第6章　点推定法

6.1　点推定量の性質

6.1.1　近隣度 (closeness)

$X_1, X_2, \ldots, X_n$ を確率（密度）関数 $f_X(x|\theta)$ からのランダム標本とする．$\Theta \subset \mathbb{R}$ を θ の母数空間とし，確率（密度）関数 $f_X(x|\theta)$ は未知母数 θ のみに依存するものとする．$\tau : \Theta \to \mathbb{R}$ とし，$\boldsymbol{X} = (X_1, X_2, \ldots, X_n)'$ とおく．

定義 6.1 $t(\boldsymbol{X})$ と $t'(\boldsymbol{X})$ を $\tau(\theta)$ の推定量とする．$t'(\boldsymbol{X})$ が $t(\boldsymbol{X})$ より集中した $\tau(\theta)$ の推定量であるとは，すべての $\lambda > 0$ に対して

$$\mathbb{P}_\theta[|t'(\boldsymbol{X}) - \tau(\theta)| \le \lambda] \ge \mathbb{P}_\theta[|t(\boldsymbol{X}) - \tau(\theta)| \le \lambda], \qquad \theta \in \Theta$$

を満足することをいう．$t^* = t^*(\boldsymbol{X}) - \tau(\theta)$ が最も集中しているとは，他のどんな推定量より集中していることをいう．

注意 6.1 t^* は最も望ましい推定量であるが，ほとんどの場合は存在しない．

定義 6.2 (Pitman's closeness) $t'(\boldsymbol{X})$ が $t(\boldsymbol{X})$ より Pitman の意味で $\tau(\theta)$ に近い推定量であるとは

$$\mathbb{P}_\theta\left[|t'(\boldsymbol{X}) - \tau(\theta)| < |t(\boldsymbol{X}) - \tau(\theta)|\right] > \frac{1}{2}, \qquad \theta \in \Theta$$

を満足することである．

$t^*(\boldsymbol{X})$ が Pitman の意味で $\tau(\theta)$ の最近隣推定量であるとは，どんな推定量 $t(\boldsymbol{X})$ よりも Pitman の意味で $\tau(\theta)$ に近いものをいう．

注意 6.2 これもなかなか存在しない．

6.1.2　平均 2 乗誤差

定義 6.3 $t(\boldsymbol{X})$ を $\tau(\theta)$ の推定量とする．推定量 $t(\boldsymbol{X})$ の平均 *2* 乗誤差を

$$\mathbb{E}_\theta[(t(\boldsymbol{X}) - \tau(\theta))^2]$$

これを $MSE(\theta, t)$ と記すことにする．で定義する．

注意 6.3 $\mathbb{E}_\theta$ の θ は考えている標本が分布族の中のどこから来ているかを明示するために用いている．すなわち

$$\begin{aligned}&\mathbb{E}_\theta[(t(\boldsymbol{X}) - \tau(\theta))^2] \\ &\quad = \int \cdots \int [t(x_1, x_2, \ldots, x_n) - \tau(\theta)]^2 f(x_1|\theta) f(x_2|\theta) \times \cdots \times f(x_n|\theta)\, dx_1\, dx_2 \cdots dx_n\end{aligned}$$

である．

注意 6.4 $t_1(\boldsymbol{X})$ と $t_2(\boldsymbol{X})$ の平均 2 乗誤差はどもに θ の関数となる．

例 6.1 $\tau(\theta)=\theta,\, \Theta \in \mathbb{R}$ とする．$\theta_0 \in \Theta$ を固定し，

$$t_0(x_1,\, x_2,\, \dots,\, x_n) = \theta_0$$

なる推定量 $t_{\theta_0}(\boldsymbol{X})$ を考える．$t_{\theta_0}(\boldsymbol{X})$ の平均 2 乗誤差は

$$MSE\ (\theta,\, t_{\theta_0}) = \mathbb{E}_\theta[(t_{\theta_0}(\boldsymbol{X})-\theta)^2] = \mathbb{E}_\theta[(\theta_0-\theta)^2] = (\theta_0-\theta)^2$$

となる．すなわち，

$$MSE\ (t_{\theta_0},\, \theta_0) = 0$$

なる．

いま，すべての推定量 $t(\boldsymbol{X})$ に対して

$$MSE\ (\theta,\, t^*(\boldsymbol{X})) \le\ MSE\ (\theta,\, t(\boldsymbol{X})), \qquad \theta \in \Theta$$

を満足するような推定量 $t^*(\boldsymbol{X})$ が存在するならば，

$$MSE\ (\theta,\, t^*(\boldsymbol{X})) = 0$$

とならなければいけない[(6-1)]．これは常に θ を正しく推定できることある．したがって，θ がわかっていることと同じである．

θ に関して一様に MSE を最小にする推定量を見つけることができない理由はすべての推定量の集合は大きすぎるからである．したがって，対象とする推定量の集合を小さくしよう．

定義 6.4 推定量 $t(\boldsymbol{X})$ が $\tau(\theta)$ の不偏推定量であるとは，すべての $\theta \in \Theta$ に対して

$$\mathbb{E}_\theta[t(\boldsymbol{X})] = \tau(\theta)$$

を満足することである．

例 6.2 $t_{\theta_0}(\boldsymbol{X})$ は不偏推定量ではない．なぜならば，$\theta \neq \theta_0$ のとき

$$\mathbb{E}_\theta[t_{\theta_0}(\boldsymbol{X})] = \theta_0 \neq \theta$$

からわかる．

定理 6.1 $t(\boldsymbol{X})$ は $\tau(\theta)$ の推定量とし，$\mathbb{E}_\theta|t(\boldsymbol{X})|^2 < \infty$ とする．このとき，

$$MSE\ (\theta,\, t(\boldsymbol{X})) = \mathbb{VAR}(t(\boldsymbol{X})) + \{\mathbb{E}_\theta[t(\boldsymbol{X})] - \tau(\theta)\}^2$$

が成立する．さらに，$t(\boldsymbol{X})$ が $\tau(\theta)$ の不偏推定量ならば，

$$MSE\ (\theta,\, t(\boldsymbol{X})) = \mathbb{VAR}(t(\boldsymbol{X}))$$

が成立する．

証明

$$\begin{aligned}\mathrm{MSE}\ (\theta,\, t(\boldsymbol{X})) &= \mathbb{E}_\theta[\{t(\boldsymbol{X}) - \mathbb{E}_\theta[t(\boldsymbol{X})] + \mathbb{E}_\theta[t(\boldsymbol{X})] - \tau(\theta)\}^2] \\ &= \mathbb{E}_\theta[\{t(\boldsymbol{X}) - \mathbb{E}_\theta[t(\boldsymbol{X})]\}^2] \\ &\quad +2(\mathbb{E}_\theta[t(\boldsymbol{X})] - \tau(\theta))\mathbb{E}_\theta[t(\boldsymbol{X}) - \mathbb{E}_\theta[t(\boldsymbol{X})]] \\ &\quad +\{\mathbb{E}_\theta[t(\boldsymbol{X})] - \tau(\theta)\}^2 \\ &= \mathbb{VAR}(t(\boldsymbol{X})) + \{\mathbb{E}_\theta[t(\boldsymbol{X})] - \tau(\theta)\}^2\end{aligned}$$

□

定義 6.5 $BIAS = \mathbb{E}_\theta[t(\boldsymbol{X}) - \tau(\theta)]$ を推定量 $t(\boldsymbol{X})$ のバイアスという．

例 6.3 $X_1, X_2, \ldots, X_n (n \geq 2)$ を $N(\mu, \sigma^2)$ からのランダム標本とする．

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i, \qquad U^2 = \frac{1}{n-1}\sum_{i=}^n (X_i - \bar{X}_n)^2 \qquad S^2 = \frac{1}{n}\sum_{i=}^n (X_i - \bar{X}_n)^2$$

