldsymbol{x})\}\,d\boldsymbol{x} = \mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\,\overset{\bullet}{\ell}_n(\theta|\,\boldsymbol{X})\big] \tag{7.23}$$

を得る．(7.22) と (7.23) を合わせると

$$\begin{aligned}\mathsf{E}_\theta\big[(\widehat{\theta}_n(\boldsymbol{X}) - \theta)\,\overset{\bullet}{\ell}_n(\theta|\,\boldsymbol{X})\big] &= \underbrace{\mathsf{E}_\theta\big[\widehat{\theta}_n(\boldsymbol{X})\,\overset{\bullet}{\ell}_n(\theta|\,\boldsymbol{X})\big]}_{=1} - \theta\,\underbrace{\mathsf{E}_\theta\big[\overset{\bullet}{\ell}_n(\theta|\,\boldsymbol{X})\big]}_{=0} \\ &= 1 \end{aligned} \tag{7.24}$$

を得る．いま

$$Y = \frac{\overset{\bullet}{\ell}_n(\theta|\,\boldsymbol{X})}{\mathcal{F}_{\boldsymbol{X}}(\theta)}$$

とおく．(7.20) に注意すると

$$\mathsf{E}_\theta\big[Y\,\overset{\bullet}{\ell}_n(\theta|\,\boldsymbol{X})\big] = \mathsf{E}_\theta\left[\frac{\big\{\overset{\bullet}{\ell}_n(\theta|\,\boldsymbol{X})\big\}^2}{\mathcal{F}_{\boldsymbol{X}}(\theta)}\right] = \frac{1}{\mathcal{F}_{\boldsymbol{X}}(\theta)}\mathsf{Var}[\overset{\bullet}{\ell}_n(\theta|\,\boldsymbol{X})] = 1$$

となる．これと (7.24) を合わせると

$$
\begin{aligned}
&\mathsf{E}_\theta\big[(\widehat{\theta}_n(\boldsymbol{X})-\theta-Y)Y\big]\\
&\quad=\frac{1}{\mathcal{F}_{\boldsymbol{X}}(\theta)}\underbrace{\mathsf{E}_\theta\big[(\widehat{\theta}_n(\boldsymbol{X})-\theta)\,\dot{\ell}_n\,(\theta|\,\boldsymbol{X})\big]}_{=1}-\frac{1}{\mathcal{F}_{\boldsymbol{X}}(\theta)}\underbrace{\mathsf{E}_\theta\big[Y\,\dot{\ell}_n\,(\theta|\,\boldsymbol{X})\big]}_{=1}\\
&\quad=0
\end{aligned}
$$

を得る．上の式を用いると

$$
\begin{aligned}
\mathsf{E}_\theta\big[\big(\widehat{\theta}_n(\boldsymbol{X})-\theta\big)^2\big]&=\mathsf{E}_\theta\big[\big(\widehat{\theta}_n(\boldsymbol{X})-\theta-Y+Y\big)^2\big]\\
&=\mathsf{E}_\theta\big[\big(\widehat{\theta}_n(\boldsymbol{X})-\theta-Y\big)^2\big]+\mathsf{E}_\theta\big[Y^2\big]\\
&\geq\mathsf{E}_\theta\big[Y^2\big]
\end{aligned}
$$

となる．よって等号成立は

$$
\begin{aligned}
\widehat{\theta}_n(\boldsymbol{X})-\theta-Y=0\ &\Leftrightarrow\ \widehat{\theta}_n(\boldsymbol{X})-\theta=\frac{\dot{\ell}\,(\theta|\,\boldsymbol{X})}{\mathcal{F}_{\boldsymbol{X}}(\theta)}\\
&\Leftrightarrow\ \mathcal{F}_{\boldsymbol{X}}(\theta)\big(\widehat{\theta}_n(\boldsymbol{X})-\theta\big)=\dot{\ell}_n\,(\theta|\,\boldsymbol{X})
\end{aligned}
$$

