今野　(2004 年 7 月 21 日)

統計解析・演習の試験問題の解答例

$\boxed{\text{問題 1}}$　連続型確率変数 X は確率密度関数

$$f_X(x) = \begin{cases} 1 & (0 < x < 1), \\ 0 & (\text{その他}), \end{cases}$$

をもつとする．このとき，以下の問いに答えよ．

(a) X の分布関数 $F_X(x) = \mathbb{P}(X \le x)$（$x \in \mathbb{R}$）と X の期待値 $\mathbb{E}[X]$ を求めよ．ただし，$F_X(x)$ は $\mathbb{R}$ 上の関数であることがわかるように答えよ．

(b) $Y = -2\log X$ としたとき，Y の分布関数 $F_Y(y) = \mathbb{P}(Y \le y)$（$y \in \mathbb{R}$）と確率密度関数 $f_Y(y)$ を求めよ．ただし，$F_Y(y)$ は $\mathbb{R}$ 上の関数であることがわかるように答えよ．また，$f_Y(y) > 0$ となる y の範囲を明示すること．

(c) Y の期待値 $\mathbb{E}[Y]$ を求めよ．

$\boxed{\text{ヒント}}$　以下は既知としてよい．　$\frac{d}{dx}\{\frac{x^2}{2}\log x - \frac{x^2}{4}\} = x\log x$,　$\frac{d}{dx}\{-xe^{-x} - e^{-x}\} = xe^{-x}$,
$\lim_{x\to\infty} xe^{-x} = 0$,　$\lim_{x\to 0} x^2 \log x = 0$.

$\boxed{\text{解答}}$　(a) $x \in (0, 1)$ に対して，

$$F_X(x) = \int_0^x dx = x$$

となる．明らかに，$x \le 0$ に対しては $F_X(x) = 0$，$x \ge 1$ に対しては $F_X(x) = 1$ となる．したがって，

$$F_X(x) = \begin{cases} 0 & (x \le 0), \\ x & (0 < x < 1), \\ 1 & (x \ge 0) \end{cases}$$

となる．

期待値の定義と与えられた確率密度関数の台より

$$\mathbb{E}[X] = \int_0^1 x\,dx = \frac{1}{2}$$

となる．

(b) $\mathbb{P}(0 < X < 1) = 1$ より $\mathbb{P}(Y > 0) = 1$ となる．このことより任意の $y \le 0$ に対して，$\mathbb{P}(Y \le y) = 0$ がわかる．任意の $y > 0$ に対して

$$\mathbb{P}(Y \le y) = \mathbb{P}(-2\log X \le y) = \mathbb{P}(X \ge e^{-y/2})$$

となる．X は連続型確率変数なので，$\mathbb{P}(X = e^{-y/2})$ と確率の性質（事象 B に対して，$\mathbb{P}(B^c) = 1 - \mathbb{P}(B)$）より

$$\mathbb{P}(X \ge e^{-y/2}) = 1 - \mathbb{P}(X \le e^{-y/2}) = 1 - e^{-y/2}$$

となる．最後の等号は $0 < e^{-y/2} < 1$ と X の分布関数の形よりわかる．したがって，

$$F_Y(y) = \begin{cases} 0 & (y \le 0), \\ 1 - e^{-y/2} & (0 < y < \infty), \end{cases}$$

を得る．

$y>0$ に対して，Y の確率密度関数は

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{1}{2}e^{-y/2}$$

となることがわかる．よって，

$$f_Y(y) = \begin{cases} \frac{1}{2}e^{-y/2} & (y>0), \\ 0 & (y \le 0), \end{cases}$$

(c) Y の確率密度関数の台から Y の期待値は

$$\mathbb{E}[Y] = \int_0^\infty y f_Y(y)\,dy = \int_0^\infty y\,\frac{1}{2}e^{-y/2}\,dy = \int_0^\infty ze^{-z}\,dz = \left[-ze^{-z} - e^{-z}\right]_0^\infty = 1$$

となる．

$\boxed{\text{問題 2}}$ 連続型確率変数 X は確率密度関数

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \qquad (-\infty < x < \infty)$$

を持つとする．ただし，μ, σ は定数で $-\infty<\mu<\infty, 0<\sigma<\infty$ とし，$\exp(x)=e^x$ である．

$$Z = \frac{X-\mu}{\sigma}$$

と定義したとき，以下の問いに答えよ．

(a) Z の確率密度関数 $f_Z(z)$ を求めよ．

(b) $10Z+50$ の期待値 $\mathbb{E}[10Z+50]$ と分散 $\mathbb{VAR}[10Z+50]$ を求めよ．ただし，講義で述べた期待値と分散の性質と $\mathbb{E}[Z]=0, \mathbb{E}[Z^2]=1$ は証明なしで用いてよい．

$\boxed{\text{解答}}$(a) 確率密度関数の定義から

$$\begin{aligned} F_Z(z) &= \mathbb{P}(Z \le z) = \mathbb{P}(X \le \sigma z + \mu) = \int_{-\infty}^{\sigma z+\mu} f_X(t)\,dt \\ &= \int_{-\infty}^{z} f_X(\sigma s+\mu)\,\sigma ds \qquad (s=(t-\mu)/\sigma \text{ とおいた}) \\ &= \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}e^{-s^2/2}\,ds \end{aligned}$$