とおく．

$\mathbb{E}_{\mu,\sigma^2}[\bar{X}_n] = \mu$ より $\bar{X}_n$ は μ の不偏推定量である．また，μ の *MSE ($\tau(\mu, \sigma^2) = \mu$)* は

$$MSE\ (\mu, \bar{X}_n) = \mathbb{E}_{\mu,\sigma^2}[(\bar{X}_n - \mu)^2] = \frac{\sigma^2}{n}$$

である．

一方，

$$\mathbb{E}_{\mu,\sigma^2}[S^2] = \mathbb{E}_{\mu,\sigma^2}[\frac{1}{n}\sum_{i=1}^n (X_i - \mu)^2 - (\bar{X}_n - \mu)^2] = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n}\sigma^2$$

となり，σ^2 の最尤推定量は σ^2 の不偏推定量ではない．しかし，

$$\mathbb{E}_{\mu,\sigma^2}[U^2] = \mathbb{E}_{\mu,\sigma^2}[\frac{1}{n-1}\sum_{i=1}^n (X_i - \mu)^2 - \frac{n}{n-1}(\bar{X}_n - \mu)^2] = \frac{n}{n-1}\sigma^2 - \frac{1}{n-1}\sigma^2 = \sigma^2$$

となり，U^2 は σ^2 の不偏推定量である．

例 6.4 $X_1, X_2, \ldots, X_n$ をベルヌーイ試行(6-2) $\mathrm{Bi}(1, p), 0 < p < 1$ からのランダム標本とする．標本平均 $\bar{X}_n$ は p の不偏推定量となる．$\tau(p) = p$ として，$\bar{X}_n$ の MSE を求めよう．

$$MSE(p, \bar{X}_n) = \mathbb{E}[(\bar{X}_n - p)^2] = \mathbb{VAR}[\bar{X}_n] = \frac{p(1-p)}{n}$$

となる．いま，$Y = \sum_{i=1}^n X_i$ として，

$$\hat{p}_B = \frac{Y + \alpha}{\alpha + \beta + n}$$

なる推定量を考えよう．ただし，α と β は *(n に依存しない)* 定数とする．定理 6.1 を利用して，

$$\begin{aligned}MSE(p, \hat{p}_B) &= \mathbb{VAR}[\hat{p}_B] + \{\mathbb{E}[\hat{p}_B] - p\}^2 \\ &= \mathbb{VAR}\left[\frac{Y+\alpha}{\alpha+\beta+n}\right] + \left[\frac{np+\alpha}{\alpha+\beta+n}\right]^2 \\ &= \frac{1}{(\alpha+\beta+n)^2}\mathbb{VAR}[Y] + + \left[\frac{np+\alpha}{\alpha+\beta+n}\right]^2 \\ &= \frac{np(1-p)}{(\alpha+\beta+n)^2} + + \left[\frac{np+\alpha}{\alpha+\beta+n}\right]^2\end{aligned}$$

となる．ここで $\alpha = \beta = \sqrt{n/4}$ とおけば，

$$MSE(p, \hat{p}_B) = \frac{n}{4(\sqrt{n} + n^2}$$

となる[(6-3)]．

6.1.3 推定量の一致性

定義 6.6 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ とする．$\tau(\theta)$ の推定量の列 $\{t_n(\boldsymbol{X})\}_{n=1}^{\infty}$ が平均 2 乗誤差の意味で $\tau(\theta)$ の一致推定量であるとは，すべての $\theta \in \Theta$ に対して，

$$\lim_{n\to\infty} \mathbb{E}_\theta[(t_n(\boldsymbol{X}) - \tau(\theta))^2] = 0$$

を満足することである．

注意 6.5 推定量 $t_n(\boldsymbol{X})$ が平均 2 乗誤差の意味で一致性を持てば，定理 6.1 から

$$\lim_{n\to\infty} \mathbb{E}_\theta[t_n(\boldsymbol{X})] = \tau(\theta), \qquad \lim_{n\to\infty} \mathbb{VAR}_\theta[t_n(\boldsymbol{X})] = 0$$

となる．

例 6.5 $X_1, X_2, \ldots, X_n$ を平均 μ，分散 σ^2 の分布からのランダム標本とする．$\{\bar{X}_n\}_{n=1}^{\infty}$ を μ の推定量の列とすれば，

$$\mathbb{E}_{\mu, sigma^2}[\{\bar{X}_n - \mu\}^2] = \mathbb{VAR}[\bar{X}_n] = \frac{\sigma^2}{n}$$

よって，$\bar{X}_n$ は平均 2 誤差の意味で一致性を持つ．

定義 6.7 $\tau(\theta)$ の推定量の列 $\{t_n(\boldsymbol{X})\}_{n=1}^{\infty}$ が弱一致性を持つとは，どんな正の数 ϵ に対しても

$$\lim_{n\to\infty} \mathbb{P}_\theta[|t_n(\boldsymbol{X}) - \tau(\theta)| < \epsilon] = 1, \qquad {}^\forall\theta \in \Theta$$

が成立することである．簡単に，このような推定量 $t_n(\boldsymbol{X})$ を一致推定量という．

注意 6.6 $t_n(\boldsymbol{X})$ が平均 2 乗誤差の意味で一致推定量であれば，$t_n(\boldsymbol{X})$ は弱一致推定量である．なぜならば，*Chebyshev* の不等式から

$$\begin{aligned}
\mathbb{P}_\theta[|t_n(\boldsymbol{X}) - \tau(\theta)| < \epsilon] &= 1 - \mathbb{P}_\theta[|t_n(\boldsymbol{X}) - \tau(\theta)| \geq \epsilon] \\
&= 1 - \mathbb{P}_\theta[|t_n(\boldsymbol{X}) - \tau(\theta)|^2 \geq \epsilon^2] \\
&\geq 1 - \frac{1}{\epsilon^2}\mathbb{E}_\theta[(t_n(\boldsymbol{X}) - \tau(\theta))^2] \nearrow 1, \qquad n \to \infty
\end{aligned}$$

からわかる．

6.2 最尤法

確率ベクトル $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ の統計モデルを $\mathcal{P} = \{P_\theta : \theta \in \Theta \subset \mathbb{R}^k\}$ とする．$\mathcal{P}$ に含まれる P_θ に対応する確率密度関数もしくは確率関数を $p(\boldsymbol{x}|\,\theta)$ と記すことにする．ただし，$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ である．$\boldsymbol{X} = \boldsymbol{x}$ が観測されたときの尤度関数を

$$L_n(\theta|\,\boldsymbol{x}) = p(\boldsymbol{x}|\,\theta), \qquad \theta \in \Theta$$

で定めることにする．$L_n(\cdot\,|\,\boldsymbol{x})$ は標本空間から $\{\theta \mapsto p(\boldsymbol{x}|\,\theta) : \boldsymbol{x} \in S\}$ なる関数族への対応となる．$\boldsymbol{x}$ があたられたとき，$L_n(\theta|\,\boldsymbol{x})$ は θ の関数とみなす．これを簡単に $L_n(\theta)$ と書くことにする．$L_n(\theta)$ は $\boldsymbol{x}$ が与えられたとき，いろいろな θ の「確からしさ」もしくは「尤もらしさ」を表現するものである．

特に，$X_1, X_2, \ldots, X_n$ が独立同一に確率密度関数 $f(x|\,\theta)$ に従うならば，尤度関数は

$$L_n(\theta|\,\boldsymbol{x}) = \Pi_{i=1}^n f(x_i|\,\theta)$$

で与えられる．

最尤法とは，与えられたデータを実現させるために「尤もらしい」母数の値を母数の推定値として用いる手法である．すなわち，$\boldsymbol{X} = \boldsymbol{x}$ が与えられたとき，尤度関数を最大にする値 $\widehat{\theta}(\boldsymbol{x})$ を見つけることである：

$$L_n(\widehat{\theta}(\boldsymbol{x})|\,\boldsymbol{x}) = p(\boldsymbol{x}|\,\widehat{\theta}(\boldsymbol{x})) = \max\{p(\boldsymbol{x}|\,\theta) : \theta \in \Theta\} = \max\{L_n(\theta|\,\boldsymbol{x}) : \theta \in \Theta\}$$

