立教大学 講義用スライド 2024/12/07

本日の講義内容

- 十分統計量
- 7.2 最尤法
- 7.3 不偏推定と情報不等式

6.4 十分統計量

$\mathbb{X} \subset \mathbb{R}$：空でない集合．標本空間

$(\Omega, \mathcal{A}, P_r)$：基本の確率空間

$n \in \mathbb{N}$ とし，$X_1, X_2, \ldots, X_n$ は確率変数 ろ'ı

$^\forall r \in \mathbb{R}$ に対し

$$\{\omega \in \Omega ; X_j(\omega) > r\} \in \mathcal{A} \quad (j = 1, \ldots, n)$$

$\Leftrightarrow$ $^\forall B \in \mathcal{B}(\mathbb{R})$ に対し

$$\{\omega \in \Omega ; X_j(\omega) \in B\} \in \mathcal{A}$$

$\underline{X} = (X_1, X_2, \ldots, X_n)^T$ は独立同一分布に従う確率変数列とする

$B \in \mathcal{B}(\mathbb{X}^n)$ に対し

$$P^{\underline{X}}(B) := \Pr(\underline{X}^{-1}(B))$$

$$= \Pr(\{\omega \in \Omega ; \underline{X}(\omega) = (X_1(\omega), \ldots, X_n(\omega))^T \in B\})$$

さらに、$B \in \mathcal{B}(\mathbb{X})$ に対し

$$P^{X_j}(B) := \Pr(X_j^{-1}(B))$$

$$= \Pr(\{\omega \in \Omega ; X_j(\omega) \in B\})$$

すると，$B_1, \ldots, B_n \in \mathcal{B}(\mathcal{X})$ に対し

$$P^{\underline{X}}(B_1 \times \cdots \times B_n) = \prod_{j=1}^{n} P^{X_1}(B_j).$$

代表で X_1 とかいて，

確率ベクトル $\underline{X}$ によって誘導された確率空間を

$(\mathcal{X}^n, \mathcal{B}(\mathcal{X}^n), P^{\underline{X}})$ とかく．

$(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ 上の統計的モデルを

$$\mathcal{P} := \{ P_\theta \,;\, \theta \in \Theta \}$$

とかく．ただし，$\Theta \subset \mathbb{R}^k$（ある $k \in \mathbb{N}$）は母数空間

仮定

真の分布

$$P^{X_1} \in \mathcal{P} \iff {}^{\exists}\theta^* \in \Theta \ \text{ s.t. } P^{X_1} = P_{\theta^*}$$

$T(\boldsymbol{X})$: 統計量

$$T : (\mathcal{X}^n, \mathcal{B}(\mathcal{X}^n)) \longrightarrow (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$$

なる可測関数で θ に依存しない.

$B_1, B_2, \ldots, B_n \in \mathcal{B}(\mathcal{X})$ に対し

・$P^{\otimes n}(B_1 \times \cdots \times B_n) = \prod_{j=1}^{n} P(B_j)$ と定める.

定義 6.12 $\underline{X} \sim P_\theta^{\otimes n}$ とする. $^\forall B \in \mathcal{B}(\mathbb{R}^n)$ に対して. 統計量 $\underline{T}(\underline{X})$ を与えたときの $\underline{X}$ の条件付き確率

$$P_\theta^{\otimes n}(B \mid \underline{T})$$

が θ に依存しないとき, $\underline{T}$ は統計的モデル $\mathcal{P}=\{P_\theta : \theta \in \Theta\}$ に対する**十分統計量**という

記号　X の同時 p.d.f. / p.m.f. を $p^{X}(x|\theta)$ とおく。

定理 6.13（Fisher-Neyman の因子分解定理）

$\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ を統計的モデルとし、

$X_1, \dots, X_n \overset{\text{i.i.d.}}{\sim} P_\theta$ とする。このとき、以下の2つの条件は同値

(1) 統計量 $T : \mathbb{R}^n \to \mathbb{R}^\ell$ は $\mathcal{P}$ に対する十分統計量

(2) $^\exists g : \mathbb{R}^\ell \to [0, \infty)$ と $h : \mathcal{X}^n \to [0, \infty)$ が存在

して

$$p^{X}(\underline{x}|\theta) = g_\theta(T(\underline{x}))\, h(\underline{x}) \qquad (\underline{x} \in \mathcal{X}^n)$$

と表現できる.

cheat proof

(1) ⇒ (2) の確認 (離散型の場合)

$p^{T}(\underline{t}|\theta)$ を $\underline{T}$ の周辺 p.m.f.

$p^{X|T}(\underline{x}|\underline{T}=\underline{t})$: $\underline{T}=\underline{t}$ を与えたときの $\underline{X}$ の条件付き

き p.m.f.

すると, 条件付き p.m.f. の定義と $\underline{t}=\underline{T}(\underline{x})$ なので.

$$p^{\underline{X}}(\underline{x}\mid\theta) = \underbrace{p^{\underline{T}}(\underline{t}\mid\theta)}_{=:\, g_\theta(\underline{T}(\underline{x}))} \; \underbrace{p^{\underline{X}\mid\underline{T}}(\underline{x}\mid\underline{T}=\underline{t})}_{\theta\text{によらない} \;=:\, h(\underline{x})}$$

と書ける.

(2) ⇒ (1) の確認

$$p^{\underline{T}}(\underline{t}\mid\theta) = \sum_{\underline{y}\in\mathcal{X}^n,\, \underline{T}(\underline{y})=\underline{t}} p^{\underline{X}}(\underline{y}\mid\theta)$$

$$= \sum_{\underline{y}\in\mathcal{X}^n,\, \underline{T}(\underline{y})=\underline{t}} g_\theta(\underline{T}(\underline{y}))\, h(\underline{y})$$

$$= g_\theta(\underline{t}) \sum_{\underline{y} \in \mathcal{X}^n,\ \underline{T}(\underline{y}) = \underline{t}} h(\underline{y}) \,. \qquad (A)$$

よって，$\underline{t} = \underline{T}(\underline{x})$ のとき

$$p^{\underline{X}|\underline{T}}(\underline{x}|\underline{t}) = \frac{p^{\underline{X}}(\underline{x}|\theta)}{p^{\underline{T}}(\underline{t}|\theta)}$$

(1)

$$= \frac{\cancel{g_\theta(\underline{t})}\, h(\underline{x})}{\cancel{g_\theta(\underline{t})} \sum_{\underline{y} \in \mathcal{X}^n,\ \underline{T}(\underline{y}) = \underline{t}} h(\underline{y})} \quad \leftarrow \theta \text{に依存しない!}$$

