確率・統計 B
演習問題の解説

若木宏文

wakaki@hiroshima-u.ac.jp
http://home.hiroshima-u.ac.jp/ wakaki/lecture/probstatB17/index.shtml

2018.2.2

table of contents

一様最小分散不偏推定量 (1/11)

定理 1 (クラメール–ラオの不等式 (1 次元パラメーターの場合))

$\boldsymbol{X} = (X_1, \ldots, X_n) \sim \mathsf{P} \in \{\mathsf{P}_\theta \,|\, \theta \in \Theta\}, \quad \Theta \subset \mathbb{R}$
$f(\boldsymbol{x}; \theta)$: P_θ の確率密度関数 (確率関数)

(仮定) 各 $\theta_0 \in \Theta$ に対して θ_0 の近傍 U と 関数 $M(\boldsymbol{x})$ が存在して, $\theta \in U$ ならば

$$\left|\frac{\partial f(\boldsymbol{x}; \theta)}{\partial \theta}\right| \leq M(\boldsymbol{x}), \text{ かつ } \mathsf{E}_\theta\left[\left\{\frac{M(\boldsymbol{X})}{f(\boldsymbol{X}; \theta)}\right\}^2\right] < \infty$$

θ の任意の不偏推定量 $\hat{\theta}$ に対して

$$\mathsf{Var}(\hat{\theta}) \geq \frac{1}{J(\theta)}$$

一様最小分散不偏推定量 (2/11)

定理 1 (クラメール－ラオの不等式 (つづき))

ただし,

$$
\mathrm{E}_\theta[g(\boldsymbol{X})] = \begin{cases} \displaystyle\iint g(\boldsymbol{x}) f(\boldsymbol{x};\theta)\,dx_1 \cdots dx_n & (\text{連続型}) \\ \displaystyle\sum_{j=1}^{\infty} g(\boldsymbol{x}_j) f(\boldsymbol{x}_j;\theta) & (\text{離散型}) \end{cases},
$$

$$
J(\theta) = \mathrm{E}_\theta\left[\left\{\frac{\partial}{\partial\theta}\log f(\boldsymbol{X};\theta)\right\}^2\right]
$$

であり, $J(\theta) > 0$ と仮定する.

- 定理 1 の $J(\theta)$ をフィッシャー情報量と呼ぶ.
- $\mathrm{Var}(\hat{\theta}) = \dfrac{1}{J(\theta)}$ となる不偏推定量が存在するとき, それを有効推定量と呼ぶ

単純仮説の場合 (2/3)

最尤推定量に基づく検定

$\hat{\boldsymbol{\theta}}_n$: Θ での $\boldsymbol{\theta}$ の最尤推定量

$$\hat{\boldsymbol{\theta}}_n = \mathrm{argmax}_{\boldsymbol{\theta}\in\Theta} \prod_{i=1}^{n} f(X_i, \boldsymbol{\theta})$$

$J(\boldsymbol{\theta})$: フィッシャー情報量行列

$$J(\boldsymbol{\theta}) = \mathsf{E}\left[\left\{\frac{\partial}{\partial\boldsymbol{\theta}}\log f(X_1, \boldsymbol{\theta})\right\}\left\{\frac{\partial}{\partial\boldsymbol{\theta}}\log f(X_1, \boldsymbol{\theta})\right\}'\right]$$

$$T_W \equiv n(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)' J(\hat{\boldsymbol{\theta}}_n)(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) > \chi^2_p(\alpha) \Rightarrow \text{棄却}$$

は適当な正則条件の下で, 有意水準が近似的に α となる.
これをワルド検定と呼ぶ. (テキストの補遺 C 参照)

単純仮説の場合 (3/3)

スコア関数に基づく検定

$$\boldsymbol{s}_n(\boldsymbol{\theta}) = \sum_{i=1}^{n} \frac{\partial \log f(X_i, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$$

をスコア関数と呼ぶ.

