今野　(2011 年 12 月 13 日)

情報統計学の問題 (その４) の解答例

$\boxed{\text{問題 1}}$　$X_1, X_2, \ldots, X_n$ は確率密度関数

$$f(x|\,\theta) = \theta x^{\theta-1}, \qquad 0 < x < 1, \qquad 0 < \theta < \infty \qquad (*)$$

からの大きさ n のランダム標本とする．

(1) $X_1 = x_1, \ldots, X_n = x_n$ を観測したのときの尤度関数を述べよ．

(2) θ の最尤推定量を求めよ．さらに，$\tau = \theta^{-1}$ の最尤推定量は

$$\widehat{\tau}_n = -\frac{1}{n}\sum_{i=1}^{n} \log X_i$$

で与えられることを示せ．

(3) 確率変数 X が確率密度関数 $(*)$ を持つ分布に従うとき，X と $T = -\theta \log X$ の分布関数を求めよ．さらに，T の積率母関数 $M_T(r)$ を求めよ．

(4) $\widehat{\tau}_n$ は τ の不偏推定量であることを示した上でその分散を求めよ．

(5) $\widehat{\tau}_n$ は τ の一様最小分散不偏推定量であることを示せ．

解答. (1)

$$L_n(\theta|\,x_1, \ldots, x_n) = \Pi_{i=1}^{n} f(x_i|\,\theta) = \Pi_{i=1}^{n} \theta x_i^{\theta-1} = \theta^n (x_1 x_2 \times \cdots \times x_n)^{\theta-1}$$

(2) 対数尤度関数

$$\ell_n(\theta|\,x_1, x_2, \ldots, x_n) = \log L_n(\theta|\,x_1, \ldots, x_n) = n \log\theta + (\theta - 1)\sum_{i=1}^{n} \log x_i$$

となる．θ に関して微分すれば、

$$\frac{d}{d\theta}\ell_n(\theta|\,x_1, x_2, \ldots, x_n) = \frac{n}{\theta} + \sum_{i=1}^{n} \log x_i = 0$$

とおけば，

$$\theta = -\frac{n}{\sum_{i=1}^{n} \log x_i}$$

となる．この点で $\ell_n(\theta|\,x_1, x_2, \ldots, x_n)$ は最大となるので，θ の最尤推定値は

$$\hat{\theta}_n = -\frac{n}{\sum_{i=1}^{n} \log x_i}$$

したがって，$\tau = 1/\theta$ の最尤推定値は

$$\hat{\tau}_n = -\frac{1}{n}\sum_{i=1}^{n}\log x_i$$

となる．よって，$\tau = 1/\theta$ の最尤推定値は

$$\hat{\tau}_n = -\frac{1}{n}\sum_{i=1}^{n}\log X_i$$

となる．
(3) $0 < x < 1$ のとき，

$$F_X(x|\,\theta) = \int_{-\infty}^{x} f(t|\,\theta)\,dt = \int_0^x f(t|\,\theta)\,dt = x^\theta$$

となるので，

$$F_X(x|\,\theta) = \begin{cases} 1 & (x \geq 1), \\ x^\theta & (0 < x < 1), \\ 0 & (x \leq 0) \end{cases}$$

つぎに，T の分布関数を求める．$0 < x < 1$ のとき，$-\theta \log x > 0$ となることに注意する．$t > 0$ に対して，$0 < \exp(-t/\theta) < 1$ なので，

$$\begin{aligned} F_T(t) &= \mathbb{P}(T \leq t) = \mathbb{P}(-\theta\log X \leq t) \\ &= \mathbb{P}\big(X \geq \exp(-t/\theta)\big) \\ &= 1 - \mathbb{P}\big(X \leq \exp(-t/\theta)\big) \\ &= 1 - F_X(\exp(-t/\theta)) \\ &= 1 - (\exp(-t/\theta))^\theta = 1 - e^{-t} \end{aligned}$$