(A)

$\square$

例 6.15 $n \in \mathbb{N}$ とし, $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathrm{Ber}(\theta)$ とする. ただし, $0 \leq \theta \leq 1$. $\underline{X} = (X_1, \ldots, X_n)^T$ の同時 p.m.f. は, $\underline{x} = (x_1, \ldots, x_n)$ に対し

$$
\begin{aligned}
p^{\underline{X}}(\underline{x} \mid \theta) &= \Pr(X_1 = x_1, \ldots, X_n = x_n) \\
&= \prod_{j=1}^{n} \Pr(X_j = x_j) \\
&= \prod_{j=1}^{n} \theta^{x_j} (1-\theta)^{1-x_j} \, \mathbb{1}_{\{0,1\}^n}(\underline{x}) \\
&= \underbrace{\theta^{\sum_{j=1}^{n} x_j} (1-\theta)^{n-\sum_{j=1}^{n} x_j}}_{} \, \underbrace{\mathbb{1}_{\{0,1\}^n}(\underline{x})}_{}.
\end{aligned}
$$

$= g_\theta(T(\underline{x}))$ $\qquad$ $= h(\underline{x})$

ただし

$$\mathbb{1}_{\{0,1\}^n}(\underline{x}) = \begin{cases} 1 & (\underline{x} \in \{0,1\}^n) \\ 0 & (\text{otherwise}) \end{cases}.$$

したがって

$$T(\underline{x}) = \sum_{j=1}^{n} x_j, \quad g_\theta(T(\underline{x})) = \theta^{T(\underline{x})}(1-\theta)^{n-T(\underline{x})}$$

$$h(\underline{x}) = \mathbb{1}_{\{0,1\}^n}(\underline{x})$$

とおけばよ..

問2

$T(\underline{X}) = X_1 + X_2 + \cdots + X_n$ は $\mathcal{P} = \{ p(\underline{x} | \theta);$

$0 \le \theta \le 1 \}$ に対する 十分統計量

例 6.16 $n \in \mathbb{N}$ と $\theta \in \mathbb{R}$ とし、

$$X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} N(\theta, 1)$$

とする。$\underline{X} = (X_1, \ldots, X_n)^T$ の同時 p.d.f. は

$$p^{\underline{X}}(\underline{x} \mid \theta) = \prod_{j=1}^{n} \underbrace{\frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x_j-\theta)^2\right\}}_{N(\theta,1)\text{ の p.d.f.}}$$

$$= \frac{1}{(\sqrt{2\pi})^n} \exp\left\{-\frac{1}{2}\sum_{j=1}^{n}(x_j-\theta)^2\right\}$$

$T(\underline{x}) = \sum_{j=1}^{n} x_j / n$ とすると

$$\sum_{j=1}^{n}(x_j-\theta)^2 = \sum_{j=1}^{n}\{x_j^2 - 2\theta x_j + \theta^2\}$$

$$= \sum_{j=1}^{n}x_j^2 - 2\theta\sum_{j=1}^{n}x_j + n\theta^2$$

$$= \sum_{j=1}^{n}x_j^2 - 2n\theta T(\underline{x}) + n\theta^2$$

よって

$$p^{\underline{X}}(\underline{x}|\theta) = \underline{\exp\left\{nT(\underline{x})\theta - \frac{n\theta^2}{2}\right\}}\ \underline{\left(\frac{1}{2\pi}\right)^{\frac{n}{2}}\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}x_j^2\right\}}$$

$=: g_\theta(T(\underline{x}))$ $\qquad = h(\underline{x})$

よって

$T(\underline{X}) = \frac{1}{n}(X_1 + \cdots + X_n)$ は $\mathcal{P} = \{N(\theta, 1) : \theta \in \mathbb{R}\}$

の十分統計量.

注意 $S(\underline{X}) = X_1 + \cdots + X_n$ も $\mathcal{P}$ の十分統計量

表現も一意でない. また

$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ を $X_1, \ldots, X_n$ を順番に

並べたものとすると, $(X_{(1)}, \ldots, X_{(n)})$ は $\mathcal{P}$ の十分統計量

例 6.17　$n \in \mathbb{N}\ (n \geq 2)$ とし，$\mu \in \mathbb{R}$, $0 < \sigma < \infty$

とする．$X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$

とする．すなわち

$$\mathcal{P} = \{ N(\mu, \sigma^2) : \mu \in \mathbb{R},\ \sigma > 0 \}.$$

$$\theta = (\mu, \sigma^2)$$

$\underline{X} = (X_1, \ldots, X_n)^T$ の同時 p.d.f. は

$$p^{\underline{X}}(\underline{x} \mid \theta) = \prod_{j=1}^{n} \left(\frac{1}{2\pi\sigma^2} \right)^{1/2} \exp\left\{ - \frac{(x_j - \mu)^2}{2\sigma^2} \right\}$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=1}^{n}(x_j-\mu)^2\right\}$$

ただし $\underline{x} = (x_1, \ldots, x_n)$.

ここで

$$\bar{x}_n = \frac{1}{n}\sum_{j=1}^{n} x_j, \quad S_n^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j-\bar{x}_n)^2$$

とおく. すなわち

$$T(\underline{x}) = (x_n, S_n^2).$$

とする.

$$\sum_{j=1}^{n}(x_j-\theta)^2=\sum_{j=1}^{n}\left\{x_j-\bar{x}_n+\bar{x}_n-\theta\right\}^2$$

$$=\sum_{j=1}^{n}\left\{(x_j-\bar{x}_n)^2+(\bar{x}_n-\theta)^2+2(\bar{x}_n-\theta)(x_j-\bar{x}_n)\right\}$$

$$=\sum_{j=1}^{n}(x_j-\bar{x}_n)^2+n(\bar{x}_n-\theta)^2+2(\bar{x}_n-\theta)\underbrace{\sum_{j=1}^{n}(x_j-\bar{x}_n)}_{=0}$$

$$=\sum_{j=1}^{n}(x_j-\bar{x}_n)^2+n(\bar{x}_n-\theta)^2$$

$$\because)\ \sum_{j=1}^{n}(x_j-\bar{x}_n)=\sum_{j=1}^{n}x_j-n\bar{x}_n=n\bar{x}_n-n\bar{x}_n=0.$$

σ^2

$$p^{\underline{X}}(\underline{x}|\underline{\theta}) = \underbrace{\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left\{-\frac{n^2}{2\sigma^2}\left(s_n^2 + (\bar{x}_n - \theta)^2\right)}_{g_\theta(T(\underline{x}))}\right\}$$

$h(\underline{x}) = 1.$

7.2 最尤法

さいゆう

$\mathbb{X} \subset \mathbb{R}$: 空でない部分集合, 標本空間

$d, n \in \mathbb{N}$

$(\mathbb{X}^n, \{ p(x|\underline{\theta}) : \underline{\theta} \in \Theta \})$: 統計的実験

$\Theta \subset \mathbb{R}^d$ で $p(x|\underline{\theta})$ は p.d.f. /p.m.f.

