今野　(2009 年 7 月 22 日)

統計解析・演習の試験問題（試験時間は 80 分）

解答上の注意

特別な指示がある場合を除き，答えが合っているかどうかよりも解答の途中過程の論理的展開を重視して採点をします．以下の点に留意して解答を作成すること．

(1) 解答の途中過程は丁寧に記述すること．

(2) 講義で述べたことは設問中で証明することを求めていない場合には証明なしに用いてよい．しかし，なにをどのように用いたかを可能な限り明示すること．

(3) 各自が理解していることを採点者にわかるように解答を作成することを心がけること．

(4) 等号の使い方に注意すること．

(5) どの問題を解答しているかが明示されていれば，解答は問題番号順でなくともよい．

答案は後期開講の「情報統計学」の第 1 回目の講義で返却をする.「情報統計学」の第 1 回目後に，受講しないものは研究室まで取りに来てください．

問題 1　確率変数 $X_1, X_2, \ldots, X_n$ の平均，分散，相関係数はつぎのように共通とする：

$$\begin{aligned} \mu &= E(X_i), \qquad \sigma^2 = V(X_i), \qquad (i = 1, 2, \ldots, n) \\ \rho &= \rho(X_i, X_j) = \frac{COV(X_i, X_j)}{\sigma^2}, \qquad (i \neq j : i, j = 1, 2, \ldots, n) \end{aligned}$$

これら n 個の確率変数の標本平均を

$$\bar{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

とするとき，以下の問に答えよ．

(1) $E(\bar{X}_n)$ を求めよ..

(2) $V(\bar{X}_n) = \dfrac{1}{n}\sigma^2 + \dfrac{n-1}{n}\rho\sigma^2$ を示せ．

(3) $X_1, X_2, \ldots, X_n$ が独立とする．このとき，$i \neq j : i, j = 1, 2, \ldots, n$ に対して，$\rho(X_i, X_j)$ の値を求めよ．

(4) 特に，$X_1, X_2, \ldots, X_n$ が独立のとき，$V(\bar{X}_n)$ を求めよ．

つぎの期待値の性質は証明なしで用いてよい：

(i) $E(aX_1 + bX_2) = a\,E(X_1) + b\,E(X_2)$．ただし，$a, b$ は定数.

(ii) $V(aX_1 + bX_2) = a^2V(X_1) + b^2V(X_2) + 2ab\,COV(X_1, X_2)$.

(iii) $COV(X_1, X_2) = E[(X_1 - E(X_1))(X_2 - E(X_2))]$.

(iv) X_1, X_2 が独立のとき，$E(X_1X_2) = E(X_1)E(X_2)$.

(v) X が非負値確率変数[1]のとき，$\mathbb{E}[X] = 0$ ならば，$\mathbb{P}(X = 0) = 1$ である．

(vi) c が定数のとき，$E(c) = c$.

(vii) $V(X_1) = E[(X_1 - E(X_1))^2]$.

ただし，上記において，いずれの期待値も存在するものと仮定する．

解答と配点　(1) 10 (2) 15 (3) 10 (4) 5 計 40

[1] $P(X \geq 0) = 1$ である．一般には，$\mathbb{E}[X] = 0$ であっても，$\mathbb{P}(X = 0) = 1$ ではないことに注意する．

(1) (i) を用いる：

$$E(\bar{X}_n) \quad = \quad E\left[\frac{X_1+\cdots+X_n}{n}\right] = \frac{1}{n}\{E(X_1)+\cdots+E(X_n)\} = \frac{n\times\mu}{n} = \mu$$

(2)　一般に，$COV(X_i,\,X_j) = COV(X_j,\,X_i)$ であることに注意して (ii) を用いる：

$$\begin{aligned} V(\bar{X}_n) &= V\left[\frac{X_1+\cdots+X_n}{n}\right] \\ &= \frac{1}{n^2}\left\{\sum_{i=1}^{n} V(X_i) + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} COV(X_i,\,X_j)\right\} \end{aligned}$$

仮定より

$$V(X_i) = \sigma^2, \qquad COV(X_i,\,X_j) = \sigma^2\rho\,(i \neq j)$$

を上の式に代入すれば，

$$\begin{aligned} V(\bar{X}_n) &= \frac{1}{n^2}\left\{\sum_{i=1}^{n}\sigma^2 + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n}\sigma^2\rho\right\} \\ &= \frac{1}{n^2}\{n\times\sigma^2 + 2\times\frac{n(n-1)}{2}\sigma^2\rho\} = \frac{1}{n}\sigma^2 + \frac{n-1}{n}\sigma^2\rho. \end{aligned}$$

(3)　$X_i,\,X_j\,(i \neq j)$ が独立ならば，(iii) と (iv) を用いると

$$COV(X_i,\,X_j) \quad = \quad E[(X_i - E(X_i))(X_j - E(X_j))] = E(X_i - E(X_i))E(X_j - E(X_j)) = 0$$