$\widehat{\theta}(\boldsymbol{x})$ を θ の最尤推定値といい，$\widehat{\theta}(\boldsymbol{X})$ を θ の最尤推定量という．

例 6.6 確率変数 X が正規分布 $N(\theta, \sigma^2)$ に従うとする．ただし，σ^2 は既知とする．このとき，尤度関数は

$$L_1(\theta|\,x) = \frac{1}{\sigma}\psi\left(\frac{x-\theta}{\sigma}\right)$$

となる．ただし，$\psi(x) = (1/\sqrt{2\pi})e^{x^2/2}$ である．このとき，最大は

$$\widehat{\theta}(x) = x$$

のとき唯一達成される．したがって，$\widehat{\theta}(X) = X$ は最尤推定量となる．

つぎに，$X_1, X_2, \ldots, X_n$ は独立同一に正規分布 $N(\theta, \sigma^2)$ に従うとする．ここでも σ^2 は既知とする．このとき，尤度関数は

$$\begin{aligned} L_n(\theta|\,\boldsymbol{x}) &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{1}{2}\sum_{i=1}^n \frac{x_i-\theta}{\sigma^2}\right] \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{1}{2}\frac{(\bar{x}_n-\theta)^2}{\sigma^2/n}\right] \exp\left[-\frac{1}{2}\sum_{i=1}^n \frac{(x_i-\bar{x}_n)^2}{\sigma^2}\right] \end{aligned}$$

となる．よって，最大は

$$\widehat{\theta}(\boldsymbol{x}) = \bar{x}_n, \qquad \bar{x}_n = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$$

で達成される．したがって，最尤推定量は $\widehat{\theta}(\boldsymbol{X}) = \bar{X}_n$ となる．ただし，$\bar{X}_n = (1/n)(X_1 + X_2 + \cdots + X_n)$ である．

尤度関数に対数をとったものを対数尤度とよび，

$$l_n(\theta) = \log L_n(\theta|\,\boldsymbol{x})$$

と記す(6-4)．特に，$X_i,\ i = 1, 2, \ldots, n,$ が独立同一に確率密度関数 $f(x|\,\theta)$ に従う場合には

$$l_n(\theta) = \log \Pi_{i=1}^n f(x_i|\,\theta) = \sum_{i=1}^n \log f(x_i|\,\theta)$$

となる．

もし，Θ が開集合で $l_n(\theta)$ が θ に関して微分可能で $\widehat{\theta}(\boldsymbol{x})$ が存在するならば，$\widehat{\theta}(\boldsymbol{x})$ は方程式

$$\frac{\partial}{\partial \theta_i} l_n(\theta) = 0, \qquad i = 1, 2, \ldots, k$$

をみたす．この方程式を尤度方程式という．

例 6.7 標識 1, 2, 3 のどれかをもつ個体から構成される母集団を考える．それぞれの標識の出現確率は Hardy-Weinberg 比率で与えられるとする：

$$p(1|\theta) = \theta^2, \qquad p(2|\theta) = 2\theta(1-\theta), \qquad p(3|\theta) = (1-\theta)^2, \qquad 0 < \theta < 1$$

たとえば，3 つの個体を観測し，$x_1 = 1,\ x_2 = 2,\ x_3 = 1$ を得たとする．このとき

$$L_3(\theta|\,\boldsymbol{x}) = p(1|\,\theta)p(2|\,\theta)p(1|\,\theta) = 2\theta^5(1-\theta)$$

となる．尤度方程式は

$$\frac{\partial}{\partial \theta} l_3(\theta) = \frac{5}{\theta} - \frac{1}{1-\theta} = 0$$

となり，唯一の解 $\widehat{\theta} = 5/6$ を得る．これは

$$\frac{\partial^2}{\partial \theta^2} l_3(\theta) = -\frac{5}{\theta^2} - \frac{1}{(1-\theta)^2} < 0, \qquad 0 < \theta < 1$$

よりわかる．

一般に，n 個の観測 $x_1, x_2, \ldots, x_n$ を得たとする．いま

$$n_j = \#\{x_i = j : i = 1, 2, \ldots, n\}, \qquad j = 1, 2, 3$$

とする．尤度関数は

$$L_n(\theta|\,\boldsymbol{x}) = \theta^{2n_1}\{2\theta(1-\theta)\}^{n_2}(1-\theta)^{2n_3} = 2^{n_2}\theta^{2n_1+n_2}(1-\theta)^{n_2+2n_3}$$

より

$$\frac{\partial}{\partial \theta} l_n(\theta) = \frac{2n_1+n_2}{\theta} - \frac{n_2+2n_3}{1-\theta} = \frac{1}{\theta(1-\theta)}\{(2n_1+n_2) - 2(n_1+n_2+n_3)\theta\}$$

より，$2n_1+n_2 > 0,\ n_2+2n_3 > 0$ のとき，最尤推定値は唯一存在して，

$$\widehat{\theta}(\boldsymbol{x}) = \frac{2n_1+n_2}{2(n_1+n_2+n_3)}$$

となる．もし，$2n_1+n_2 = 0$ のとき，尤度関数は

$$2^{n_2}(1-\theta)^{(2n_1+n_2)+(n_2+2n_3)} = 2^{n_2}(1-\theta)^{2(n_1+n_2+n_3)}$$

となり，$\theta = 0$ のとき，尤度関数は最大になり，$\Theta = (0, 1)$ なので，最尤推定値は存在しない．また，$n_2+2n_3 = 0$ のときは，尤度関数は $2^{n_2}\theta^{2(n_1+n_2+n_3)}$ なり，最尤推定値は存在しない．

例 6.8 $X_1, X_2, \ldots, X_n$ は独立同一分布に従い，各 $X_i, i = 1, 2, \ldots, n,$ は $1, 2, \ldots, k$ の値をとり，その確率は $\theta_j = \mathbb{P}\{X_i = j\}, j = 1, 2, \ldots, k,$ で与えられるとする．ここで，$n \geq k-1$ を仮定する．いま，$N_j = \sum_{i=1}^n I\{X_i = j\}$ おく．$\boldsymbol{X} = \boldsymbol{x}$ が与えられたとする．このとき，$n_j = \sum_{i=1}^n I\{x_i = j\}$ とおけば，対数尤度は

$$l_n(\boldsymbol{\theta}) = \log L_n(\boldsymbol{\theta}|\,\boldsymbol{x}) = \log p(\boldsymbol{x}|\,\boldsymbol{\theta}) = \sum_{j=1}^k n_j \log \theta_j$$

となる．ただし，$\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$ で

$$\theta_k = 1 - \sum_{j=1}^{k-1} \theta_j \tag{6.1}$$

である．

まず，$n_j > 0, j = 1, 2, \ldots, k,$ を仮定する．このとき，θ_j のどれかがゼロならば，$p(\boldsymbol{x}|\,\boldsymbol{\theta}) = 0$ となる．したがって，最尤推定値は $\theta_j > 0$ となるので，上の仮定のもとでは，最尤推定値は $[0, 1]^k$ の内点である．したがって，最尤推定値は尤度方程式

$$\frac{\partial}{\partial \theta_j} l_n(\boldsymbol{\theta}) = \frac{\partial}{\partial \theta_j} \sum_{\ell=1}^k n_\ell \log \theta_\ell = \sum_{\ell=1}^k \frac{n_\ell}{\theta_\ell} \frac{\partial \theta_\ell}{\partial \theta_j} = 0, \qquad j = 1, 2, \ldots, k-1$$