$\underline{\theta}^* \in \Theta$ とし

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

とする.

定義7.4 $X_1=x_1, X_2=x_2, \ldots, X_n=x_n$ を観測したときの $\underline{\theta}$ の尤度関数 $\mathrm{lik}_n(\underline{\theta} \mid \underline{x})$ を

$$\mathrm{lik}_n(\underline{\theta} \mid \underline{x}) = \prod_{j=1}^{n} p(x_j \mid \underline{\theta}) \qquad \underline{x} = (x_1, \ldots, x_n)$$

で定義し、対数尤度(関数) $l_n(\underline{\theta} \mid \underline{x})$ を

$$l_n(\underline{\theta} \mid \underline{x}) = \log \mathrm{lik}_n(\underline{\theta} \mid \underline{x})$$

で定義する。

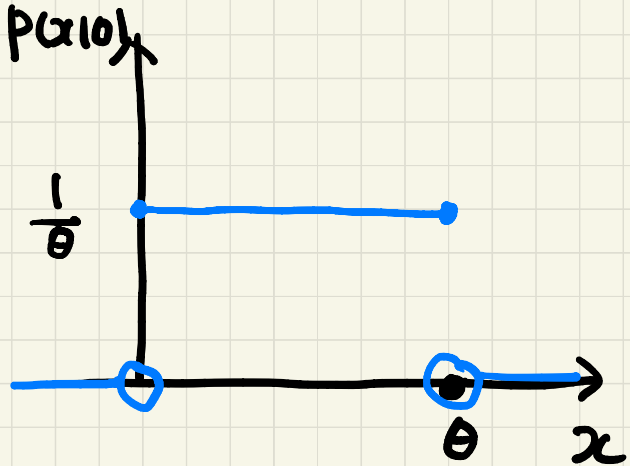
例　$\theta > 0$ として

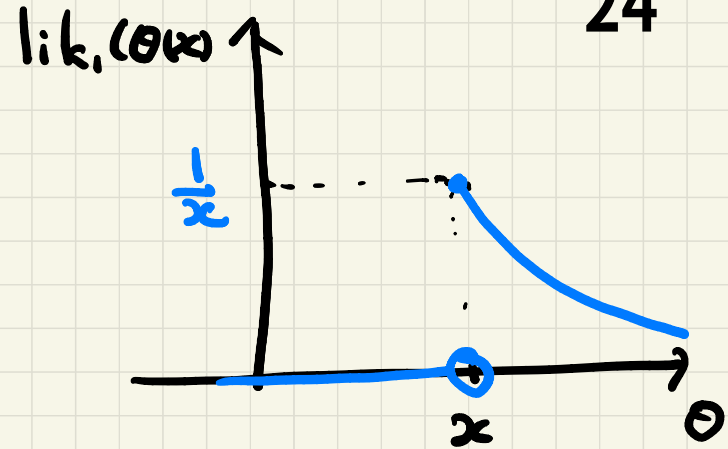
$$p(x \mid \theta) = \frac{1}{\theta} \mathbb{1}_{[0,\theta]}(x) = \begin{cases} \dfrac{1}{\theta} & (0 \leq x \leq \theta) \\ 0 & (\text{otherwise}) \end{cases}$$

とする．すると，$X_1 = x\ (x > 0)$ を観測したとき．

$$\mathrm{lik}_1(\theta \mid x) = \frac{1}{\theta} \mathbb{1}_{[0,\theta)}(x) = \frac{1}{\theta} \mathbb{1}_{[x,\infty)}(\theta)$$

とする．





記法 関数 $g: \mathbb{R} \to \mathbb{R}$ の最大値を取る点の集合を

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

と記す.

たとえば、$g(x) = -(x-1)^2$ のとき

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

$g(x) = \sin x$ のとき

$$\underset{0 \leq x \leq 4\pi}{\arg\max}\ g(x) = \left\{\frac{\pi}{2}, \frac{5}{2}\pi\right\}$$

となる。

$g(x) = x^2$ のとき

$$\underset{x \in \mathbb{R}}{\arg\max}\ g(x) = \phi$$ (存在しない)

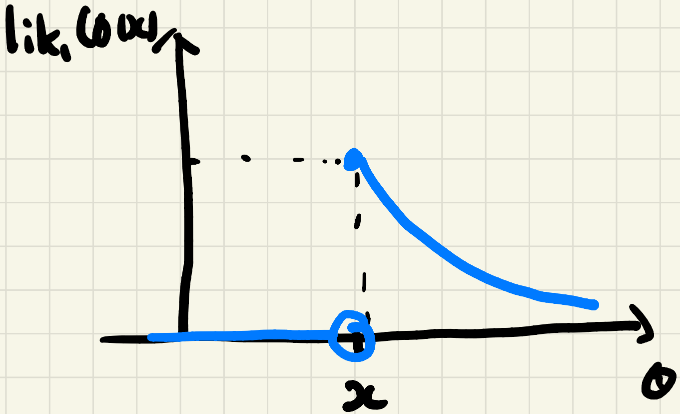
定義7.6 $X_1=x_1, X_2=x_2, \ldots, X_n=x_n$ を観測したとき.

$\boldsymbol{\theta}^*$ の最尤推定値を $\mathrm{lik}_n(\boldsymbol{\theta}\,|\,\boldsymbol{x})$ を最大にする点の値 $\hat{\boldsymbol{\theta}}_n(\boldsymbol{x})$ で定める. すなわち

$$\hat{\boldsymbol{\theta}}_n(\boldsymbol{x}) \in \operatorname*{arg\,max}_{\boldsymbol{\theta}\in\Theta} \mathrm{lik}_n(\boldsymbol{\theta}\,|\,\boldsymbol{x})$$

である. $\boldsymbol{x}=(x_1,\ldots,x_n)$ に $\boldsymbol{X}=(X_1,\ldots,X_n)$ を形式的に代入したもの $\hat{\boldsymbol{\theta}}_n(\boldsymbol{X})$ を $\boldsymbol{\theta}^*$ の最尤推定量という.

例1の続き



$\Rightarrow \hat{\theta}(x) = x.$

離散分布のときはめんどう.

