2003 年 12 月 17 日　（今野）

レポート問題（情報統計学）

1. 以下の $\boxed{\text{問題 1}}$ から $\boxed{\text{問題 12}}$ の中から 2 題以上を選んだ上，解答せよ．
2. 締め切りは 2004 年 1 月 15 日（木）16 時半
3. レポートの提出先は数研の今野の郵便受け．提出後は，確認のために下記までメールをお願いします．
4. 問題に不明な点があれば，email: konno@fc.jwu.ac.jp まで連絡してください．
5. 訂正等は

   http://mp-w3math.jwu.ac.jp/~konno/stati.html

   に提示します．

$\boxed{\text{問題 1}}$ 　X と Y は独立で $\mathbb{E}[X^2] < \infty$, $\mathbb{E}[Y^2] < \infty$ とする．このとき，

$$\mathbb{VAR}[aX + bY] = a^2\mathbb{VAR}[X] + b^2\mathbb{VAR}[Y]$$

を示せ．ただし，a, b は定数とする．

$\boxed{\text{問題 2}}$ 　$\epsilon_1, \epsilon_2, \ldots, \epsilon_n$, $n \geq 2$ を成功する確率が $p\,(0 < p < 1)$ の独立なベルヌーイ試行とする．すなわち，

$$\mathbb{P}(\epsilon_1 = 1) = p, \qquad \mathbb{P}(\epsilon_1 = 0) = 1 - p$$

である．このとき

$$\mathbb{VAR}\left(\frac{1}{n}(\epsilon_1 + \epsilon_2 + \cdots + \epsilon_n)\right) = \frac{p(1-p)}{n}$$

を示せ．

$\boxed{\text{問題 3}}$ 　$\Phi(x)$ を標準正規分布の分布関数とする．すなわち，

$$\Phi(x) = \int_{-\infty}^{x} (1/\sqrt{2\pi}) e^{-t^2/2}\, dt$$

である．ただし，x は実数である．このとき，$G(x) = \Phi(ax)$ は正規分布 $N(0,\, 1/a^2)$ の分布関数となることを示せ．ただし，$a > 0$ とする．

$\boxed{\text{問題 4}}$ 　Z が標準正規分布に従うとき，$\mathbb{E}[Z^k]\,(k = 1,\, 2,\, 3,\, 4)$ を求めよ．

$\boxed{\text{問題 5}}$ 　$X = \sigma Z + \mu$ とおく．ただし，Z は標準正規分布に従い，$-\infty < \mu < \infty$，$\sigma > 0$ とする．このとき，

$$\text{(a)} \quad \mathbb{COV}(X,\, X^2) \qquad \text{(b)} \quad \mathbb{VAR}(X^2)$$

を求めよ．

$\boxed{\text{問題 6}}$ X は平均 λ のポアソン分布

$$f_X(x|\,\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \qquad x = 0,\, 1,\, 2,\, \ldots$$

に従うとする．ただし，$\lambda > 0$ は未知とする．このとき，この分布は指数分布族に属することを示せ．また，自然母数および自然母数空間は明示せよ．

$\boxed{\text{問題 7}}$ $X_1, X_2, \ldots, X_n$ をベルヌーイ分布

$$f_\theta(x|\,\theta) = \theta^x (1-\theta)^{1-x}, \qquad x = 0,\, 1, \qquad \theta \in (0,\, 1)$$

からのランダム標本とし，$\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ とし，$S(\boldsymbol{X}) = \sum_{i=1}^n X_i$ とする．$S = s$ が与えられたときの $\boldsymbol{X}$ の条件付き確率

$$P(\boldsymbol{X} = \boldsymbol{x} | S = s)$$

が θ に依存しないことを示すことにより，S が θ の十分統計量であることを確認せよ．

$\boxed{\text{問題 8}}$

(i) X を一様分布

$$f_X(x|\,\theta) = \frac{1}{\theta} I_{[0,\,\theta]}(x)$$

に従う確率変数とする．ただし，$\theta > 0$ であるが未知とし，

$$I_{[0,\,\theta]}(x) = \begin{cases} 1, & (0 \le x \le \theta), \\ 0, & (\text{その他}), \end{cases} \tag{1}$$