となる．(6.1) から $\partial \theta_k / \partial \theta_j = -1, j = 1, 2, \ldots, k-1,$ となる．よって，尤度方程式は

$$\frac{\widehat{\theta}_k}{\widehat{\theta}_j} = \frac{n_k}{n_j}, \qquad j = 1, 2, \ldots, k-1$$

となる．これを再度 (6.1) に代入すれば，$\widehat{\theta}_k = n_k/n$ となり，

$$\widehat{\theta}_j = \frac{n_j}{n}, \qquad j = 1, 2, \ldots, k$$

となる．ただし，$n = \sum_{\ell=1}^k n_\ell$ とした．つぎに，$\theta_j = n_j/n, j = 1, 2, \ldots, k$ が実際に $l_n(\boldsymbol{\theta})$ を最大にしていることを確認するために，$l_n(\boldsymbol{\theta})$ は $(\theta_1, \theta_2, \ldots, \theta_{k-1})$ に関して concave であることを示す．$1 \leq r \leq k-1, 1 \leq j \leq k-1$ に対して，

$$\frac{\partial}{\partial \theta_r} \frac{\partial}{\partial \theta_j} l_n(\boldsymbol{\theta}) = \frac{\partial}{\partial \theta_r} \left(\frac{n_j}{\theta_j} - \frac{n_k}{\theta_k} \right) = \begin{cases} -\left(\frac{n_r}{\theta_r^2} + \frac{n_k}{\theta_k^2} \right) < 0, & r = j \\ -\frac{n_k}{\theta_k^2} < 0, & r \neq j \end{cases}$$

となる．

ある j に対して，$n_j = 0$ のとき，$\widehat{\theta}_j = n_j/n$ が最尤推定値であることも確認することができる．

定理 6.2 $T(\boldsymbol{X})$ を未知母数 θ の十分統計量とする．このとき，θ の最尤推定量が一意に存在するならば，θ の最尤推定量は T の関数である．

証明 $p(\boldsymbol{x}|\,\theta)$ を確率関数または確率密度関数とする．因子分解定理から θ と T を通してのみ $\boldsymbol{x}$ に依存する関数 g と $\boldsymbol{x}$ のみに依存する関数 h が存在して，

$$p(\boldsymbol{x}|\,\theta) = h(\boldsymbol{x}) g(T(\boldsymbol{x})|\,\theta)$$

と書ける．これより θ に関して $p(\boldsymbol{x}|\,\theta)$ を最大化することは θ に関して $g(T(\boldsymbol{x})|\,\theta)$ を最大化することと同値になる．また，$T(\boldsymbol{x}) = t$ と与えられたとき，$g(t|\,\theta)$ が 2 つ以上 θ で最大になるとすれば，それに対応する $\boldsymbol{x}$ において $g(T(\boldsymbol{x})|\,\theta)$ も 2 つ以上の θ で最大化されるので，仮定と矛盾する．したがって，$g(t|\,\theta)$ を θ について最大にする点はひとつである． □

定理 6.3 関数 $g:\mathbb{R}\to\mathbb{R}$ が 1 対 1 のとき，$\widehat{\theta}$ が θ の最尤推定量であれば，$g(\widehat{\theta})$ は $g(\theta)$ の最尤推定量である．

証明 g は 1 対 1 だから，逆関数 g^{-1} が存在し，$\tau=g(\theta)$ のとき，$\theta=g^{-1}(\tau)$ となる．これより

$$L_n(\theta|\,\boldsymbol{x})=p(\boldsymbol{x}|\,\theta)=p(\boldsymbol{x}|\,g^{-1}(\tau))=\tilde{L}_n(\tau|,\boldsymbol{x})$$

と書けるので，

$$\sup_\tau \tilde{L}_n(\tau|,\boldsymbol{x})=\sup_\tau L_n(g^{-1}(\tau)|,\boldsymbol{x})=\sup_\theta L_n(\theta|\,\boldsymbol{x})$$

よって，$\widehat{\theta}=g^{-1}(\widehat{\tau})$ のとき最大化される．したがって，$\widehat{\tau}=g(\widehat{\theta})$ のとき最大化される． □

例 6.9 （指数分布族の最尤推定量） $\boldsymbol{X}=(X_1,\,X_2,\,\ldots\,,\,X_n)'$ は k 母数指数分布族

$$f(x|\,\boldsymbol{\eta})=S(x)\exp[\sum_{j=1}^k \eta_j T_j(x)-A(\boldsymbol{\eta})]$$

からの大きさ n のランダム標本とする．ただし，$\boldsymbol{\eta}=(\eta_1,\,\eta_2,\,\ldots\,,\,\eta_k)$ とし，自然母数空間 $\eta\in\mathcal{E}$ は $\mathbb{R}^k$ の開集合とする．である．このとき，$\boldsymbol{X}$ の同時確率（密度）関数は

$$\begin{aligned} p(\boldsymbol{x}|\,\boldsymbol{\eta}) &= \Pi_{i=1}^n f(x_i|\,\boldsymbol{\eta}) \\ &= \Pi_{i=1}^n S(x_i)\exp[\sum_{j=1}^k n\eta_j \bar{T}_j(\boldsymbol{x})-nA(\boldsymbol{\eta})] \end{aligned}$$

となる．ただし，$\bar{T}_j(\boldsymbol{x})=\sum_{i=1}^n T_j(x_i),\ j=1,\,2,\,\ldots\,,\,k$ である．したがって，対数尤度は

$$l_n(\boldsymbol{\eta})=n\sum_{j=1}^k \eta_j\bar{T}_j(\boldsymbol{x})-nA(\boldsymbol{\eta})+（定数項）$$

となる．系 **??** を用いれば，

$$\frac{\partial}{\partial\eta_j}l_n(\boldsymbol{\eta})=n\bar{T}_j(\boldsymbol{x})-n\frac{\partial}{\partial\eta_j}A(\boldsymbol{\eta})=n\bar{T}_j(\boldsymbol{x})-n\mathbb{E}_\eta[T_j(X)]$$

と

$$\frac{\partial^2}{\partial\eta_j\eta_m}l_n(\boldsymbol{\eta})=-n\frac{\partial^2}{\partial\eta_j\eta_m}A(\boldsymbol{\eta})=-n\mathbb{COV}_\eta(T_j(X),\,T_m(X))$$

となるので，行列 $((\partial^2/\partial\boldsymbol{\eta}'\boldsymbol{\eta})A(\boldsymbol{\eta}))$ は負の定符号となる．したがって，尤度方程式

$$\frac{1}{n}\sum_{i=1}^n T_j(x_i)=\frac{\partial}{\partial\eta_j}A(\boldsymbol{\eta})$$

または

$$\frac{1}{n}\sum_{i=1}^n T_j(x_i)=\mathbb{E}_\eta[T_j(X)],\qquad j=1,\,2,\,\ldots\,,\,k$$

は唯一の解を持ち，これは $\boldsymbol{\eta}$ の最尤推定値となる．

6.3 不偏推定と Cramér-Rao の定理

$X_1, X_2, \ldots, X_n$ を確率（密度）関数 $f(x|\theta), \theta \in \Theta$ からのランダム標本とし，$T = T(\boldsymbol{X})$ を標本 $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ に基づく $\tau(\theta)$ の推定量とする．T の MSE は

$$\mathbb{E}_\theta[\{T(\boldsymbol{X}) - \tau(\theta)\}^2] = \mathbb{VAR}\theta[T(\boldsymbol{X})] + \{\mathbb{E}_\theta[T(\boldsymbol{X})] - \tau(\theta)\}^2$$