例4.7.7 $X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} \mathrm{Ber}(\theta^*)$ とする

ただし $\Theta = (0, 1) \ni \theta^*$ とする すなわち

$$p(x \mid \theta^*) = (\theta^*)^x (1-\theta^*)^{1-x} \mathbb{I}_{\{0,1\}}(x).$$

$X_j = x_j \in \{0, 1\}$ $(j = 1, \ldots, n)$ を観測したとき.

$$\mathrm{lik}_n(\theta \mid \underline{x}) = \prod_{j=1}^{n} p(x_j \mid \theta)$$

$$= \prod_{j=1}^{n} \theta^{x_j} (1-\theta)^{1-x_j} \mathbb{I}_{\{0,1\}}(x_j)$$

$$= \theta^{t_n}(1-\theta)^{n-t_n} \mathbb{I}_{\{0,1,\dots,n\}}(t_n)\}9$$

となる. ただし, $t_n = \sum_{j=1}^{n} x_j$ である.

よって

$$l_n(\theta \mid \underline{x}) = \{t_n \log \theta + (n-t_n)\log(1-\theta)\} \mathbb{I}_{\{0,\dots,n\}}(t_n)$$

$$(0 < \theta < 1)$$

なので, $0 < t_n < n$ のとき

$$\frac{t_n}{n} \in \underset{0<\theta<1}{\arg\max}\; l_n(\theta \mid \underline{x})$$

一方，$t_n = 0$ のとき

$$l_n(\theta \mid \underline{x}) = n \log(1-\theta)$$

となり

$$\underset{0<\theta<1}{\arg\max}\, l_n(\theta \mid \underline{x}) = \phi$$

$t_n = n$ のとき，$l_n(\theta \mid \underline{x}) = n \log \theta$ となり

$$\underset{0<\theta<1}{\arg\max}\, l_n(\theta \mid \underline{x}) = \phi$$

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

標準偏差で母数化

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

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

$$\mathrm{lik}_n(\underline{\theta} \mid \underline{x}) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{ -\frac{(x_j - \mu)^2}{2\sigma^2} \right\}$$

$$= \left(\frac{1}{\sqrt{2\pi}} \right)^n \frac{1}{\sigma^n} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{n} (x_j - \mu)^2 \right\}$$

$$= \left(\frac{1}{\sqrt{2\pi}}\right)^n \frac{1}{\sigma^n} \exp\left\{-\frac{nS_n^2}{2\sigma^2}\right\} \exp\left\{-\frac{n(\bar{x}_n-\mu)^2}{2\sigma^2}\right\}$$

と書き直せる. ただし

$$S_n^2 = \frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x}_n)^2, \quad \bar{x}_n = \frac{1}{n}\sum_{j=1}^{n} x_j$$

$$\sum_{j=1}^{n}(x_j - \bar{x}_n)^2 > 0 .$$

$$\because) \quad \Pr\left(\sum_{j=1}^{n}(x_j - \bar{x}_n)^2 = 0\right) = 0 .$$

対数尤度は

$$l_n(\mu, \sigma \mid \underline{x}) = -n\log\sigma - \frac{nS_n^2}{2\sigma^2} - \frac{n(\bar{x}_n-\mu)^2}{2\sigma^2}$$

$+$ (定数項).

となる. よって

$$\begin{cases} \dfrac{\partial l_n(\mu, \sigma \mid \underline{x})}{\partial \mu} = \dfrac{\bar{x} - \mu}{\sigma^2} = 0 \\ \dfrac{\partial l_n(\mu, \sigma \mid \underline{x})}{\partial \sigma} = -\dfrac{n}{\sigma} + \dfrac{n s_n^2}{\sigma^3} + \dfrac{n(\bar{x}_n - \mu)^2}{\sigma^3} \end{cases}$$

を解くと

$$\mu = \bar{x}_n, \quad \sigma = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x}_n)^2} = \sqrt{S_n^2} = |S_n| = S_n$$

$\because \; S_n > 0.$

$l_n(\mu, \sigma \mid \underline{x})$ の Hessian を求めると

$$\underline{H} = \begin{pmatrix} \frac{\partial l_n^2}{\partial \mu^2}(\bar{x}_n, S_n \mid \underline{x}) & \frac{\partial l_n^2}{\partial \mu \partial \sigma}(\bar{x}_n, S_n \mid \underline{x}) \\ \frac{\partial l_n^2}{\partial \sigma \partial \mu}(\bar{x}_n, S_n \mid \underline{x}) & \frac{\partial l_n^2}{\partial \sigma \partial \sigma}(\bar{x}_n, S_n \mid \underline{x}) \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{1}{2 S_n^2} & 0 \\ 0 & -\frac{2n}{S_n^2} \end{pmatrix}$$

より、$-\underline{H}$ は正定値となるので

$\Theta \ni (\mu, \sigma) \longmapsto l_n(\mu, \sigma \mid \underline{x})$

は $(\mu, \sigma) = (\bar{x}_n, S_n)$ で最大となる

よって，(μ^*, σ^*) の最尤推定量は

$$\hat{\mu}_n = \frac{1}{n}\sum_{j=1}^{n} X_j, \quad \hat{S}_n = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X}_n)^2}$$

注意　$g: \mathbb{R}\times[0,\infty) \to \mathbb{R}\times[0,\infty)$ は1対1とする。

すると $\hat{\theta}_n$ が θ^* の最尤推定量ならば、$g(\hat{\theta}_n)$ は $g(\theta^*)$ の最尤推定量であることがわかる。すなわち、最尤推定量は関数の変換に関して不変!

よって

$(\hat{\mu}_n, \sigma_n^2)$ は $(\mu^*, (\sigma^*)^2)$ の最尤推定量である。

定理7.10 $\mathcal{X} \subset \mathbb{R}$を空でない部分集合とする

$$X_1, X_2, \dots, X_n \overset{\text{i.i.d.}}{\sim} p(x|\eta) = h(x)\exp\{\eta^* T(x) - u(\eta)\}\mathbb{I}_{\mathcal{X}}(x)$$

$$(\eta \in \mathcal{E}^\circ)$$

とする.

$$\mathcal{E} := \left\{\eta \in \mathbb{R};\ u(\eta) = \int_{\mathcal{X}} h(x)\exp\{\eta T(x)\}dx < \infty\right\}$$

の内部を $\mathcal{E}^\circ$ とした. さらに

$$\ddot{u}(\eta^*) = \frac{d^2 u}{d\eta^2}(\eta^*) > 0 \quad (\eta^* \in \mathcal{E}^\circ)$$

と仮定する. このとき, η^*の最尤推定量 $\hat{\eta}_n$ は

十分大きな n に対して確率1で一意的に存在

$$\hat{\eta}_n \xrightarrow{a.s.} \eta^*$$

$$\sqrt{n}(\hat{\eta}_n - \eta^*) \to N\left(0, \frac{1}{\ddot{u}(\eta^*)}\right)$$

となる。

<u>証明</u> $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ を観測したとき

$$\mathrm{lik}_n(\eta \mid \underline{x}) = \left\{\prod_{j=1}^{n} f(x_j)\right\} \exp\left[\eta \sum_{j=1}^{n} T(x_j) - n\,u(\eta)\right]$$

$\times \mathbb{1}_{\mathcal{X}_n}(\underline{x})$

となる。ただし $\underline{x} = (x_1, \ldots, x_n)^T$.