$$T_S \equiv \frac{1}{n}\{\boldsymbol{s}_n(\boldsymbol{\theta}_0)\}' J(\boldsymbol{\theta}_0)^{-1}\{\boldsymbol{s}_n(\boldsymbol{\theta}_0)\} > \chi_p^2(\alpha) \Rightarrow \text{棄却}$$

は適当な正則条件の下で, 有意水準が近似的に α となる.
これをスコア検定と呼ぶ. (テキストの補遺 C 参照)

メモ用紙

問題 1 (1)

$$\mathsf{P}(X_{(i)} \le x) = \mathsf{P}(X_1, \ldots, X_n \text{ の内 } i \text{ 個以上が } x \text{ 以下 })$$
$$= \sum_{k=i}^{n} \frac{n!}{k!(n-k)!} \{F(x)\}^k \{1 - F(x)\}^{n-k}$$

$$\frac{d}{dx}\mathsf{P}(X_{(i)} \le x) = \sum_{k=i}^{n} \frac{n!}{(k-1)!(n-k)!} \{F(x)\}^{k-1} \{1 - F(x)\}^{n-k} f(x)$$
$$- \sum_{k=i}^{n-1} \frac{n!}{k!(n-k-1)!} \{F(x)\}^{k} \{1 - F(x)\}^{n-(k+1)} f(x)$$
$$= \sum_{k=i}^{n} \frac{n!}{(k-1)!(n-k)!} \{F(x)\}^{k-1} \{1 - F(x)\}^{n-k} f(x)$$
$$- \sum_{k'=i+1}^{n} \frac{n!}{(k'-1)!(n-k')!} \{F(x)\}^{k'-1} \{1 - F(x)\}^{n-k'} f(x)$$

問題 1 (1), 続き

$$= \frac{n!}{(i-1)!(n-i)!}\{F(x)\}^{i-1}\{1-F(x)\}^{n-i}f(x)$$

指数分布の分布関数は

$$F(x) = \int_0^x \frac{1}{\mu}e^{-t/\mu}dt = 1 - e^{-x/\mu}$$

$n = 2m-1, i = m$ として $X_{(m)}$ の確率密度関数は

$$\frac{(2m-1)!}{(m-1)!(m-1)!}\{1-e^{-x/\mu}\}^{m-1}\{e^{-x/\mu}\}^{m-1}\frac{e^{-x/\mu}}{\mu}$$

問題 1 (2)

$$\mathsf{P}(X_{(m)} \le x) = \frac{\Gamma(2m)}{\Gamma(m)\Gamma(m)} \int_{-\infty}^{x} \{1 - e^{-t/\mu}\}^{m-1} \{e^{-t/\mu}\}^{m-1} \frac{e^{-t/\mu}}{\mu} dt$$

$$\left(u = 1 - e^{-t/\mu} \text{ と変数変換 } \right)$$

$$= \frac{\Gamma(2m)}{\Gamma(m)\Gamma(m)} \int_{0}^{1-e^{-x/\mu}} u^{m-1}(1-u)^{m-1} du$$

$$= \mathsf{P}(U \le 1 - e^{-x/\mu})$$

問題 1 (3)

$Z_1, \ldots, Z_{2m} \overset{i.i.d.}{\sim} N(0,1)$ とすると大数の法則から $m \to \infty$ のとき

$$\frac{1}{2m}\sum_{i=1}^{2m} Z_i^2 \xrightarrow{p} \mathsf{E}(Z_1^2) = 1 \quad \therefore \quad \frac{1}{2m}\sum_{i=1}^{2m} Z_i^2 \xrightarrow{d} 1$$

分布収束は, 分布関数で定義されるので

$\sum_{i=1}^{2m} Z_i^2 \sim \chi^2_{2m}$ であることから $\quad \frac{1}{2m} Y_1 \xrightarrow{d} 1$

スラスキーの定理から

$$\frac{Y_1}{Y_1 + Y_2} = \frac{\frac{Y_1}{2m}}{\frac{Y_1}{2m} + \frac{Y_2}{2m}} \xrightarrow{p} \frac{1}{1+1} = \frac{1}{2}$$

