確率・統計 B 練習問題 2

2017.2.3

問 1 $X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$, $\hat{\mu} = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} X_j$ とする.

(1) $\hat{\mu}$ の偏りを求めよ.

(2) $\hat{\mu}$ の分散を求めよ.

(3) $\hat{\mu}$ の平均 2 乗誤差を求めよ.

(4) $\hat{\mu}$ の集中確率 $\mathrm{P}(|\hat{\mu} - \mu| \leq a)$ を a, n, σ と 標準正規分布関数 Φ を用いて表わせ.

(5) $\hat{\mu}$ は一致推定量であることを示せ.

問 2 X_1, X_2 を独立に $E(1/\mu)$ に従う確率変数とする.

(1) $\mathrm{E}(X_1), \mathrm{E}(X_1^2), \mathrm{Var}(X_1)$ を求めよ.

(2) (X_1, X_2) の同時確率密度関数 $f(x_1, x_2;\ \mu)$, および μ の Fisher 情報量を求めよ.

(3) $T = X_1 + X_2$ は μ の十分統計量である. 理由を述べよ.

(4) (X_1, T) の同時確率密度関数を求めよ.

(5) $T = t$ を与えたときの X_1 の条件付き確率密度関数を求めよ.

(6) $g(T) = \mathrm{E}[X_1 \,|\, T]$ を求めよ.

(7) $g(T)$ は μ の一様最小分散不偏推定量であることを示せ.

問 3 $X_1, \ldots, X_n \overset{i.i.d.}{\sim} G(p)$ とする.

(1) $\mathrm{E}(X_1)$ を求めよ.

(2) p の最尤推定量 $\hat{p}_n$ を求めよ.

(3) $\hat{p}_n$ は p の一致推定量であることを示せ.

問 4 (1) 手作りのサイコロを 100 回投げたところ, 偶数の目が 64 回でた. このサイコロを投げるとき偶数がでる確率を p とする. p の信頼係数 95% 近似信頼区間を求めよ.

(2) 中学校 2 年生男子 100 人を無作為に選んで, 身長を測定したところ, その平均は 155.0cm, 標準偏差 10.0 cm であった. 身長の母平均の 95% 信頼区間を求めよ.

z_α を標準正規分布の上側 α 点とするとき, $z_{0.025} = 1.96$ とせよ.

指数分布 : $E(\lambda)$ $\quad f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$

幾何分布 : $G(p)$ $\quad f(x;p) = p(1-p)^x \ (x = 0, 1, 2, \ldots)$

解答

問 1 (1) $\mathrm{E}[\hat{\mu}] - \mu = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n} \mathrm{E}(X_j) - \mu = 0$

(2) $\mathrm{Var}\Big(\dfrac{1}{n}\displaystyle\sum_{j=1}^{n} X_j\Big) = \displaystyle\sum_{j=1}^{n}\Big(\dfrac{1}{n}\Big)^2 \mathrm{Var}(X_j) = \dfrac{1}{n}\sigma^2$

(3) 不偏推定量については, 平均 2 乗誤差と分散は同じだから $\dfrac{\sigma^2}{n}$

(4) $\dfrac{\sqrt{n}(\hat{\mu}-\mu)}{\sigma} \sim N(0,1)$ であるから

$$\mathrm{P}(|\hat{\mu}-\mu| \le a) = \mathrm{P}\Big(\Big|\frac{\sqrt{n}(\hat{\mu}-\mu)}{\sigma}\Big| \le \frac{\sqrt{n}a}{\sigma}\Big) = \Phi\Big(\frac{\sqrt{n}a}{\sigma}\Big) - \Phi\Big(-\frac{\sqrt{n}a}{\sigma}\Big)$$

(5) 任意の $\varepsilon > 0$ に対して

$$\mathrm{P}(|\hat{\mu}-\mu| < \varepsilon) = \mathrm{P}(|\hat{\mu}-\mu| \le \varepsilon) = 1 - \Phi\Big(\frac{\sqrt{n}a}{\sigma}\Big) + \Phi\Big(-\frac{\sqrt{n}a}{\sigma}\Big)$$
$$\lim_{n\to\infty} \mathrm{P}(|\hat{\mu}-\mu| < \varepsilon) = 1 - \lim_{n\to\infty}\Phi\Big(\frac{\sqrt{n}a}{\sigma}\Big) + \lim_{n\to\infty}\Phi\Big(-\frac{\sqrt{n}a}{\sigma}\Big) = 0$$