よって

$$\ell_n(\eta \mid \underline{x}) = \log \mathrm{lik}_n(\eta \mid \underline{x})$$

$$= n\{\overline{T}_n - u(\eta)\} + (\mathrm{constant}).$$

ただし $\overline{T}_n = n^{-1}\sum_{j=1}^{n} T(x_j)$.

したがって

$$\dot{\ell}_n(\eta \mid \underline{x}) = n(\overline{T}_n - \dot{u}(\eta)) = 0$$

とおくと

$$\overline{T}_n = \dot{u}(\eta) = E[T(X_1)].$$

形式的に x_j に X_j $(j=1,\dots,n)$ を代入する.

$$\overline{T}_n = \frac{1}{n}\sum_{j=1}^{n} T(X_j) \xrightarrow{a.s.} E[T(X_1)] = \dot{u}(\eta^*)$$

$$(n \to \infty)$$

$\because$) 大数の強法則より

$\dot{u}(\mathcal{E}^\circ) := \{\dot{u}(\eta);\ \eta \in \mathcal{E}^\circ\}$ とおくと

$$\Pr\left(\lim_{n\to\infty}\overline{T}_n \in \dot{u}(\mathcal{E}^\circ)\right) = 1$$

となる.

さらに, $U(\eta^*)$ を η^* の開近傍とする. $U(\eta^*)$ を十分小さく取れば, n が大きいとき,

$$\dot{u}(\eta) > 0 \quad (\eta \in U(\eta^*))$$

とできる. したがって

$$U(\eta^*) \ni \eta \longmapsto \dot{u}(\eta)$$

は狭義単調増加関数である.

近傍 $U(\eta^*)$ の元 η が存在して

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

となる. $\dot{\mu}(\eta) > 0$ $(\eta \in U(\eta^*))$ なので、近傍 $U(\eta^*)$

で $\dot{\mu}$ の逆写像 $\dot{\mu}^{-1}$ が存在して

$$\hat{\eta}_n := \dot{\mu}^{-1}(\bar{T}_n)$$

と定める. さらに

$$\bar{T}_n \xrightarrow{a.s.} \dot{\mu}(\eta^*) \qquad (n \to \infty)$$

かつ

$$\dot{\mu}^{-1} : \dot{\mu}(U(\eta^*)) \to U(\eta^*)$$

で連続

よって

$$\hat{\eta}_n = \dot{\psi}^{-1}(\overline{T}_n) \xrightarrow{a.s.} \dot{\psi}^{-1}(\dot{\psi}(\eta^*)) = \eta^*$$

がわかる.

一方

$$E[T(X_1)] = \dot{\psi}(\eta^*),$$

$$\mathrm{Var}[T(X_1)] = \ddot{\psi}(\eta^*)$$

であったので, 中心極限定理を用いると

$$\sqrt{n}\left(\overline{T}_n - \dot{\mu}(\eta^*)\right) = \sqrt{n}\left(\overline{T}_n - E[T(X_1)]\right)$$

$$\rightsquigarrow N\left(0, \ddot{\mu}(\eta^*)\right)$$

がわかる.

$\hat{\eta}_n = \dot{\mu}^{-1}(\overline{T}_n)$ に対して，デルタ法を用いると

$$\sqrt{n}\left(\hat{\eta}_n - \eta^*\right) = \sqrt{n}\left\{\dot{\mu}^{-1}(\overline{T}_n) - \dot{\mu}^{-1}(E[T(X_1)])\right\}$$

$$\rightsquigarrow N\left(0, \frac{\ddot{\mu}(\eta^*)}{\left\{\frac{\partial \mu^{-1}}{\partial \eta}(E[T(X_1)])\right\}^2}\right)$$

例1.

$$\frac{\partial \mu^{-1}}{\partial \eta}(E[T(x_1)]) = \ddot{\mu}(\eta^*)$$

よって

$$\sqrt{n}(\hat{\eta}_n - \eta^*) \rightsquigarrow N\left(0, \frac{1}{\ddot{\mu}(\eta^*)}\right). \qquad \square$$

注意

$$\frac{\partial l_n}{\partial \eta}(\eta \mid \underline{X}) = n\{\bar{T}_n - \dot{u}(\eta)\}$$

となるので、

$$E_{\eta^*}\left[\left\{\frac{\partial l_1}{\partial \eta}(\eta^* \mid \underline{X})\right\}\right] = E_{\eta}\left[\{T(X_1) - \dot{u}(\eta)\}^2\right]$$

$$= \mathrm{Var}_{\eta^*}[T(X_1)]$$

$$= \ddot{u}(\eta^*)$$

よって

$$\hat{I}(\eta) = \ddot{\ell}(\eta^*)$$

なので

$$\sqrt{n}\left(\hat{\eta}_n - \eta^*\right) \rightsquigarrow N\left(0, \frac{1}{\hat{I}(\eta^*)}\right)$$

とも書ける. $\square$

7.3 不偏推定と情報不等式

$\mathcal{X} \subset \mathbb{R}$: 空でない部分集合

$(\Omega, \mathcal{A}, P_r)$: 確率空間

$X_1, X_2, \ldots, X_n$: $(\Omega, \mathcal{A}, P_r)$ 上の確率変数列

$\underline{X} = (X_1, X_2, \ldots, X_n)^T$

$P := P_r \circ X_1^{-1}$, $\quad P^{\otimes n} := P_r \circ \underline{X}^{-1}$

$\Leftrightarrow$ ${}^\forall B \in \mathcal{B}(\mathcal{X})$ に対し

$$P(B) = P_r \circ X_1^{-1}(B) = P_r\left(\omega \in \Omega : X_1(\omega) \in B\right)$$

$B_1, \dots, B_n \in \mathcal{B}(\mathcal{X})$ に対し

$$P^{\otimes n}(B_1 \times \cdots \times B_n) = P_r(\omega \in \Omega ; X_1(\omega) \in B_1, \dots, X_n(\omega) \in B_n).$$

すると

$$P^{\otimes n} = P \times \cdots \times P$$

$\mathcal{P} := \{ P_\theta ; \theta \in \Theta, P_\theta \text{は} (\mathcal{X}, \mathcal{B}(\mathcal{X}))\text{上の分布} \}$

は正則母数モデルとする.