となる．このことより，$\overline{\theta}$, $\underline{\theta}\in\mathbb{X}$ とすると

$$
\begin{aligned}
\ell_n(\overline{\theta}|\,\boldsymbol{X})-\ell_n(\underline{\theta}|\,\boldsymbol{X})&=\int_{\underline{\theta}}^{\overline{\theta}}\dot{\ell}_n\,(\boldsymbol{X}|\,\theta)\,\mathrm{d}\theta=\int_{\underline{\theta}}^{\overline{\theta}}\mathcal{F}_{\boldsymbol{X}}(\theta)\big(\widehat{\theta}_n(\boldsymbol{X})-\theta\big)\,\mathrm{d}\theta\\
&=\widetilde{A}(\overline{\theta},\,\underline{\theta})\widehat{\theta}_n(\boldsymbol{X})-\widetilde{B}(\overline{\theta},\,\underline{\theta})
\end{aligned}
$$

と書ける．ただし

$$
\widetilde{A}(\overline{\theta},\,\underline{\theta}):=\int_{\underline{\theta}}^{\overline{\theta}}\mathcal{F}_{\boldsymbol{X}}(\theta)\,\mathrm{d}\theta;\qquad\widetilde{B}(\overline{\theta},\,\underline{\theta}):=\int_{\underline{\theta}}^{\overline{\theta}}\theta\mathcal{F}_{\boldsymbol{X}}(\theta)\,\mathrm{d}\theta
$$

である．ここで $n=1$, $\overline{\theta}=\theta$, $\underline{\theta}=1$, $X_1=x$ とおくと

$$
\ell_1(\theta|\,x)=T(x)A(\theta)-B(\theta)+g(x)
$$

と書ける．ただし

$$
T(x)=\widehat{\theta}_1(x),\quad g(x)=\ell_1(1|\,x),\quad A(\theta)=\widetilde{A}(\theta,\,1),\quad B(\theta)=\widetilde{B}(\theta,\,1)
$$

と書ける．□

注意 7.23. 定理 7.22 から，R 有効な推定量が存在する統計的モデルは指数型分布族となることがわかる．□

7.4 章末注釈と参考文献

定理 7.11 の証明は $[7, \mathrm{pp}.241-242]$ を借用した．定理 7.14 の証明は $[42, \mathrm{pp}.35-41]$ を借用した．定理 7.22 は $[30, \mathrm{pp}.39-40]$ を借用した．

7.5 演習問題

演習問題 7.1. 離散型確率変数 X は, p.m.f.

$$\mathsf{p}(x|\,\theta) = \begin{cases} \left(\dfrac{\theta}{2}\right)^{|x|}(1-\theta)^{1-|x|} & (x=-1,\,0,\,1) \\ 0 & (\text{その他の場合}) \end{cases}$$

からの標本の大きさが 1 のランダム標本とする．ただし, $0<\theta<1$ で, X は確率空間 $(\Omega,\,\mathcal{A},\,\mathsf{Pr})$ 上で定義された確率変数とする．ふたつの統計量

$$S = S(X) = |X|, \qquad T = T(X) = \begin{cases} 2 & (X=1) \\ 0 & (\text{その他の場合}) \end{cases}$$

を考える．
(1) S の確率分布を求めよ[3]．
(2) T の確率分布を求めよ．
(3) S が与えられたときの X の条件付確率分布を求め[4], S は θ の十分統計量かどうかを調べよ．
(4) S は θ の不偏推定量かどうかを調べよ．
(5) T は θ の不偏推定量かどうかを調べよ．
(6) S と T の平均 2 乗誤差 $\mathsf{MSE}_S(\theta) = \mathsf{E}[(S-\theta)^2]$ と $\mathsf{MSE}_T(\theta) = \mathsf{E}[(T-\theta)^2]$ を求めよ．S と T の平均 2 乗誤差 $\mathsf{MSE}_S(\theta)$ と $\mathsf{MSE}_T(\theta)$ の大小の比較をせよ．(横軸を θ とし, 縦軸を MSE の値として, S と T の平均 2 乗誤差 $\mathsf{MSE}_S(\theta)$ と $\mathsf{MSE}_T(\theta)$ のグラフを描き比較すること.)