したがって

$$\mathsf{P}(U \le u) = \mathsf{P}\left(\frac{Y_1}{Y_2 + Y_3} \le u\right) \to \begin{cases} 1 & u > 1/2 \\ 0 & u < 1/2 \end{cases}$$

問題 1 (3), 続き

$x > x_0$ のとき, $1 - e^{-x/\mu} > \frac{1}{2}$ であるから $\mathsf{P}(U \leq 1 - e^{-x/\mu}) \to 1$

$x < x_0$ のとき, $1 - e^{-x/\mu} < \frac{1}{2}$ であるから $\mathsf{P}(U \leq 1 - e^{-x/\mu}) \to 0$

したがって,

$$\mathsf{P}(X_{(m)} \leq x) = \mathsf{P}(U \leq 1 - e^{-x/\mu}) \to \begin{cases} 1 & x > x_0 \\ 0 & x < x_0 \end{cases}$$

問題 2 (1), (2)

$$\mathsf{E}\left[\sum_{i=1}^{n} c_i X_i\right] = \sum_{i=1}^{n} c_i \mathsf{E}[X_i] = \sum_{i=1}^{n} c_i \lambda = \lambda \quad \text{より} \quad \sum_{i=1}^{n} c_i = 1$$

$$\mathsf{Var}\left[\sum_{i=1}^{n} c_i X_i\right] = \sum_{i=1}^{n} \mathsf{Var}[c_i X_i] = \sum_{i=1}^{n} c_i^2 \mathsf{Var}[X_i] = \sum_{i=1}^{n} c_i^2 \lambda = \|\boldsymbol{c}\|^2 \lambda$$

$$\because \mathsf{E}[X_i^2] = \mathsf{E}[X_i(X_i - 1)] + \mathsf{E}[X_i]$$

$$\mathsf{E}[X_i(X_i - 1)] = e^{-\lambda} \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=2}^{\infty} \lambda^2 \frac{\lambda^{k-2}}{(k-2)!}$$

$$= \lambda^2 e^{-\lambda} \sum_{k'=0}^{\infty} \frac{\lambda^{k'}}{k'!} = \lambda^2$$

$$\mathsf{Var}[X_i] = \mathsf{E}[X_i^2] - \mathsf{E}[X_i]^2 = \lambda^2 + \lambda - \lambda^2$$

問題 2 (3)

$\mathbf{1} = (1, \ldots, 1)$ とおくとシュワルツの不等式より

$$\|\boldsymbol{c}\|^2 \|\mathbf{1}\|^2 \geq (\boldsymbol{c}, \mathbf{1})^2 = \left\{ \sum_{i=1}^{n} c_i \right\}^2 = 1$$

$\therefore \|c\|^2 \geq \dfrac{1}{n}$ (等号が成り立つのは)$\boldsymbol{c} = k\mathbf{1}$.

このとき $\displaystyle\sum_{i=1}^{n} c_i = nk$ より $k = \dfrac{1}{n}$ $\therefore \boldsymbol{c} = \dfrac{1}{n}(1, \ldots, 1)$

問題 2 (4)

$t(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i$ とおく.

$X_1, \ldots, X_n$ の同時確率密度関数は

$$\begin{aligned}
\mathsf{P}(X_1 = x_1, \ldots, X_n = x_n) &= \prod_{i=1}^{n} e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} \\
&= e^{-n\lambda} \frac{1}{x_1! \cdots x_n!} \lambda^{\sum_{i=1}^{n} x_i} \\
&= \frac{1}{x_1! \cdots x_n!} \times e^{-n\lambda} \lambda^{t(x_1, \ldots, x_n)}
\end{aligned}$$

となるから, 分解定理より $T = t(X_1, \ldots, X_n)$ は十分統計量である.