(1) $\Theta \subset \mathbb{R}$ は開集合

(2) P_θ の p.m.f./p.d.f. を $p(\cdot \mid \theta)$ として、

$\{x \in \mathcal{X} ; p(x \mid \theta) > 0\}$ は θ に依存しない。

(3) $\frac{d}{d\theta} p(x \mid \theta)$, $\frac{d^2}{d\theta^2} p(x \mid \theta)$ は $x \in \mathcal{X}$ に関して連続

(4) $\frac{d}{d\theta} p(x \mid \theta)$, $\frac{d^2}{d\theta^2} p(x \mid \theta)$ は P_θ 可積分

(5) $\frac{d}{d\theta} p(x \mid \theta)$, $\frac{d^2}{d\theta^2} p(x \mid \theta)$ は θ の微分記号と

x の積分記号の交換が可.

記号　P_θ のもとの期待値と分散を $E_\theta[\cdot]$ $Var_\theta[\cdot]$ と書く.

定義 7.11　$\underline{X} := (X_1, X_2, \ldots, X_n) \sim P_\theta^{\otimes n}$ $(\theta \in \Theta \subset \mathbb{R})$ とする. 統計量 $T(\underline{X})$ が ${}^\forall\theta \in \Theta$ に対して

$$E_\theta[T(\underline{X})] = \theta$$

をみたすとき, $T(\underline{X})$ は θ の**不偏推定量**という

例 7.12 $n \geq 2$とし，$X_1, X_2, \ldots, X_n \overset{\text{i.i.d.}}{\sim} N(\theta, 1)$ ($\theta \in \Theta = \mathbb{R}$) とする．このとき，$\bar{X}_n = \frac{1}{n}\sum_{j=1}^{n} X_j$ は θ の不偏推定量である．

定理 7.13 $\underline{X} = (X_1, \ldots, X_n)^T \sim P_\theta^{\otimes n}$ ($\theta \in \Theta \subset \mathbb{R}$) で $T = T(\underline{X})$ は θ の十分統計量とする．$\hat{\theta}_n(\underline{X})$ は θ の任意の不偏推定量で，有限の分散 $\mathrm{Var}_\theta[\hat{\theta}_n(\underline{X})] < \infty$ を持つとする．推定量 $\tilde{\theta}_n(T)$ を

$$\tilde{\theta}_n(T) = E_\theta[\hat{\theta}(\underline{X}) \mid T]$$

により定めたとき，次の(1)と(2)が成立する.

(1) $\tilde{\theta}_n(T)$ は θ の不偏推定量.

(2) $Var_\theta[\tilde{\theta}_n(T)] \leq Var_\theta[\hat{\theta}_n(\underline{X})]$
$({}^\forall \theta \in \Theta)$.

(1)の証明　T は十分統計量なので，$\underline{X} \mid T = t$

の分布に θ に依存しない. よ

$p^{\underline{X} \mid T}(\underline{x} \mid t)$ をその p.d.f. とおく $\longrightarrow$ 離散型のときは

$\int \to \Sigma$

$$E_\theta[\hat{\theta}_n(\underline{X}) \mid T = t] = \underbrace{\int_{\mathcal{X}^n} \underbrace{\hat{\theta}_n(\underline{x})\, p^{\underline{X}|T}(\underline{x}|t)}_{\theta\text{ に関係ない}}\, d\underline{x}}_{\text{計算結果は } t \text{ の関数 たとえば } g(t) \text{ とおく}}$$

よって、

$$E_\theta[\hat{\theta}_n(\underline{X}) \mid T] := g(T)$$

と表現できるので、Tのみに依存する量.

定理 2.28(2) より

$$E_\theta[\tilde{\theta}_n(T)] = E_\theta[\, E_\theta[\hat{\theta}_n(\underline{X}) \mid T] \,]$$

$$= E_\theta[\hat{\theta}_n(\underline{X})] = \theta$$

よって，$\tilde{\theta}_n(T)$はθの不偏推定量．

(2)の証明　定理2.28(4)より

$$E_\theta[\{\hat{\theta}_n(\underline{X}) - \tilde{\theta}_n(T)\}\tilde{\theta}_n(T) \mid T]$$

$$= \tilde{\theta}_n(T) E_\theta[\hat{\theta}_n(\underline{X}) - \tilde{\theta}_n(T) \mid T]$$

$$= \tilde{\theta}_n(T)\{E_\theta[\hat{\theta}_n(\underline{X}) \mid T] - E_\theta[\tilde{\theta}_n(T) \mid T]\}$$

$$= \tilde{\theta}_n(T)\{\hat{\theta}_n(T) - \hat{\theta}_n(T)\} = 0. \quad (※)$$

となる. よって

$$\mathrm{Var}_\theta[\hat{\theta}_n(\underline{X})]$$

$$= E_\theta\left[\left\{\hat{\theta}_n(\underline{X}) - \underbrace{E_\theta[\hat{\theta}_n(\underline{X})]}_{=\theta}\right\}^2\right]$$

$$= E_\theta\left[\left\{\hat{\theta}_n(\underline{X}) - \tilde{\theta}_n(T) + \tilde{\theta}_n(T) - \theta\right\}^2\right]$$

$$= E_\theta\left[\left\{\hat{\theta}_n(\underline{X}) - \tilde{\theta}_n(T)\right\}^2\right]$$

$$+ E_\theta\left[\left\{\tilde{\theta}_n(T) - \theta\right\}^2\right]$$

$$+ 2 E_\theta\left[\{ \hat{\theta}_n(\underline{x}) - \tilde{\theta}_n(T) \} \{ \tilde{\theta}_n(T) - \theta \} \right]$$

$$= \underbrace{E_\theta\left[\{ \hat{\theta}_n(\underline{x}) - \tilde{\theta}_n(T) \} \tilde{\theta}_n(T) \right]}_{=0 \text{ by } (※)} - \theta \underbrace{E_\theta\left[\hat{\theta}_n(x) - \tilde{\theta}_n(T) \right]}_{=0}$$

$$= \underbrace{E\left[\{ \hat{\theta}_n(\underline{x}) - \tilde{\theta}_n(T) \}^2 \right]}_{\geq 0} + \mathrm{Var}_\theta\left[\tilde{\theta}_n(T) \right]$$

$$\geq \mathrm{Var}_\theta\left[\tilde{\theta}_n(T) \right]. \qquad (\forall \theta \in \Theta). \qquad \square$$