演習問題 7.2. $\theta>0$ とする．連続型確率変数 X は p.d.f.

$$\mathsf{p}(x|\,\theta) = \begin{cases} \theta x^{\theta-1} & (0<x<1) \\ 0 & (\text{その他}) \end{cases}$$

からの大きさ 1 のランダム標本とする．ただし, X は確率空間 $(\Omega,\,\mathcal{A},\,\mathsf{Pr})$ 上で定義された確率変数とする．
(1) 連続型確率変数 X の期待値 $\mathsf{E}[X]$ を求めよ．

[3]確率関数または確率分布表を求めること．
[4]条件付き p.m.f. を求めるか, X が与えられた値ごとの確率分布表を求めればよい．

(2) 連続型確率変数 X の分布関数 $\mathsf{F}(x) = \mathsf{Pr}(X \leq x)\,(x \in \mathbb{R})$ を求めよ.
(3) $X = x\,(0 < x < 1)$ を観測したのときの尤度関数 $\mathsf{lik}(\theta|\,x)$ を述べよ (答えのみでよい).
(4) θ の最尤推定値を求めよ.

演習問題 7.3. $m,\, n \geq 2$ を整数とし

$$\begin{aligned} X_1,\, X_2,\, \ldots,\, X_m &\overset{\text{i.i.d.}}{\sim} N(\mu_1,\, \sigma^2) \\ Y_1,\, Y_2,\, \ldots,\, Y_n &\overset{\text{i.i.d.}}{\sim} N(\mu_2,\, \sigma^2) \end{aligned}$$

とする. さらに

$$\begin{aligned} \overline{X}_m &= \frac{1}{m}\sum_{i=1}^{m} X_i \\ \overline{Y}_n &= \frac{1}{n}\sum_{i=1}^{n} Y_i \\ \widehat{\sigma}^2 &= \frac{1}{m+n-2}\left\{\sum_{i=1}^{m}(\overline{X}_m - X_i)^2 + \sum_{i=1}^{n}(\overline{Y}_n - Y_i)^2\right\} \end{aligned}$$

とおく. これらの確率変数は同じ確率空間 $(\Omega,\, \mathcal{A},\, \mathsf{Pr})$ 上で定義されたものとする. このとき, 以下の問いに答えよ.
(1)

$$\overline{X}_m - \overline{Y}_n \sim \mathsf{N}\left(\mu_1 - \mu_2,\, \frac{m+n}{mn}\sigma^2\right)$$

となることを説明 (証明) せよ.
(2)

$$\mathsf{E}[\widehat{\sigma}^2] = \sigma^2$$

となることを説明（証明）せよ.

演習問題 7.4. $n \geq 2$ を整数とし, $X_1,\, X_2,\, \ldots,\, X_n \overset{\text{i.i.d.}}{\sim} \mathsf{Po}(\lambda)$ とする[5]. ただし, $\lambda > 0$ で, これらの確率変数は確率空間 $(\Omega,\, \mathcal{A},\, \mathsf{Pr})$ 上で定義されているとする.
(1) $\sum_{i=1}^{n} X_i$ の分布は $\mathsf{Po}(n\lambda)$ となることを積率母関数を計算することで示せ.

[5]

$$\mathsf{Pr}(X_1 = x) = \frac{\lambda^x}{x!}e^{-\lambda} \qquad (x = 0,\, 1,\, \ldots)$$

である. ただし, これらの確率変数は確率空間 $(\Omega,\, \mathcal{A},\, \mathsf{Pr})$ 上で定義されたものとする. その積率母関数は

$$\mathsf{M}_{X_1}(t) = \mathsf{E}[e^{tX_1}] = e^{(e^t-1)\lambda}$$

である.

(2) λ の任意の不偏推定量の分散についてのその下限（Cramér-Rao の下限）を求めよ.
(3) λ の最尤推定量を求め, その分散が Cramér-Rao の下限に到達することを確認せよ.

演習問題 7.5. $X_1,\, X_2,\, \ldots,\, X_n$ を母集団分布 (平均は θ, 分散は 1^2 の正規分布)

$$\mathsf{p}(x|\,\theta) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x-\theta)^2}{2}\right)$$