問 2 (1) $\mathrm{E}(X_1) = \displaystyle\int_0^\infty \frac{x}{\mu}e^{-x/\mu}dx = \mu\int_0^\infty ye^{-y}dy = \mu \quad (x = \mu y,\ dx = \mu dy)$

$\mathrm{E}(X_1^2) = \displaystyle\int_0^\infty \frac{x^2}{\mu}e^{-x/\mu}dx = \mu^2\int_0^\infty y^2e^{-y}dy = 2\mu^2$

$\mathrm{Var}(X_1) = \mathrm{E}(X_1^2) - \{\mathrm{E}(X_1)\}^2 = \mu^2$

(2) $x_1 \ge 0, x_2 \ge 0$ のとき

$f(x_1,x_2;\mu) = \dfrac{1}{\mu^2}e^{-x_1/\mu}e^{-x_2/\mu} = \dfrac{1}{\mu^2}e^{-(x_1+x_2)/\mu}$

$x_1 < 0$ または $x_2 < 0$ のとき $f(x_1,x_2;\mu) = 0$

$\log f(x_1,x_2;\mu) = -\dfrac{x_1+x_2}{\mu} - 2\log\mu$

$\dfrac{\partial \log f(x_1,x_2;\mu)}{\partial\mu} = \dfrac{x_1+x_2}{\mu^2} - \dfrac{2}{\mu}$

$\mathrm{E}\Big[\Big(\dfrac{\partial \log f(X_1,X_2;\mu)}{\partial\mu}\Big)\Big] = \dfrac{\mathrm{E}(X_1)+\mathrm{E}(X_2)}{\mu^2} - \dfrac{2}{\mu} = 0$

$J(\mu) = \mathrm{E}\Big[\Big(\dfrac{\partial \log f(X_1,X_2;\mu)}{\partial\mu}\Big)^2\Big] = \mathrm{Var}\Big[\Big(\dfrac{\partial \log f(X_1,X_2;\mu)}{\partial\mu}\Big)\Big]$

$\quad = \dfrac{1}{\mu^4}(\mathrm{Var}(X_1) + \mathrm{Var}(X_2)) = \dfrac{2}{\mu^2}$

(3) $f(x_1,x_2;\mu)$ $t(x_1,x_2) = x_1 + x_2$ の関数であるから分解定理より, $T = t(X_1,X_2)$ は十分統計量である.

(4) $Y=X_1$ とすると $X_2=T-Y$, $T\geq 0,\ 0\leq Y\leq T$

$$\begin{vmatrix}\dfrac{\partial X_1}{\partial Y_1} & \dfrac{\partial X_2}{\partial Y_1}\\ \dfrac{\partial X_1}{\partial T} & \dfrac{\partial X_2}{\partial T}\end{vmatrix}=\begin{vmatrix}1 & -1\\ 0 & 1\end{vmatrix}=1$$

$$f_{X_1,T}(x_1,t)=f_{Y,T}(x_1,t)=\begin{cases}\dfrac{1}{\mu^2}e^{-t/\mu} & (0\leq x_1\leq t)\\ 0 & \text{その他}\end{cases}$$

(5) T の周辺確率密度関数は $t\geq 0$ のとき

$$f_T(t)=\int_0^t \frac{1}{\mu^2}e^{-t/\mu}dx_1=\frac{t}{\mu^2}e^{-t/\mu}$$

X_1 の $T=t$ が与えたときのの条件付き確率密度関数は $0\leq x_1\leq t$ のとき

$$\frac{f_{X_1,T}(x_1,t)}{f_T(t)}=\frac{1}{t}$$

(6) X_1 の $T=t$ のときの条件付き分布は $[0,t]$ 上の一様分布であるから

$$\mathrm{E}(X_1|T=t)=\frac{t}{2}$$

$$g(T)=\mathrm{E}(X_1|T)=\frac{T}{2}$$

(7) (1) より $\mathrm{E}[g(T)]=\mu$, $\mathrm{Var}[g(T)]=\dfrac{1}{4}\mathrm{Var}X_1+\dfrac{1}{4}\mathrm{Var}X_2=\dfrac{\mu^2}{2}$

(2) とクラメール－ラオの不等式から 任意の不偏推定量の分散の下限は $\dfrac{\mu^2}{2}$.

したがって $g(T)$ は有効推定量であるから, 一様最小分散不偏推定量である.