定義7.14 $\underline{X} = (X_1, X_2, \ldots, X_n)^T \sim P_\theta^{\otimes n}$ ($\theta \in \Theta \subset \mathbb{R}$) とする. θ の任意の不偏推定量 $\tilde{\theta}_n(\underline{X})$ に対して

$$\mathrm{Var}_\theta[\hat{\theta}_n(\underline{X})] \leq \mathrm{Var}_\theta[\tilde{\theta}_n(\underline{X})] \quad (\forall \theta \in \Theta)$$

をみたすとき. $\hat{\theta}_n(\underline{X})$ を θ の一様最小分散不偏推定量という UMVUE

注意7.15 存在するかどうかは不明だが. 存在すれば一意的.

いま

$$g : \mathbb{R} \longrightarrow \mathbb{R}$$

は連続微分可能とする

$X \sim P_\theta^{\otimes n}$ に基づき $g(\theta)$ の推定を考える

定理7.16 $\{P_\theta ; \theta \in \Theta \subset \mathbb{R}\}$ は正則統計モデルとする．

$\underline{X} \sim P_\theta^{\otimes n}$ $(\theta \in \Theta)$ とし，$\underline{X}$ の同時 p.d.f. / p.m.f.

を $p^{\underline{X}}(\cdot \mid \theta)$ とする．$\hat{\theta}_n := \hat{\theta}_n(\underline{X})$ を $g(\theta)$ の任意の

不偏推定量とし

$$A(\hat{\theta}, \theta) := \mathrm{Var}_\theta\left[\frac{p^{\underline{X}}(\underline{X} \mid \hat{\theta})}{p^{\underline{X}}(\underline{X} \mid \theta)}\right] \quad (\hat{\theta}, \theta \in \Theta)$$

$$A(\hat{\theta}, \theta) > 0 \quad (\hat{\theta} \neq \theta)$$

とおく．このとき

$$\mathrm{Var}_\theta[\hat{\theta}_n(\underline{X})] \geq \sup_{\tilde{\theta}\in\Theta} \frac{\{g(\theta)-g(\tilde{\theta})\}^2}{A(\tilde{\theta},\theta)}$$

が成り立つ.

証明は連続型の場合を書いておく

$\mathbb{X}^n$: $\underline{X}$ の値域

正則条件から θに依存しない

$\tilde{\mathbb{X}}^n := \{\underline{x} \in \mathbb{R}^n ; p^{\underline{X}}(\underline{x}|\theta) > 0\}$ とおく

$\hat{\theta}_n(\underline{X})$ は $g(\theta)$ の不偏推定量なので

$$E_\theta[\hat{\theta}_n(\underline{X})] = \int_{\mathcal{X}^n} \hat{\theta}_n(\underline{x})\, p^{\underline{X}}(\underline{x}|\theta)\, d\underline{x} = g(\theta)$$

となる. よって

Cauchy-Schwarz の不等式 を使いたい

$$E_\theta\left[\hat{\theta}_n(\underline{X})\left\{\frac{p^{\underline{X}}(\underline{x}|\hat{\theta})}{p^{\underline{X}}(\underline{x}|\theta)} - 1\right\}\right]$$

$$= \int_{\mathcal{X}^n} \hat{\theta}_n(\underline{x})\left\{\frac{p^{\underline{X}}(\underline{x}|\hat{\theta})}{p^{\underline{X}}(\underline{x}|\theta)} - 1\right\} p^{\underline{X}}(\underline{x}|\theta)\, d\underline{x}$$

$$= \int_{\mathcal{X}^n} \hat{\theta}_n(\underline{x})\, p^{\underline{X}}(\underline{x}|\hat{\theta})\, d\underline{x} - \int_{\mathcal{X}^n} \hat{\theta}_n(\underline{x})\, p^{\underline{X}}(\underline{x}|\theta)\, d\underline{x}$$

$= g(\tilde{\theta}) - g(\theta)$ ∵) $\hat{\theta}_n$の不偏性

$\nwarrow$ (7.7)

となる。一方

$$E_\theta\left[\frac{p^{\underline{X}}(\underline{x}|\tilde{\theta})}{p^{\underline{X}}(\underline{x}|\theta)}\right] = \int_{\mathcal{X}^n} \frac{p^{\underline{X}}(\underline{x}|\tilde{\theta})}{p^{\underline{X}}(\underline{x}|\theta)} p^{\underline{X}}(\underline{x}|\theta)\, d\underline{x}$$

$$= \int_{\mathcal{X}^n} p^{\underline{X}}(\underline{x}|\tilde{\theta})\, d\underline{x} = 1 \qquad (7.8)$$

と (7.7) より

$$g(\hat{\theta}) - g(\theta) \overset{(7.7)}{=} E_\theta\left[\hat{\theta}_n(\underline{x})\left\{\frac{p^{\underline{x}}(\underline{x}|\hat{\theta})}{p^{\underline{x}}(\underline{x}|\theta)} - 1\right\}\right]$$

$$= E_\theta\left[\left\{\hat{\theta}_n(\underline{x}) - E_\theta[\hat{\theta}_n(\underline{x})] + E_\theta[\hat{\theta}_n(\underline{x})]\right\}\left\{ \quad \right\}\right]$$

$$= E_\theta\left[\left\{\hat{\theta}_n(\underline{x}) - E_\theta[\hat{\theta}_n(\underline{x})]\right\}\left\{ \quad \right\}\right]$$

$$+ E_\theta[\hat{\theta}_n(\underline{x})]\, \underbrace{E_\theta\left[\frac{p^{\underline{x}}(\underline{x}|\hat{\theta})}{p^{\underline{x}}(\underline{x}|\theta)} - 1\right]}_{=0}$$

$\because$ (7.8)

$$= E_\theta\left[\left\{\hat{\theta}_n(\underline{x}) - E_\theta[\hat{\theta}_n(\underline{x})]\right\}\left\{\frac{p^{\underline{x}}(\underline{x}|\tilde{\theta})}{p^{\underline{x}}(\underline{x}|\theta)} - 1\right\}\right] \quad 65$$