とする．このとき，$X = x$ が与えられたときの尤度関数 $L_1(\theta|\,x)$ のグラフ（横軸が θ で縦軸が尤度関数の値）を作図した上で，最尤推定値を求めよ．

(ii) $X_1, X_2, \ldots, X_n\,(n \ge 2)$ は確率密度関数が (1) で与えられる一様分布からの標本の大きさが n のランダム標本とする．すなわち，これらの確率変数列は互いに独立で，各 $X_i\,(i = 1, 2, \ldots, n)$ は確率密度関数が (1) で与えられる一様分布 に従う．このとき，$X_i = x_i\,(i = 1, 2, \ldots, n)$ が与えられたときの尤度関数 $L_n(\theta|\,\boldsymbol{x}) = \Pi_{i=1}^n f_X(x_i|\,\theta)$ を $\max(\boldsymbol{x})$ を用いて θ の関数としてわかるように明示的に表現せよ．さらに，尤度関数 $L_n(\theta|\,\boldsymbol{x})$ のグラフ（横軸が θ で縦軸が尤度関数の値）を作図した上で，最尤推定値を求めよ．ただし，$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ とし，$\max(\boldsymbol{x})$ は $x_1, x_2, \ldots, x_n$ の最大値である．

$\boxed{\text{問題 9}}$ $X_1, X_2, \ldots, X_n$ は母数 $p\,(0 < p < 1)$ のベルヌーイ分布に従うとする．すなわち，

$$\mathbb{P}(X_1 = 0) = 1 - p, \qquad \mathbb{P}(X_1 = 1) = p$$

である．このとき，

$$U_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \xrightarrow{a.s.} p(1-p), \qquad \text{ただし } \bar{X}_n = \frac{1}{n} X_i$$

が成立することを示せ．(大数の強法則を用いる)

$\boxed{\text{問題 10}}$

(i)　確率変数 Y が母数 λ の指数分布

$$f_Y(y|\,\lambda) = \lambda e^{-\lambda y} I_{(0,\,\infty)}$$

に従うとする．ただし，$\lambda > 0$ である．このとき，Y の期待値 $\mathbb{E}[Y]$ を求めよ．

(ii) 確率変数 X は確率密度関数

$$f_X(x|\,\alpha) = \alpha x^{-\alpha-1} I_{(1,\,\infty)}(x), \qquad I_{(1,\,\infty)}(x) = \begin{cases} 1, & (x > 1), \\ 0, & (\text{その他}), \end{cases} \tag{2}$$

を持つとする．ただし，$\alpha > 0$ である．$\log X$ は母数 α の指数分布にしたことを示せ．

(iii) 確率変数列 $X_1, X_2, \ldots, X_n$ は独立同一に分布に従い，その分布は確率密度関数 (2) を持つとする．

$$S_n = (\Pi_{i=1}^n X_i)^{1/n}$$

とおく．$n \to \infty$ のとき

$$S_n \xrightarrow{P} e^{1/\alpha}$$

を示せ．($\log S_n$ に対して，大数の法則と定理 5.8 を利用する)

$\boxed{\text{問題 11}}$　確率変数 X はパレート分布

$$f_X(x|\,\theta) = \theta k^\theta x^{-(\theta+1)} I_{(k,\,\infty)}(x) \tag{3}$$

に従うとする．ただし，$\theta > 0$ は未知で，$k > 1$ は既知とする．また，実数 $a, b\,(a < b)$ に対し

$$\begin{aligned} I_{(a,\,b)}(x) &= \begin{cases} 1, & (a < x < b), \\ 0, & \text{その他} \end{cases}, \quad I_{[a,\,b)}(x) = \begin{cases} 1, & (a \le x < b), \\ 0, & \text{その他} \end{cases}, \\ I_{(a,\,\infty)}(x) &= \begin{cases} 1, & (x > a), \\ 0, & \text{その他} \end{cases} \end{aligned}$$