問題 2 (4), 続き

完備性を示すために，まず，特性関数を利用して $T \sim p(n\lambda)$ であることを示す．$X \sim p(\lambda)$ の特性関数は

$$\varphi_X(t;\lambda) = \mathsf{E}[e^{itX}] = e^{-\lambda}\sum_{x=0}^{\infty} e^{itx}\frac{\lambda^x}{x!}$$

$$= e^{-\lambda}\sum_{x=0}^{\infty}\frac{\left(e^{it}\lambda\right)^x}{x!} = e^{-\lambda}\exp\{e^{it}\lambda\}$$

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} p(\lambda)$ より

$$\varphi_T(t) = \mathsf{E}[e^{it\sum_{k=1}^{n} X_i}] = \mathsf{E}\left[\prod_{k=1}^{n} e^{itX_k}\right] = \{\varphi_X(t;\lambda)\}^n$$

$$= e^{-n\lambda}\exp\{e^{it}n\lambda\}$$ これは $p(n\lambda)$ の特性関数だから

$T \sim p(n\lambda)$

問題 2 (4), 続き

任意の $\lambda > 0$ に対して $\mathsf{E}[g(T)] = 0$ ならば
任意の非負の整数 k に対して $g(k) = 0$ であることを示せばよい.

$\mathsf{E}[g(T)] = e^{-n\lambda}\sum_{k=0}^{\infty} g(k)\frac{(n\lambda)^k}{k!} = 0$ より $G(\lambda) := \sum_{k=0}^{\infty}\left\{g(k)\frac{n^k}{k!}\right\}\lambda^k = 0$

$\mathsf{E}[g(T)]$ が存在することから, 級数は絶対収束する.

優収束定理から $0 = \lim_{\lambda\to 0+0} G(\lambda) = \sum_{k=0}^{\infty}\left\{g(k)\frac{n^k}{k!}\right\}\lim_{\lambda\to 0+0}\lambda^k = g(0)$

$g(0) = g(1) = \cdots g(k-1) = 0$ と仮定すると

$$G(\lambda) = \sum_{j=k}^{\infty}\left\{g(j)\frac{n^j}{j!}\right\}\lambda^j = \lambda^k\sum_{j=0}^{\infty}\left\{g(j+k)\frac{n^{j+k}}{(j+k)!}\right\}\lambda^j$$

より $\sum_{j=0}^{\infty}\left\{g(j+k)\frac{n^{j+k}}{(j+k)!}\right\}\lambda^j = 0,\ \lambda \to 0+0$ とすると $g(k) = 0$

問題 2 (5), (6)

$X_1, \ldots, X_n$ は独立に同じ分布に従うので
$T = t$ を与えたときの X_j の条件付き平均は j に依らない.
従って

$$\begin{aligned} \mathsf{E}[X_1 \mid T = t] &= \frac{1}{n}\mathsf{E}[X_1 + \cdots + X_n \mid T = t] \\ &= \frac{1}{n}\mathsf{E}[T \mid T = t] = \frac{t}{n} \end{aligned}$$

$\bar{X}_n$ は完備十分統計量の関数で, かつ, 不偏推定量であるから一様最小分散不偏推定量である.

問題 2 (7)

対数尤度関数

$$\begin{aligned}\ell(\lambda; x_1, \ldots, x_n) &= \log f(x_1, \ldots, x_n; \lambda) \\ &= -n\lambda + t(x_1, \ldots, x_n) \log \lambda - \log(x_1! \cdots x_n!)\end{aligned}$$

スコア関数

$$s_n(\lambda) = -n + \frac{t(x_1, \ldots, x_n)}{\lambda}$$

フィッシャー情報量 ($X_1, \ldots, X_n$ の同時分布のもの)

$$J(\lambda) = \mathrm{Var}[s_n(\lambda)] = \mathrm{Var}\left[\frac{T}{\lambda}\right] = \frac{1}{\lambda^2}(n\lambda) = \frac{n}{\lambda}$$

クラメール・ラオの不等式

$$\mathrm{Var}[\lambda \text{ の不偏推定量 }] \geq \frac{1}{J(\lambda)} = \frac{\lambda}{n}$$

$\mathrm{Var}[\bar{X}_n] = \dfrac{\lambda}{n}$ であるから有効推定量である.