を得る。ここで、Cauchy-Schwarz の不等式を用いる

$$|E[WY]| \le \sqrt{E[W^2]}\sqrt{E[Y^2]}$$

$$|g(\tilde{\theta}) - g(\theta)| = \left| \qquad \right|$$

$$\le \sqrt{E_\theta\left[\left\{\hat{\theta}_n(\underline{x}) - E_\theta[\hat{\theta}_n(\underline{x})]\right\}^2\right]} \times \sqrt{E_\theta\left[\left\{\frac{p^{\underline{x}}(\underline{x}|\tilde{\theta})}{p^{\underline{x}}(\underline{x}|\theta)} - 1\right\}^2\right]}$$

$E\left[\frac{p^{\underline{x}}(\underline{x}|\tilde{\theta})}{p^{\underline{x}}(\underline{x}|\theta)}\right] = 1$ ← (7.14)

$$= \sqrt{\mathrm{Var}_\theta[\hat{\theta}_n(\underline{x})]}\ \sqrt{\underbrace{\mathrm{Var}_\theta\left[\frac{p^{\underline{x}}(\underline{x}\mid\tilde{\theta})}{p^{\underline{x}}(\underline{x}\mid\theta)}\right]}_{=A(\tilde{\theta},\,\theta)}}$$

を得る. 以上から

$$\mathrm{Var}_\theta[\hat{\theta}_n(\underline{x})] \geq \frac{\{g(\tilde{\theta}) - g(\theta)\}^2}{A(\tilde{\theta},\,\theta)}$$

□

定理7.17（Cramér-Raoの不等式）次を仮定する

(1) $\{P_\theta : \theta \in \Theta\}$ は正則

(2) $g : \mathbb{R} \to \mathbb{R}$ は連続微分可能で、その導関数より

$\dot{g}(\theta) \neq 0$ をみたす。さらに

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

が存在する。ただし、

$$A(\tilde{\theta}, \theta) = Var_\theta \left[\frac{p^{\underline{x}}(\underline{x} \mid \tilde{\theta})}{p^{\underline{x}}(\underline{x} \mid \theta)} \right]$$

(3) $\forall \theta \in \Theta$ と十分小さな $\varepsilon > 0$ に対し

$\hat{\theta}$ を θ の ε 近傍から取ると、ある関数

$$G(\cdot \mid \theta) : \mathbb{R}^n \longrightarrow \mathbb{R}$$

が存在して

$$\left| \frac{p^{\underline{X}}(\underline{x} \mid \hat{\theta}) - p^{\underline{X}}(\underline{x} \mid \theta)}{(\hat{\theta} - \theta)\, p^{\underline{X}}(\underline{x} \mid \theta)} \right| < G(\underline{x}, \theta)$$

かつ

$$E_\theta[G^2(X \mid \theta)] < \infty$$

をみたす.

このとき

$$\mathrm{Var}_\theta\left[\hat{\theta}_n(\underline{x})\right] \geq \frac{\{\dot{g}(\theta)\}^2}{\tilde{h}_{\underline{x}}(\theta)} .$$

指数型分布のとき

$n\ddot{\psi}(\eta)$

ただし

$$\tilde{h}_{\underline{x}}(\theta) = E_\theta\left[\left\{\frac{\partial}{\partial\theta}\log p^{\underline{x}}(\underline{x}\mid\theta)\right\}^2\right]$$

Fisher情報量

証明　定理7.16と仮定(2)より

$$\mathrm{Var}_\theta[\hat{\theta}_n(X)] \geq \lim_{\tilde{\theta}\to\theta} \frac{\{g(\tilde{\theta})-g(\theta)\}^2}{A(\tilde{\theta},\theta)}$$

$$= \lim_{\tilde{\theta}\to\theta} \frac{\left\{\dfrac{g(\tilde{\theta})-g(\theta)}{\tilde{\theta}-\theta}\right\}^2}{\dfrac{A(\tilde{\theta},\theta)}{(\tilde{\theta}-\theta)^2}} = \frac{\{\dot{g}(\theta)\}^2}{J(\theta)}.$$

最後に

$$J(\theta) = \lim_{\hat{\theta} \to \theta} \frac{A(\hat{\theta}, \theta)}{(\hat{\theta} - \theta)^2}$$

$$= \lim_{\hat{\theta} \to \theta} \int_{\mathcal{X}^n} \left\{ \frac{p^{X}(\boldsymbol{x} \mid \hat{\theta})}{p^{X}(\boldsymbol{x} \mid \theta)} - 1 \right\}^2 \frac{p^{X}(\boldsymbol{x} \mid \theta)}{(\hat{\theta} - \theta)^2} \, d\boldsymbol{x}$$

$\because$ (3)

$$= \lim_{\hat{\theta} \to \theta} \int_{\mathcal{X}^n} \left\{ \frac{p^{X}(\boldsymbol{x} \mid \hat{\theta}) - p^{X}(\boldsymbol{x} \mid \theta)}{(\hat{\theta} - \theta)\, p^{X}(\boldsymbol{x} \mid \theta)} \right\}^2 p^{X}(\boldsymbol{x} \mid \theta) \, d\boldsymbol{x}$$

$$= \int_{\tilde{\mathcal{X}}^n} \left\{ \lim_{\hat{\theta}\to\theta} \frac{p^X(\underline{x}|\hat{\theta}) - p^X(\underline{x}|\theta)}{(\hat{\theta}-\theta)\, p^X(\underline{x}|\theta)} \right\}^2 p^X(\underline{x}|\theta)\, d\underline{x}$$

$$\to \frac{\partial}{\partial\theta} p^X(\underline{x}|\theta) \quad (\hat{\theta}\to\theta)$$

$$= \int_{\tilde{\mathcal{X}}^n} \left\{ \frac{\partial}{\partial\theta} \log p^X(\underline{x}|\theta) \right\}^2 p^X(\underline{x}|\theta)\, d\underline{x}$$

$$= \frac{1}{p^X(\underline{x}|\theta)} \frac{\partial}{\partial\theta} p^X(\underline{x}|\theta)$$

$$= F_{\underline{X}}(\theta)$$

となる. ただし

$$\tilde{\mathcal{X}}^n := \{ \underline{x} \in \mathcal{X}^n ;\ p^X(\underline{x}|\theta) > 0 \}$$

以上から；

$$\mathrm{Var}_\theta[\hat{\theta}_n(\underline{X})] \geq \frac{\{\dot{g}(\theta)\}^2}{\tilde{I}_X(\theta)}. \qquad \square$$

定義 7.19

$\hat{\theta}_n(\underline{X})$ は θ の R 有効

$$\iff \mathrm{Var}_\theta[\hat{\theta}(\underline{X})] = \frac{\{\dot{g}(\theta)\}^2}{\tilde{I}_X(\theta)}.$$