より $\rho = 0$ となる．
(4)　$\rho = 0$ と (2) より，

$$V(\bar{X}_n) = \frac{1}{n}\sigma^2$$

問題 2　2 つの連続型確率変数 X, Y は独立で，それぞれの確率密度関数は $f_X(x),\, f_Y(y)$ であるとする．

(1) 確率変数 X, Y の同時確率密度関数 $f_{(X,\,Y)}(x,\,y)$ を $f_X(x),\, f_Y(y)$ で表せ．答えのみでよい．

(2) $Z=X+Y$ と $W=Y$ としたとき，Z, W の同時確率密度関数を $f_X,\, f_Y$ を用いて表せ．

(3) Z の周辺確率密度関数 f_Z を求めよ．

ヒント　x_1x_2 平面から y_1y_2 平面の中への 1 対 1 変換

$$y_1 \;=\; \varphi_1(x_1,\,x_2),\, y_2=\varphi_2(x_1,\,x_2)$$

と

$$x_1=\psi_1(y_1,\,y_2),\, x_2=\psi_2(y_1,\,y_2)$$

に対して，

$$J=\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2}-\frac{\partial x_1}{\partial y_2}\frac{\partial x_2}{\partial y_1}$$

としたとき，$Y_1=\varphi_1(X_1,\,X_2),\, Y_2=\varphi_2(X_1,\,X_2)$ の確率密度関数は，

$$f_{(Y_1,\,Y_2)}(y_1,\,y_2)=f_{(X_1,\,X_2)}(\psi_1(y_1,\,y_2),\,\psi_2(y_1,\,y_2))|J|$$

で与えられる．ただし，$f_{(X_1,\,X_2)}$ は $X_1,\, X_2$ の同時確率密度関数である．

解答と配点　(1) 10 (2) 10 (3) 10 計 30

(1) $f_{(X,\,Y)}(x,\,y)=f_X(x)f_Y(y)$.

(2)　$x=z-w,\, y=z$ に注意する：

$$J=\begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} \frac{\partial}{\partial z}(z-w) & \frac{\partial}{\partial w}(z-w) \\ \frac{\partial}{\partial z}w & \frac{\partial}{\partial w}w \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = 1$$

よって，

$$f_{(Z,\,W)}(z,\,w)=f_{(Z,\,W)}(z-w,\,w)|J|=f_X(z-w)f_Y(w).$$

(3)

$$f_Z(z)=\int_{-\infty}^{\infty} f_{(Z,\,W)}(z,\,w)\,dw=\int_{-\infty}^{\infty} f_X(z-w)f_Y(w)\,dw$$

問題 3　離散型確率変数 X は母数 $p\,(0<p<1)$ のベルヌーイ分布に従う：

$$f_X(x)=\begin{cases} p^x(1-p)^{1-x}, & x=0,\,1; \\ 0 & (\text{その他}). \end{cases}$$

このとき，以下の問いの答えよ．

(1) X のモーメント母関数 $M_X(t)=E(e^{tX})$ を求めよ．ただし，t は $E(e^{tX})<\infty$ をみたす数である．

(2) これを用いて $E(X),\, E(X^2)$ を求めよ．

(3) $V(X)$ を求めよ．問題 1 のヒントも用いてよい．

ヒント 確率変数 X は確率関数または確率密度関数　$f_X(x)$ を持つとし，$g(x)$ を実数値関数とする．このとき，$g(X)$ の期待値を

$$\mathbb{E}[g(X)]=\begin{cases} \sum_{x\in S} g(x)f_X(x), & (\text{離散型}) \\ \int_{-\infty}^{\infty} g(x)f_X(x)\,dx, & (\text{連続型}) \end{cases}$$

で定義する．ただし，$\sum_{x\in S}|g(x)|f_X(x)<\infty$ もしくは $\int_{-\infty}^{\infty}|g(x)|f_X(x)\,dx<\infty$ のとき，$g(X)$ の期待値を定義することにする．期待値が定義されるとき，$g(X)$ の期待値が存在するという．ただし，X が離散型のとき，$S=\{x\in\mathbb{R}:\, f_X(x)>0\}$ とし，これは高々可算集合となる．

解答と配点　(1) 10 (2) 10 (3) 10 計 30

(1)

$$
\begin{aligned}
E(e^{tX}) &= \sum_{x=0,\,1} e^{tx} f_X(x) = e^{t\cdot 0} f_X(0) + e^{t\cdot 1} f_X(1) \\
&= e^{t\cdot 0}(1-p) + e^{t\cdot 1} p = (1-p) + pe^t.
\end{aligned}
$$

(2)　期待値とモーメント母関数の関係より

$$
\begin{aligned}
E(X) &= \frac{d}{dt} M(t)\Big|_{t=0} = pe^t\Big|_{t=0} = p; \\
E(X^2) &= \frac{d^2}{dt^2} M(t)\Big|_{t=0} = pe^t\Big|_{t=0} = p.
\end{aligned}
$$

(3)　(2) の結果と問題 1 のヒント (vii), (i), および (vi) より

$$
\begin{aligned}
V(X) &= E[(X-E(X))^2] = E[(X-p)^2] = E(X^2 - 2pX + p^2) \\
&= E(X^2) - 2pE(X) + p^2 = p - 2p^2 + p^2 = p(1-p)
\end{aligned}
$$

==

90 点以上:A+・89〜70 点 ：A・69〜50 点：B・49〜20 点：C

==