と書けた．さらに，T が $\tau(\theta)$ の不偏推定量であれば，

$$\mathbb{E}_\theta[\{T((\boldsymbol{X}) - \tau(\theta)\}^2] = \mathbb{VAR}\theta[T(\boldsymbol{X})]$$

となる．

定義 6.8 (一様最小分散不偏推定量　Uniformly Minimum Variance Unbiased Estimator)　$X_1, X_2, \ldots, X_n$ を確率（密度）関数 $f(x|,\theta), \theta \in \Theta$ からのランダム標本とし，$\tau(\theta)$ の推定量 $T^* = T^*(X_1, X_2, \ldots, X_n)$ が $\tau(\theta)$ の一様最小分散不偏推定量（略して UMVUE）であるとはつぎを満足することである．

(i)　$\mathbb{E}_\theta[T^*] = \tau(\theta)$

(ii)　他の $\tau(\theta)$ の不偏推定量 $T = T(X_1, X_2, \ldots, X_n)$ に対し

$$\mathbb{VAR}_\theta[T^*] \le \mathbb{VAR}_\theta[T].$$

6.3.1 不偏推定量の分散の下限

$X_1, X_2, \ldots, X_n$ を確率密度関数 $f(x|\theta), \theta \in \Theta \subset \mathbb{R}$ からのランダム標本とする．以下の仮定をおく．

(A1)　集合 $\{x :\ f(x|\theta) > 0\}$ は θ に依存しない．

(A2)　すべての x と θ に対し，$(\partial/\partial\theta)f(x|\theta)$ は存在する．

(A3)　$\mathbb{E}_\theta[T(X_1, X_2, \ldots, X_n)] < \infty$ なる推定量 T に対し

$$\frac{\partial}{\partial\theta}\int\cdots\int T(x_1, x_2, \ldots, x_n)\Pi_{i=1}^n f(x_i|\theta)\,dx_i = \int\cdots\int T(x_1, x_2, \ldots, x_n)\frac{\partial}{\partial\theta}\Pi_{i=1}^n f(x_i|\theta)\,dx_i.$$

(A4)　すべての $\theta \in \Theta$ に対し

$$0 < \mathbb{E}_\theta\left[\left\{\frac{\partial}{\partial\theta}\log f(X_1|\theta)\right\}^2\right] < \infty.$$

(A5)　$\tau(\theta)$ は微分可能．このとき，θ に関して一様に

$$\mathbb{VAR}_\theta[T] \ge \frac{[\tau'(\theta)]^2}{\mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\theta)\right)^2\right]}.$$

が成立する．また，等号成立は

$$\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}\,\theta) = K(\theta, n)[t(\boldsymbol{X} - \tau(\theta)] \tag{6.2}$$

のときに限る．

証明 $f(\boldsymbol{x}|\,\theta) = \Pi_{i=1}^n f(x_i|\,\theta)$ と $d\boldsymbol{x} = dx_1 dx_2 \cdots dx_n$ 記すことにする．

$$\begin{aligned}
\tau'(\theta) &= \frac{\partial}{\partial\theta}\int\cdots\int T(\boldsymbol{x}) f(\boldsymbol{x}|\,\theta)\,d\boldsymbol{x} \\
&= \int\cdots\int T(\boldsymbol{x})\frac{\partial}{\partial\theta} f(\boldsymbol{x}|\,\theta)\,d\boldsymbol{x} - \tau(\theta)\frac{\partial}{\partial\theta}\int\cdots\int f(\boldsymbol{x}|\,\theta)\,d\boldsymbol{x} \\
&= \int\cdots\int [T(\boldsymbol{x}) - \tau(\theta)]\frac{\partial}{\partial\theta} f(\boldsymbol{x}|\,\theta)\,d\boldsymbol{x} \\
&= \int\cdots\int [T(\boldsymbol{x}) - \tau(\theta)]\left(\frac{\partial}{\partial\theta}\log f(\boldsymbol{x}|\,\theta)\right) f(\boldsymbol{x}|\,\theta)\,d\boldsymbol{x} \\
&= \mathbb{E}_\theta\left[\{T(\boldsymbol{X}) - \tau(\theta)\}\left(\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\,\theta)\right)\right]
\end{aligned}$$

となる．Cauchy-Schwarz の不等式から

$$[\tau'(\theta)]^2 \le \mathbb{E}_\theta[\{T(\boldsymbol{X}) - \tau(\theta)\}^2]\,\mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\,\theta)\right)^2\right]$$

よって

$$\mathbb{VAR}_\theta[T] \ge \frac{[\tau'(\theta)]^2}{\mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\,\theta)\right)^2\right]}.$$

しかし

$$\begin{aligned}
\mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\,\theta)\right)^2\right] &= \mathbb{E}_\theta\left[\left\{\sum_{i=1}^n \frac{\partial}{\partial\theta}\log f(X_i|\,\theta)\right\}\right] \\
&= \sum_{i=1}^n\sum_{j=1}^n \mathbb{E}_\theta\left[\frac{\partial}{\partial\theta}\log f(X_i|\,\theta)\frac{\partial}{\partial\theta}\log f(X_j|\,\theta)\right] \\
&= n\mathbb{E}_\theta\left[\left\{\frac{\partial}{\partial\theta}\log f(X_1|\,\theta)\right\}^2\right].
\end{aligned}$$

なぜならば，$i \ne j$ のとき X_i と X_j は独立であることと

$$\begin{aligned}
\mathbb{E}_\theta\left[\frac{\partial}{\partial\theta}\log f(X_1|\,\theta)\right] &= \int\left(\frac{\partial}{\partial\theta}\log f(x|\,\theta)\right) f(x|\,\theta)\,dx = \int\frac{\partial}{\partial\theta} f(x|\,\theta)\,dx \\
&= \frac{\partial}{\partial\theta}\int f(x|\,\theta)\,dx = 0
\end{aligned}$$

最後に Cauchy-Schwarz の不等式の等号成立条件から

$$\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}\,\theta) = K(\theta, n)[t(\boldsymbol{x}) - \tau(\theta)] + b(\theta) \tag{6.3}$$

が等号成立条件となることがわかる．しかし，両辺の期待値をとれば，

$$0 = \mathbb{E}_\theta\left[\frac{\partial}{\partial\theta}\log f(\boldsymbol{X}\,\theta)\right] = K(\theta, n)\mathbb{E}_\theta\left[t(\boldsymbol{x}) - \tau(\theta)\right] + b(\theta) = b(\theta)$$

となるので，$b(\theta) = 0$ がわかる．よって，等号成立条件は (6.2) であることが示せた． □

注意 6.7 T が $\tau(\theta)$ の一様最小分散不偏推定量ならば，一意である．これは以下の議論からわかる．T' を別の $\tau(\theta)$ の一様最小分散不偏推定量とし，$T^* = (1/2)(T + T')$ とおく．すると，$\mathbb{E}[T^*] = \tau(\theta)$ となり，T^* も $\tau(\theta)$ の不偏推定量となる．したがって，すべての θ に対して，

$$\mathbb{VAR}_\theta(T^*) \ge \mathbb{VAR}_\theta(T)$$

となる．しかし，

$$\begin{aligned}\mathbb{VAR}_\theta(T^*) &= \mathbb{VAR}_\theta((1/2)T+(1/2)T') \\ &= \frac{1}{4}\mathbb{VAR}_\theta(T)+\frac{1}{4}\mathbb{VAR}_\theta(T')+\frac{1}{2}\mathbb{COV}_\theta(T,\,T') \\ &\le \frac{1}{4}\mathbb{VAR}_\theta(T)+\frac{1}{4}\mathbb{VAR}_\theta(T')+\frac{1}{2}\sqrt{\mathbb{VAR}_\theta(T)}\sqrt{\mathbb{VAR}_\theta(T')} \\ &= \mathbb{VAR}_\theta(T) \qquad (6.4)\end{aligned}$$