問 3 (1) $\mathrm{E}(X_1)=p\displaystyle\sum_{k=0}^{\infty}k(1-p)^k$

$$\sum_{k=0}^{\infty}(1-p)^k=\frac{1}{1-(1-p)}=\frac{1}{p}$$

の左辺を項別微分すると

$$-\sum_{k=0}^{\infty}k(1-p)^{k-1}=-\sum_{k=0}^{\infty}(k+1)(1-p)^k$$

右辺の級数は絶対収束するから,

$$-\sum_{k=0}^{\infty}(k+1)(1-p)^k=\frac{d}{dp}\left(\frac{1}{p}\right)=-\frac{1}{p^2}$$

したがって

$$\sum_{k=0}^{\infty}k(1-p)^k=\sum_{k=0}^{\infty}(k+1)(1-p)^k-\sum_{k=0}^{\infty}(1-p)^k=\frac{1}{p^2}-\frac{1}{p}$$

$$\mathrm{E}(X_1)=\frac{1}{p}-1$$

(別解) X_1 の 特性関数は

$$\phi_{X_1}(t) = \mathrm{E}(e^{itX_1}) = p\sum_{k=0}^{\infty} e^{itk}(1-p)^k = p\sum_{k=0}^{\infty}\{e^{it}(1-p)\}^k = p\{1-e^{it}(1-p)\}^{-1},$$

$\phi'_{X_1}(t) = -p\{1-e^{it}(1-p)\}^{-2}\{-ie^{it}(1-p)\}$ より $\mathrm{E}(X_1) = \dfrac{1}{i}\phi'_{X_1}(0) = \dfrac{1-p}{p}$

(2) $\boldsymbol{X} = (X_1,\ldots,X_n)'$ の同時確率密度関数は

$f(\boldsymbol{x};p) = p^n(1-p)^{\sum_{j=1}^n x_j}$

$$\ell(p;\boldsymbol{x}) = \log f(\boldsymbol{x};p) = n\log p + \sum_{j=1}^n x_j \log(1-p)$$

$\dfrac{\partial \ell(p;\boldsymbol{x})}{\partial p} = \dfrac{n}{p} - \displaystyle\sum_{j=1}^n x_j \dfrac{1}{1-p} = 0$ より

$$\hat{p} = \frac{n}{n+\sum_{j=1}^n X_j} = \frac{1}{1+\bar{X}}, \quad \bar{X} = \frac{1}{n}\sum_{j=1}^n X_j$$

(3) 大数の法則より $\bar{X} \xrightarrow{p} \mathrm{E}(X_1) = p^{-1}-1 \ (n\to\infty)$

したがって, $\hat{p} \xrightarrow{p} \dfrac{1}{1+(p^{-1}-1)} = p \ (n\to\infty)$

問 4 (1) 100 回投げたときの偶数の出る回数を X , $\hat{p} = \dfrac{X}{100}$ とする. $X \sim B(100,p)$ であるから $\mathrm{E}(\hat{p}) = p$, $\mathrm{Var}(\hat{p}) = \dfrac{p(1-p)}{100}$. したがって, 中心極限定理と $\hat{p}$ の一致性より

$$\mathrm{P}\left(\left|\frac{10(\hat{p}-p)}{\sqrt{\hat{p}(1-\hat{p})}}\right| \le 1.96\right) \approx 0.95$$

となり, 左辺括弧内の不等式を p について解くと, 95%近似信頼区間は

$$\left[\hat{p} - 1.96\frac{\sqrt{\hat{p}(1-\hat{p})}}{10}, \hat{p} + 1.96\frac{\sqrt{\hat{p}(1-\hat{p})}}{10}\right]$$

$\hat{p} = \dfrac{64}{100}$ を代入すると

$[0.64 - 1.96 \times 0.048, 0.64 + 1.96 \times 0.048] \approx [0.546, 0.734]$

(2) 標本数を n, 標本平均を $\bar{x}$, 標本分散を s^2, 母平均を μ, 母分散を σ^2 とすると, 中心極限定理と $s \xrightarrow{p} \sigma$ であることから

$$\mathrm{P}\left(\left|\frac{\sqrt{n}(\bar{x}-\mu)}{s}\right| \le 1.96\right) \approx 0.95$$

となり, 左辺括弧内の不等式を μ について解くと, 95%近似信頼区間は

$$\left[\bar{x} - 1.96\frac{s}{\sqrt{n}}, \bar{x} + 1.96\frac{s}{\sqrt{n}}\right]$$

$n = 100, \bar{x} = 155.0, s = 10.0$ を代入すると $[153.04, 156.96]$