等と記号を定めることにする．

(i)　X の分布関数 $F_X(x) = P(X \le x)$ を求めよ．

(ii)　実数 a に対して，$\lfloor a \rfloor$ を a 以下の最大の整数とする．このとき，$P(\lfloor X \rfloor = y)$ を求めよ．ただし，$y \ge \lfloor k \rfloor$ とする．

(iii)　$\lfloor X \rfloor = y$ が与えられたときの X の条件付確率密度関数が

$$k_\theta(x|\,y) = \frac{\theta x^{-(\theta+1)}}{y^{-\theta} - (y+1)^{-\theta}} I_{[y,\,y+1)}(x)$$

で与えられることを示せ．ただし，$y \ge \lfloor k \rfloor$ とする．

(iv)　$\lfloor X \rfloor = y$ が与えられたときの条件付き期待値が

$$\mathbb{E}[\,\log X|\,\lfloor X \rfloor = y] = \frac{1}{\theta} + g(y,\,\theta)$$

となることを示せ．ただし，

$$g(y,\,\theta)=\frac{\log y}{1-\{y/(y+1)\}^{\theta}}-\frac{\log(y+1)}{\{(y+1)/y)\}^{\theta}-1}$$

とした

(v)　確率変数列 $X_1,\,X_2,\,\ldots,\,X_n$ は確率密度関数 (3) を持つ分布に独立同一に従うとする．$\lfloor X_i\rfloor=y_i\,(i=1,\,2,\,\ldots,\,n)$ を観測したとする．観測 $y_1,\,y_2,\,\ldots,\,y_n$ に基づいて θ の最尤推定値を EM アルゴリズムで求めることを考える．初期推定値を $\widehat{\theta}^{(0)}$ としたとき，m 番目の逐次推定量は

$$\widehat{\theta}^{(m)}=\frac{1}{\{\widehat{\theta}^{(m-1)}\}^{-1}+(1/n)\sum_{i=1}^{n}g(y_i,\,\widehat{\theta}^{(m-1)})-\log k},\qquad m=0,\,1,\,2,\,\ldots$$

で与えられることを示せ．

問題 12

(i)　$X_1,\,X_2,\,\ldots,X_n\,(n\geq 2)$ を正規分布 $N(\mu,\,\sigma^2)$ からの大きさ n のランダム標本とし，$X_{(1)},\,X_{(2)},\,\ldots,X_{(n)}\,(X_{(1)}\leq X_{(2)}\leq\cdots\leq X_{(n)})$ をその順序統計量とする．

$$m_n=\begin{cases}\frac{n+1}{2}, & (n\text{ が奇数のとき})\\ \frac{n}{2}, & (n\text{ が偶数のとき})\end{cases}$$

とおき，$Z_n=X_{(m_n)}$ を標本メデアンとする．X_1 のメデアンは μ となることを示せ．さらに，$\sqrt{n}(Z_n-\mu)$ はどのような分布に分布収束するかを答えよ．

(ii)　$X_1,\,X_2,\,\ldots,X_n\,(n\geq 2)$ をコーシー分布

$$f(x)=\frac{1}{\pi\sigma}\,\frac{1}{1+[(x-\mu)/\sigma]^2}I_{(-\infty,\infty)}$$

からの大きさ n のランダム標本とし，$X_{(1)},\,X_{(2)},\,\ldots,X_{(n)}\,(X_{(1)}\leq X_{(2)}\leq\cdots\leq X_{(n)})$ をその順序統計量とする．

$$m_n=\begin{cases}\frac{n+1}{2}, & (n\text{ が奇数のとき})\\ \frac{n}{2}, & (n\text{ が偶数のとき})\end{cases}$$

とおき，$Z_n=X_{(m_n)}$ を標本メデアンとする．X_1 のメデアンは μ となることを示せ．さらに，$\sqrt{n}(Z_n-\mu)$ はどのような分布に分布収束するかを答えよ．