となる．不等号は Cauchy-Schwartz の不等式からわかり，最後の等号は $\mathbb{VAR}_\theta(T)=\mathbb{VAR}_\theta(T')$ よりわかる．(6.4) と (6.4) から

$$\mathbb{VAR}_\theta(T)=\mathbb{VAR}_\theta(T^*) \qquad (6.5)$$

が成立することがわかる．したがって，(6.4) の Cauchy-Schwartz の不等式は等号が成立している．Cauchy-Schwartz の不等式の等号成立条件から，ある定数 $a(\theta)$ と $b(\theta)$ が存在して

$$T'=a(\theta)T+b(\theta)$$

と書ける．しかし，$\mathbb{VAR}_\theta(T')=a(\theta)\mathbb{VAR}_\theta(T)$ と (6.5) から $a(\theta)=1$ となる．さらに，

$$\tau(\theta)=\mathbb{E}_\theta(T')=\mathbb{E}_\theta(T)+b(\theta)=\tau(\theta)+b(\theta)$$

から $b(\theta)=0$ となる．したがって，$T=T'$ となることがわかる．

例 6.10 $X_1,\,X_2,\,\dots,\,X_n$ を指数分布 $f(x|\,\theta)=\theta e^{-\theta x}I_{(0,\,\infty)}(x)$ からの大きさ n のランダム標本とする．ただし，$\theta>0$ である．

まず，$\tau(\theta)=\theta$ とする．θ の推定量 $T=T(X_1,\,X_2,\,\dots,\,X_n)$ の Cramér-Rao) の下限は

$$\mathbb{VAR}_\theta(T)\ge\frac{1}{n\mathbb{E}_\theta[\{(\partial/\partial\theta)\log f(X_1|\,\theta)\}^2]}$$

となる．さらに，

$$\frac{\partial}{\partial\theta}\log f(x|\,\theta)=\frac{1}{\theta}-x$$

より

$$\mathbb{E}_\theta\left[\left\{\frac{\partial}{\partial\theta}\log f(X_1|\,\theta)\right\}^2\right]=\mathbb{E}_\theta\left[\left\{\frac{1}{\theta}-X_1\right\}^2\right]=\mathbb{VAR}(X)=\frac{1}{\theta^2}$$

となる(6-5)．したがって，

$$\mathbb{VAR}_\theta\ge\frac{1}{n(1/\theta^2)}=\frac{\theta^2}{n}$$

となる．

つぎに，$\tau(\theta)=1/\theta$ とする．$\tau(\theta)=1/\theta)$ の推定量 $S=S(X_1,\,X_2,\,\dots,\,X_n)$ の Cramér - Rao の下限は

$$\mathbb{VAR}_\theta(S)\ge\frac{\{\tau'(\theta)\}^2}{n(1/\theta^2)}=\frac{1}{n\theta^2}$$

となる．さらに，

$$\sum_{i=1}^n\frac{\partial}{\partial\theta}\log f(X_i|\,\theta)=\sum_{i=1}^n\frac{\partial}{\partial\theta}(\log\theta-\theta X_i)=\sum_{i=1}^n\left(\frac{1}{\theta}-X_i\right)=-n\left(\bar{X}_n-\frac{1}{\theta}\right)$$

より，(6.2) において $K(\theta,\,n)=-n$ とおけば，$S=\bar{X}_n$ は $1/\theta$ の UMVUE となる．

例 6.11 $X_1, X_2, \ldots, X_n$ をポアソン分布 $f(x|\theta) = e^{-\lambda}\lambda^x/x!(x = 0, 1, ldots)$ からの大きさ n のランダム標本とする．ただし，$\theta > 0$ である．

$\tau(\theta) = e^{-\theta} = \mathbb{P}[X_1 = 0]$ とする．

$$\frac{\partial}{\partial\theta}\log f(x|\theta) = -1 + \frac{x}{\theta}$$

に注意すれば，

$$\mathbb{E}_\theta\left[\left\{\frac{\partial}{\partial\theta}\log f(X_1|\theta)\right\}^2\right] = \mathbb{E}_\theta\left[\left\{\frac{X_1}{\theta} - 1\right\}^2\right] = \frac{1}{\theta^2}\mathbb{E}_\theta[(X_1 - \theta)^2] = \frac{1}{\theta}$$

となる(6-6)．よって，$\tau(\theta)$ の推定量推定量 $T = T(X_1, X_2, \ldots, X_n)$ の Cramér-Rao) の下限は

$$\mathbb{VAR}_\theta(T) \geq \frac{(-e^{-\theta})^2}{n(1/\theta)} = \frac{\theta e^{-2\theta}}{n}$$

となる．

いま，$\tau(\theta)$ の推定量推定量

$$T_0 = \frac{1}{n}\sum_{i=1}^{n} I\{X_i = 0\}$$

を考える．

$$\mathbb{E}_\theta[T_0] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_\theta[I\{X_i = 0\}] = \frac{1}{n}\sum_{i=1}^{n}\big[1\cdot\mathbb{P}_\theta(X_i = 0) + 0\cdot\mathbb{P}_\theta(X_i \neq 0)\big] = e^{-\theta}$$

となる．したがって，T_0 は $\tau(\theta) = e^{-\theta}$ の不偏推定量となる．また，

$$\begin{aligned}
\mathbb{VAR}_\theta &= \mathbb{E}_\theta\left[\left\{\frac{1}{n}\sum_{i=1}^{n} I\{X_i = 0\}\right\}^2\right] - \{\mathbb{E}_\theta(T_0)\}^2 \\
&= \sum_{i=1}^{n}\mathbb{E}_\theta\left[\frac{1}{n^2}I\{X_i = 0\}\right] + \sum_{i\neq j}\mathbb{E}_\theta\left[\frac{1}{n^2}I\{X_i = 0\}I\{X_j = 0\}\right] - e^{-2\theta} \\
&= \frac{1}{n}e^{-\theta} + \frac{n(n-1)}{n^2}e^{-2\theta} - e^{-2\theta} = \frac{1}{n}e^{-\theta}(1 - e^{-\theta})
\end{aligned}$$

となる．$\theta > 0$ に対し

$$\frac{1}{n}e^{-\theta}(1 - e^{-\theta}) \geq \frac{1}{n}\theta e^{-2\theta}$$

である(6-7)．

つぎに，$\tau(\theta) = \theta$ とおく．

$$\sum_{i=1}^{n}\frac{\partial}{\partial\theta}\log f(X_i|\theta) = \sum_{i=1}^{n}\left(-1 + \frac{X_i}{\theta}\right) = \frac{n}{\theta}(\bar{X}_n - \theta)$$

となり，(6.2) において $K(\theta, n) = n/\theta$ とおけば，$\bar{X}_n$ は $\tau(\theta) = \theta$ の UMVUE となることがわかる．

6.4 十分性と完備性

定理 6.4 (Rao-Blackwell) $X_1, X_2, \ldots, X_n$ を確率（密度）関数 $f(x|\theta), \theta \in \Theta$ からの大きさ $n\,(n \geq 2)$ のランダム標本とし，$S_i = S_i(X_1, X_2, \ldots, X_n), i = 1, 2, \ldots, k$ を十分統計量とする．$T = T(X_1, X_2, \ldots, X_n)$ を $\tau(\theta)$ の不偏推定量で $\mathbb{VAR}(T) < \infty$ なるものとする．推定量 $T' = T'(X_1, X_2, \ldots, X_n)$ を $T' = \mathbb{E}[T|\, S_1, \ldots, S_k]$ で定義する．このとき，

(i) T' は統計量で $S_1, \ldots, S_k$ を通してのみ $(X_1, X_2, \ldots, X_n)$ に依存する．したがって，$T' = T'(S_1, \ldots, S_k)$ を記すことにする．