したがって，

$$F_T(t) = \begin{cases} 1 - e^{-t} & (t > 0), \\ 0 & (t \leq 0) \end{cases}$$

したがって，T の確率密度関数は

$$f_T(t) = \begin{cases} e^{-t} & (t > 0), \\ 0 & (t \leq 0) \end{cases}$$

となるので，T の積率母関数は $\xi < 1$ に対して，

$$M_T(\xi) = \mathbb{E}[\exp(\xi T)] = \int_0^\infty e^{\xi t}e^{-t}\,dt = \left[-\frac{1}{1-\xi}e^{-(1-\xi)t}\right]_0^\infty = \frac{1}{1-\xi}$$

このことから

$$\begin{aligned} \mathbb{E}[T] &= \left.\frac{d}{d\xi}\right|_{\xi=0} M_T(\xi) = \left.\frac{1}{(1-\xi)^2}\right|_{\xi=0} = 1 \\ \mathbb{E}[T^2] &= \left.\frac{d^2}{d\xi^2}\right|_{\xi=0} M_T(\xi) = \left.\frac{2}{(1-\xi)^3}\right|_{\xi=0} = 2 \end{aligned}$$

したがって，
$$\mathbb{VAR}[T] = \mathbb{E}[T^2] - \{\mathbb{E}[T]\}^2 = 1$$
に注意する．
(4) $T_i = -\theta \log X_i$ とおけば，T_i の確率密度関数は
$$f_T(t) = \begin{cases} e^{-t} & (t > 0), \\ 0 & (t \leq 0) \end{cases}$$
となることに注意すれば，
$$\begin{aligned} \mathbb{E}[\hat{\tau}_n] &= \frac{\tau}{n}\sum_{i=1}^{n} \mathbb{E}\left[-\theta \log X_i\right] = \frac{\tau}{n}\sum_{i=1}^{n} \mathbb{E}\left[T_i\right] = \tau \\ \mathbb{VAR}[\hat{\tau}_n] &= \frac{\tau^2}{n^2}\sum_{i=1}^{n} \mathbb{VAR}\left[-\theta \log X_i\right] = \frac{\tau^2}{n^2}\sum_{i=1}^{n} \mathbb{VAR}\left[T_i\right] = \frac{\tau^2}{n} \end{aligned}$$
(5) Cramér-Rao の不等式から，τ に不偏推定量 S に対して，
$$\mathbb{VAR}_\theta[S] \geq \frac{(\tau'(\theta))^2}{n\mathbb{E}_\theta\left[\left(\frac{d}{d\theta}\log f(X|\,\theta)\right)^2\right]}$$
であった．
$$\frac{d}{d\theta}\log f(X|\,\theta) = \frac{d}{d\theta}\left\{\log\theta + (\theta - 1)\log X\right\} = \frac{1}{\theta} + \log X = \frac{1}{\theta}(1 + \theta \log X) = \tau(1 - T)$$
となるので，
$$\mathbb{E}_\theta\left[\left(\frac{d}{d\theta}\log f(X|\,\theta)\right)^2\right] = \tau^2\mathbb{E}[(T-1)^2] = \tau^2\mathbb{VAR}[T] = \tau^2$$
となる．さらに，
$$\tau'(\theta) = -\frac{1}{\theta^2} = -\tau^2$$
なので，
$$\mathbb{VAR}_\theta[S] \geq \frac{(\tau'(\theta))^2}{n\mathbb{E}_\theta\left[\left(\frac{d}{d\theta}\log f(X|\,\theta)\right)^2\right]} = \frac{\tau^4}{n\tau^2} = \frac{\tau^2}{n}$$
したがって，τ の最尤推定量 $\hat{\tau}_n = -\frac{1}{n}\sum_{i=1}^{n}\log X_i$ の分散は Cramér-Rao の下限に到達するので，τ の一様最小分散不偏推定量になることがわかる．