問題 2 (8),(9),(10)

大数の法則より $\bar{X}_n \xrightarrow{p} \mathsf{E}[X_j] = \lambda$

$$s_n(\lambda) = -n + \frac{T}{\lambda} = 0 \text{ より, 最尤推定量は } \hat{\lambda} = \frac{T}{n} = \bar{X}_n$$

$$Z_n = \frac{\bar{X}_n - \mathsf{E}[\bar{X}_n]}{\sqrt{\mathsf{Var}[\bar{X}_n]}} \cdot \frac{\sqrt{\lambda}}{\sqrt{\bar{X}_n}}$$

中心極限定理とスラスキーの定理より, $N(0,1)$ に分布収束する.

$-1.96 \leq Z_n \leq 1.96$ を λ について解くと

$$\bar{X}_n - 1.96\sqrt{\frac{\bar{X}_n}{n}} \leq \lambda \leq \bar{X}_n + 1.96\sqrt{\frac{\bar{X}_n}{n}}$$

問題3 (1)

$X_1, \ldots, X_n$ の同時確率密度関数：

$$f(x_1, \ldots, x_n; \lambda) = e^{-n\lambda} \frac{1}{x_1! \cdots x_n!} \lambda^{\sum_{i=1}^n x_i}$$

最強力検定の棄却域は

$$\mathscr{W}_c = \left\{ (x_1, \ldots, x_n) \,\middle|\, \frac{f(x_1, \ldots, x_n; 2)}{f(x_1, \ldots, x_n; 1)} \geq c \right\}.$$

$$\frac{f(x_1, \ldots, x_n; 2)}{f(x_1, \ldots, x_n; 1)} = e^{-n}\, 2^{\sum_{i=1}^n x_i} \geq c \Leftrightarrow \sum_{i=1}^n x_i \geq \frac{n + \log c}{\log 2}$$

帰無仮説の下で $\displaystyle\sum_{i=1}^n X_i \sim p(n)$ なので

$\dfrac{n + \log c}{\log 2} = k$ となるように c をとれば

有意水準 α_k の最強力検定となる.

問題3 (2)

$\lambda' > 1$ として，

帰無仮説 $H_0 : \lambda = 1$，　対立仮説 $H_1 : \lambda = \lambda'$

の最強力検定を求める

最強力検定の棄却域は

$$\mathscr{W}_c = \left\{ (x_1, \ldots, x_n) \,\middle|\, \frac{f(x_1, \ldots, x_n; \lambda')}{f(x_1, \ldots, x_n; 1)} \geq c \right\}.$$

$$\frac{f(x_1, \ldots, x_n; 2)}{f(x_1, \ldots, x_n; 1)} = e^{-n(\lambda'-1)} \lambda'^{\sum_{i=1}^{n} x_i} \geq c$$

$$\Leftrightarrow \sum_{i=1}^{n} x_i \geq \frac{n(\lambda' - 1) + \log c}{\log \lambda'}$$

有意水準を α_k とするには

$\dfrac{n(\lambda') + \log c}{\log \lambda'} = k$ となるように c をとることになるので

棄却域は λ' によらない．したがって一様最強力検定である

問題 3 (3),(4)

対立仮説の下での λ の尤度比検定統計量は

$\hat{\lambda} = \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_n$ であるから，尤度比検定統計量は

$$\frac{f(X_1, \ldots, X_n; 1)}{f(X_1, \ldots, X_n; \bar{X}_n)} = e^{n(\bar{X}_n - 1)} \bar{X}_n^{-n\bar{X}_n}$$

フィッシャー情報量 (X_1 1 個分) は $J(\lambda) = \frac{1}{\lambda}$，スコア関数は

$s_n(\lambda) = \frac{\sum_{i=1}^{n} X_i}{\lambda} - n$ なので

スコア検定統計量：$T_S = \frac{1}{n} s_n(1) J(1)^{-1} s_n(1) = n(\bar{X}_n - 1)^2$

ワルド検定統計量：$T_W = n(\bar{X}_n - 1) J(\bar{X}_n)(\bar{X}_n - 1) = \frac{n(\bar{X}_n - 1)^2}{\bar{X}_n}$