(ii) $\mathbb{E}_\theta[T'] = \tau(\theta)$ となる．すなわち，T' は $\tau(\theta)$ の不偏推定量である．

(iii) すべての $\theta \in \Theta$ に対し

$$\mathbb{VAR}(T') \leq \mathbb{VAR}(T)$$

が成立する．ある θ に対し，$\mathbb{P}(T = T') \neq 1$ ならば，

$$\mathbb{VAR}(T') < \mathbb{VAR}(T)$$

が成立する．

証明 $S_1, \ldots, S_k$ の十分性から $S_1, \ldots, S_k$ を与えたときの $(X_1, X_2, \ldots, X_n)$ の条件付き分布は θ に依存しない．したがって，$S_1, \ldots, S_k$ の十分性から $S_1, \ldots, S_k$ を与えたときの T の条件付き分布は θ に依存しないので，$\mathbb{E}_\theta[T|\, S_1, \ldots, S_k]$ は θ に依存しないので，統計量となる．また，これは $S_1, \ldots, S_k$ の関数となることは明らである．

$$\mathbb{T}' = \mathbb{E}_\theta[\mathbb{E}[T|\, S_1, \ldots, S_k]] = \mathbb{E}_\theta[T] = \tau(\theta)$$

となる．したがって，T' は $\tau(\theta)$ の不偏推定量となる．

$$\begin{aligned}
\mathbb{VAR}_\theta[T] &= \mathbb{E}_\theta[(T - \mathbb{E}_\theta(T))^2] \\
&= \mathbb{E}_\theta[(T - T' + T' - \mathbb{E}_\theta(T))^2] \\
&= \mathbb{E}_\theta[\{T - T'\}^2] + 2\mathbb{E}_\theta[(T - T')(T' - \mathbb{E}_\theta(T))] + \mathbb{VAR}_\theta(T')
\end{aligned}$$

となる．しかし，

$$\begin{aligned}
\mathbb{E}_\theta[(T - T')(T' - \mathbb{E}_\theta(T))] &= \mathbb{E}_\theta[\mathbb{E}[(T - T')(T' - \mathbb{E}_\theta(T))|\, S_1, \ldots, S_k]] \\
&= \mathbb{E}_\theta[(T' - \mathbb{E}_\theta(T))\mathbb{E}[T - T'|\, S_1, \ldots, S_k]] = 0
\end{aligned}$$

となる．よって，

$$\mathbb{VAR}_\theta[T] = \mathbb{E}_\theta[\{T - T'\}^2] + \mathbb{VAR}_\theta(T') \geq \mathbb{VAR}_\theta(T')$$

を得る．さらに，

$$\mathbb{E}_\theta[\{T - T'\}^2] \Longleftrightarrow \mathbb{P}_\theta(T - T' = 0) = 1$$

に注意すればよい． □

注意 6.8 Rao-Blackwell の定理の証明の最後の部分は以下のように確認できる．W を確率変数とし，$\mathbb{E}[W^2] < \infty$ とする．このとき，

$$\mathbb{E}[W^2] \Longleftrightarrow P(W = 0) = 1$$

を示せばよい．まず，$\Leftarrow$ は自明である．したがって，$\Rightarrow$ を示せばよい．任意の正数 ϵ に対し

$$\mathbb{E}[W^2] \geq \mathbb{E}[W^2 I\{W \geq \epsilon\}] \geq \epsilon^2 \mathbb{E}[I\{W \geq \epsilon\}] = \epsilon^2 \mathbb{P}\{W \geq \epsilon\}$$

となることより，$\mathbb{P}\{W \geq \epsilon\} = 0$ となる．同様にすれば，$\mathbb{P}\{W \leq \epsilon\} = 0$ を得る．したがって，$\mathbb{P}\{-\epsilon \leq W \leq \epsilon\} = 1$ となる．いま，$A_n = \{-1/n \leq W \leq 1/n\}\,(n = 1, 2, \ldots)$ とおけば，$A_1 \supset A_2 \supset \cdots$ となる．また，$\{W = 0\} = \cap_{n=1}^{n} A_n$ であり，$\mathbb{P}_\theta(A_n) = 1$ である．これらより

$$\mathbb{P}_\theta(W = 0) = \mathbb{P}_\theta\{\cap_{n=1}^{n} A_n\} = \lim_{n\to\infty} P(A_n) = 0$$

となることよりわかる．

例 6.12 $X_1, X_2, \ldots, X_n$ をベルヌーイ試行 $f(x|\,\theta) = \theta^x(1-\theta)^{1-x} I_{\{0,1\}}(x)$ からの大きさ $n(n \geq 2)$ のランダム標本とし，$\tau(\theta) = \theta$ の推定を考える．明らかに，$T(X_1, X_2, \ldots, X_n) = X_1$ は θ の不偏推定量になる．しかし，この推定量は効率がわるいので，十分統計量 $S = \sum_{i=1}^n X_i$ を用いて，修正することを考える．

$$\tilde{T} = \mathbb{E}_[X_1|\,S]$$

とおく．この推定量を具体的に計算してみよう．

$$\begin{aligned}
\mathbb{P}[X_1 = 0|\sum_{i=1}^n X_i = s] &= \frac{\mathbb{P}[X_1 = 0 \quad \text{かつ} \quad \sum_{i=1}^n X_i = s]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
&= \frac{\mathbb{P}[X_1 = 0 \quad \text{かつ} \quad \sum_{i=2}^n X_i = s]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
&= \frac{\mathbb{P}[X_1 = 0]\mathbb{P}[\sum_{i=2}^n X_i = s]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
&= \frac{(1-\theta)\begin{pmatrix} n-1 \\ s \end{pmatrix}\theta^s(1-\theta)^{n-1-s}}{\begin{pmatrix} n \\ s \end{pmatrix}\theta^s(1-\theta)^{n-s}} = \frac{n-s}{n}
\end{aligned}$$

となる．また，

$$\begin{aligned}
\mathbb{P}[X_1 = 1|\sum_{i=1}^n X_i = s] &= \frac{\mathbb{P}[X_1 = 1 \quad \text{かつ} \quad \sum_{i=1}^n X_i = s]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
&= \frac{\mathbb{P}[X_1 = 1 \quad \text{かつ} \quad \sum_{i=2}^n X_i = s-1]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
&= \frac{\mathbb{P}[X_1 = 1]\mathbb{P}[\sum_{i=2}^n X_i = s-1]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
&= \frac{\theta\begin{pmatrix} n-1 \\ s-1 \end{pmatrix}\theta^s(1-\theta)^{n-1-(s-1)}}{\begin{pmatrix} n \\ s \end{pmatrix}\theta^s(1-\theta)^{n-s}} = \frac{n}{n}
\end{aligned}$$

となる．したがって

$$\tilde{T} = \mathbb{E}_[X_1|\,S] = 0 \cdot \mathbb{P}[X_1 = 0|\,S] + 1 \cdot \mathbb{P}[X_1 = 1|\,S] = \frac{S}{n}$$

定義 6.9 (完備性) $X_1, X_2, \ldots, X_n$ を確率 (密度) 関数 $f(x|\,\theta), \theta \in \Theta$ からの大きさ n のランダム標本とし，$T = T(X_1, X_2, \ldots, X_n)$ を統計量とする．統計量 T の標本分布の確率 (密度) 関数の族が完備であるとは，次の条件をみたすことである．統計量 T の値域上で定義された可測関数 g が，すべての $\theta \in \Theta$ に対し，$\mathbb{E}_\theta[g(T)] = 0$ をみたすならば，$\mathbb{P}_\theta(g(T) = 0) = 1$ が成立する．