からの大きさ n のランダム標本とする. ただし, 母数 $\theta\,(-\infty < \theta < \infty)$ は未知とする. このとき, 以下の問いに答えよ.
(1) $X_1 = x_1,\, X_2 = x_2,\, \ldots,\, X_n = x_n$ を観測したときの θ の尤度関数 $\mathsf{lik}_n(\theta|\, x_1,\, x_2,\, \ldots,\, x_n)$ と対数尤度関数 $\mathsf{lik}_n(\theta|\, x_1,\, x_2,\, \ldots,\, x_n)$ を書け.
(2) θ の最尤推定値 $\widehat{\theta}_n(x_1,\, x_2,\, \ldots,\, x_n)$ を求めよ.
(3) θ の最尤推定量 $\widehat{\theta}_n(X_1,\, X_2,\, \ldots,\, X_n)$ の平均 $\mathsf{E}[\widehat{\theta}_n(X_1,\, X_2,\, \ldots,\, X_n)]$ を求めよ.
(4) θ の最尤推定量 $\widehat{\theta}_n(X_1,\, X_2,\, \ldots,\, X_n)$ の分散 $\mathsf{Var}[\widehat{\theta}_n(X_1,\, X_2,\, \ldots,\, X_n)]$ を求めよ.
(5) 任意の正数 ϵ に対して,

$$\lim_{n\to\infty} \mathsf{Pr}\big(|\widehat{\theta}_n(X_1,\, X_2,\, \ldots,\, X_n) - \theta| \geq \epsilon\big) = 0$$

を示せ.

演習問題 7.6. 確率空間 $(\Omega,\, \mathcal{A},\, P)$ 上の確率変数 $X_1,\, X_2,\, X_3$ は独立同一に平均が μ, 分散が σ^2 の正規分布 $\mathsf{N}(\mu,\, \sigma^2)$ に従うとする. ただし, $\mu \in \mathbb{R},\, 0 < \sigma < \infty$ である. このとき, 以下の問に答えよ.
(1) X_1 の積率母関数 $M(t) = \mathsf{E}[e^{tX_1}]$ が

$$M(t) = \exp\left(\mu t + \frac{\sigma^2}{2} t^2\right)$$

で与えられることを示せ. ただし, 任意の $\mu \in \mathbb{R}$ と $\sigma > 0$ に対して

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x = 1$$

を証明なしで用いてよい.
(2) X_1 の期待値 $\mathsf{E}[X_1]$ と 2 次の積率 $\mathbb{E}[X_1^2]$ を (1) で求めた積率母関数を利用して求めよ.
(3) $a,\, b,\, c,\, d \in \mathbb{R}$ に対して, σ^2 の推定量の族

$$\begin{aligned} T := g(X_1,\, X_2,\, X_3) = a\big\{X_1^2 + X_2^2 + X_3^2\big\} + b\big\{X_1X_2 + X_1X_3 + X_2X_3\big\} \\ + c\big\{X_1 + X_2 + X_3\big\} + d \end{aligned}$$

を考える.

$$\mathsf{E}[T] = \sigma^2$$

をみたす a, b, c, d を求めよ.
(4) (3) で求めた a, b, c, d に対する T は

$$T = \frac{1}{2}\Big((X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + (X_3 - \overline{X}^2)\Big); \quad \overline{X} = \frac{1}{3}\big(X_1 + X_2 + X_3\big)$$

の形に変形できることを確認せよ.

演習問題 7.7. T_1 と T_2 を母数 θ の最小分散不偏推定量とする. このとき, 以下の問伊に答えよ.
(1) $\mathsf{E}\big[\frac{1}{2}(T_1 + T_2)\big]$ を計算せよ.
(2) $\mathsf{Var}\big[\frac{1}{2}(T_1 + T_2)\big]$ を $\mathsf{Var}[T_1]$, $\mathsf{Cov}[T_1, T_2]$ で表現せよ.
(3) $\mathsf{Cov}[T_1, T_2] = \mathsf{Var}[T_1]$ を証明せよ.
(4) $\mathsf{Var}[T_1 - T_2]$ の値を求めよ.
(5) $\mathsf{Pr}(T_1 = T_2) = 1$ を示せ.

ヒント (3) T_1 は最小分散不偏推定量であることと Cauchy-Scharz の不等式を $\mathsf{Cov}[T_1, T_2]$ に適用するとよい.
(5) 非負値確率変数 W に対して

$$\mathsf{E}[W] = 0 \Rightarrow \mathsf{Pr}(W = 0) = 1$$

を証明なしで用いてよい.

演習問題 7.8.

演習問題 7.9.