さらに，統計量 T が完備であるとは，統計量 T の標本分布の確率 (密度) 関数の族が完備であることをいう．

例 6.13 $X_1, X_2, \ldots, X_n$ をベルヌーイ試行 $f(x|\theta) = \theta^x(1-\theta)^{1-x}I_{\{0,1\}}(x), \theta(0, 1)$ からの大きさ $n\,(n \geq 2)$ のランダム標本とする．たとえば，統計量 $\tilde{T} = X_1 - X_2$ は完備ではない．なぜならば，$\mathbb{E}_\theta[X_1 - X_2] = 0$ であるが，$\mathbb{P}(X_1 - X_2 = 0) \neq 1$ であることよりわかる．

次に，$T = \sum_{i=1}^n X_i$ を考える．$g(T)$ をすべての $\theta \in (0, 1)$ に対し $\mathbb{E}_\theta[g(T)] = 0$ をみたす可測関数とする．T が完備であることを示すために，$t = 0, 1, \ldots, n$ に対して $g(t) = 0$ を示せばよい．いま，$0 < \theta < 1$ に対し

$$0 = \mathbb{E}_\theta[g(T)] = \sum_{t=0}^n \begin{pmatrix} n \\ t \end{pmatrix} \theta^t(1-\theta)^{n-t} = (1-\theta)^t \sum_{t=0}^n g(t) \begin{pmatrix} n \\ t \end{pmatrix} \left(\frac{\theta}{1-\theta}\right)$$

に注意すれば，すべての $0 < \theta < 1$ に対し，

$$\sum_{t=0}^n g(t) \begin{pmatrix} n \\ t \end{pmatrix} \left(\frac{\theta}{1-\theta}\right)^t = 0$$

となることがわかる．または

$$\sum_{t=0}^n g(t) \begin{pmatrix} n \\ t \end{pmatrix} \alpha^t = 0, \qquad \alpha = \frac{\theta}{1-\theta}$$

となる．$\{\alpha^t, t = 0, 1 \ldots, n\}$ は一次独立なので，

$$g(t) \begin{pmatrix} n \\ t \end{pmatrix} = 0, \qquad t = 0, 1 \ldots, n$$

となる．$\begin{pmatrix} n \\ t \end{pmatrix} \neq 0$ なので，$g(t) = 0\,(t = 0, 1 \ldots, n)$ となる．

例 6.14 $X_1, X_2, \ldots, X_n$ を一様分布 $f(x) = I_{(0,1)}(x)$ からの大きさ n のランダム標本とする．ただし，$\Theta = \{\theta, \theta > 0\}$ である．$Y_n = \max(X_1, X_2, \ldots, X_n)$ が完備であることを示そう．そのために，すべての $\theta \in \Theta$ に対し，$\mathbb{E}_\theta[g(Y_n)] = 0$ ならば，$\mathbb{P}_\theta\{g(Y_n) = 0\} = 1$ を示せばよい．

$$\mathbb{E}_\theta[g(Y_n)] = \int g(y) f_{Y_n}(y)\, dy = \int g(y)\theta^{-n} n y^{n-1} I_{(0,\theta)}(y)\, dy$$

となり，すべての $\theta \in \Theta$ に対し

$$\mathbb{E}_\theta[g(Y_n)] = 0 \Longleftrightarrow \frac{n}{\theta^n}\int_0^\theta g(y) y^{n-1}\, dy = 0$$

となる．したがって，θ に関して微分すれば

$$g(y)\theta^{n-1} = 0 \Longleftrightarrow g(y) = 0$$

となることがわかる．

定理 6.5 Lehmann-Sheffée $X_1, X_2, \ldots, X_n$ を $f(x|\theta)$ からの大きさ n のランダム標本とし，$\tau : \Theta \to \mathbb{R}$ とする．$S = S(X_1, X_2, \ldots, X_n)$ が完備十分統計量で $T^* = T^*(S)$ が $\tau(\theta)$ の不偏推定量ならば，T^* は $\tau(\theta)$ の UMVUE である．

証明 $T' = T'(S)$ を $\tau(\theta)$ の不偏推定量とする．すると任意の $\theta \in \Theta$ に対し

$$\mathbb{E}_\theta[T' - T^*] = 0$$

となる．S の完備性から

$$\mathbb{P}_theta[T'(S) = T^*(S)] = 0$$

となる．S の関数である不偏推定量は唯一である．

次に，T を任意の $\tau(\theta)$ の推定量とする．Rao-Blackwell の定理から $\mathbb{E}[T|\,S]$ は $\tau(\theta)$ の不偏推定量で S の関数となるので

$$T^* = \mathbb{E}[T|\,S], \quad a.s.$$

となる．さらに，Rao-Blackwell の定理から

$$\mathbb{VAR}_\theta[T^*] \leq \mathbb{VAR}_\theta[T]$$

となる．よって，T^* は UMVUE である． □

例 6.15 $X_1, X_2, \ldots, X_n$ をポアソン分布 $f(x|\,\theta) = e^{-\theta}\theta^x/x!\,(x = 0, 1, \ldots)$ からの大きさ $n\,(n \geq 2)$ のランダム標本とする．$\sum_{i=1}^n X_i$ は母数 $n\theta$ のポアソン分布に従う．いま，$T = \sum_{i=1}^n X_i$ とする．

$$\begin{aligned}
\mathbb{E}_\theta[g(T)] &= \sum_{t=0}^\infty g(t)\frac{e^{-n\theta}(n\theta)^t}{t!} = e^{-n\theta}\sum_{t=0}^\infty g(t)\frac{n^t}{t!}\theta^t = 0 \\
&\Leftrightarrow \sum_{t=0}^\infty g(t)\frac{n^t}{t!}\theta^t = 0 \qquad \theta \text{ の多項式は一次独立} \\
&\Leftrightarrow g(t) = 0
\end{aligned}$$

となる．したがって，$\sum_{i=1}^n X_i$ は完備統計量である．

$\tau(\theta) = e^{-\theta} = \mathbb{P}_\theta\{X_1 = 0\}$ の推定を考えよう．$I_{\{0\}}(X_1)$ は $e^{-\theta}$ の不偏推定量なので，Lehmann-Sheffée の定理から $\mathrm{T}^* = \mathbb{E}[I_{\{0\}}(X_1)|\sum_{i=1}^n X_i]$ は $\tau(\theta)$ の UMVUE となる．T^* がどのような推定量になるかを計算しよう．

$$\begin{aligned}
\mathbb{P}\{X_1 = 0\sum_{i=1}^n X_i = s\} &= \frac{\mathbb{P}\{X_1 = 0 \quad \text{かつ} \quad \sum_{i=1}^n X_i = s\}}{\mathbb{P}\{\sum_{i=1}^n X_i = s\}} \\
&= \frac{\mathbb{P}\{X_1 = 0 \quad \text{かつ} \quad \sum_{i=2}^n X_i = s\}}{\mathbb{P}\{\sum_{i=1}^n X_i = s\}} \\
&= \frac{\mathbb{P}\{X_1 = 0\}\mathbb{P}\{\sum_{i=2}^n X_i = s\}}{\mathbb{P}\{\sum_{i=1}^n X_i = s\}} \\
&= \frac{e^{-\theta}e^{-(n-1)\theta}[(n-1)\theta]^s/s!}{e^{-n\theta}[n\theta]^s/s!} = \left(\frac{n-1}{n}\right)^s
\end{aligned}$$

となる．したがって

$$\mathbb{E}[I_{\{0\}}(X_1)|\sum_{i=1}^n X_i = s]\mathbb{P}\{X_1 = 0|\sum_{i=1}^n X_i = s\} = \left(\frac{n-1}{n}\right)^s$$

となる．故に

$$T^* = \left(\frac{n-1}{n}\right)^{\sum_{i=1}^n X_i}$$

は $\tau(\theta) = e^{-\theta}$ の UMVUE である．