となる．したがって，

$$f_Z(z) = \frac{d}{dz}F_Z(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2/2}, \qquad (-\infty<z<\infty)$$

となる．

(b) 分散公式から

$$\mathbb{VAR}[Z] = \mathbb{E}[Z] - \{\mathbb{E}[Z]\}^2 = 1$$

となる．期待値と分散の性質より

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}(10Z+50) = 10\mathbb{E}(Z)+50 = 50 \\ \mathbb{VAR}(X) &= \mathbb{VAR}(10Z+50) = 10^2\mathbb{VAR}Z) = 100 \end{aligned}$$

となる．

なお，期待値の定義から

$$
\begin{aligned}
\mathbb{E}(Z) &= \int_{-\infty}^{\infty} z f_Z(z)\,dz = \lim_{M\to\infty}\int_0^M z f_Z(z)\,dz + \lim_{M\to\infty}\int_{-M}^{\infty} z f_Z(z)\,dz \\
&= \lim_{M\to\infty}\left[\frac{1}{\sqrt{2\pi}} - \frac{1}{\sqrt{2\pi}}e^{-M^2/2}\right] + \lim_{M\to\infty}\left[\frac{1}{\sqrt{2\pi}}e^{-M^2/2} - \frac{1}{\sqrt{2\pi}}\right] = 0 \\
\mathbb{E}(Z^2) &= \int_{-\infty}^{\infty} z^2 f_Z(z)\,dz = \lim_{M\to\infty}\int_0^M z^2 f_Z(z)\,dz + \lim_{M\to\infty}\int_{-M}^{\infty} z^2 f_Z(z)\,dz \\
&= \lim_{M\to\infty}\left\{\left[-\frac{1}{\sqrt{2\pi}} z e^{-z^2/2}\right]_{-M}^{0} + \int_{-M}^{0}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\,dz\right\} \\
&\quad + \lim_{M\to\infty}\left\{\left[-\frac{1}{\sqrt{2\pi}} z e^{-z^2/2}\right]_{0}^{M} + \int_{0}^{M}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\,dz\right\} \\
&= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\,dz = 1
\end{aligned}
$$

となることからヒントはわかる．

問題 3　離散型確率変数 X, Y の同時確率関数 $f_{X,Y}(x, y)$ は下の表のように与えられているとする．

$x \backslash y$	-1	0	1
-1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$
0	$\frac{3}{16}$	0	$\frac{3}{16}$
1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$

(a) X と Y の周辺確率関数 $f_X(x)$ と $f_Y(y)$ を求めよ．ただし，答えのみでよい．

(b) X の周辺分布関数 $F_X(x) = \mathbb{P}(X \leq x)$（$y \in \mathbb{R}$）を求め，$F_X(x)$ のグラフを作図せよ．ただし，答えのみでよい．

(c) X と Y は独立か従属かを調べよ．さらに，その理由を述べること．

解答　(a)

x	-1	0	1	計
$f_X(x)$	$\frac{5}{16}$	$\frac{6}{16}$	$\frac{5}{16}$	1

y	-1	0	1	計
$f_Y(y)$	$\frac{5}{16}$	$\frac{6}{16}$	$\frac{5}{16}$	1

(b)

$$
F_X(x) = \begin{cases} 0 & (x < -1) \\ \dfrac{5}{16} & (-1 \leq x < 0) \\ \dfrac{11}{16} & (0 \leq x < 1) \\ 1 & (x \geq 1) \end{cases}
$$

(c) X と Y が独立であるためには，すべての x と y に対して

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

が成立しなければならない．たとえば，$x=1,\,y=1$ で調べてみる．

$$f_{X,Y}(1,\,1)=\frac{1}{16}\qquad f_X(1)f_Y(1)=\left(\frac{5}{16}\right)^2$$

したがって，X と Y は従属．

$\boxed{\text{問題 4}}$ 連続型確率変数 $X,\,Y$ は同時確率密度関数

$$f_{X,Y}(x,\,y)=\begin{cases} x+y & (0<x<1,\,0<y<1), \\ 0 & (\text{その他}), \end{cases}$$

を持つとする．

(a) X の周辺確率密度関数 $f_X(x)$ を求めよ．なお，解答には $f_X(x)>0$ となる x の範囲を明示すること．

(b) X の期待値 $\mathbb{E}[X]$，分散 $\mathbb{VAR}[X]$，XY の期待値 $\mathbb{E}[XY]$ および X と Y の共分散 $\mathbb{COV}[X,\,Y]$ を求めよ．ただし，最終的な数値を求めるために必要な分数の足し算や引き算の計算（通分）はしなくともよい．

(c) $X=x$ が与えられたときの Y の条件付確率密度関数 $f_{Y|X}(y|\,x)$ を求めよ．ただし，解答には $f_{Y|X}(y|\,x)$ が定義される x の範囲と $f_{Y|X}(y|\,x)>0$ となる y の範囲を明示すること．

(d) $X=x$ が与えられたときの Y の条件付期待値 $\mathbb{E}(Y|\,x)$ を求めよ．

(e) $U=X+Y$ と $V=X-Y$ とおいたとき，U と V の同時確率確率密度関数 $f_{U,V}(u,\,v)$ を求めよ．ただし，$f_{U,V}(u,\,v)>0$ となる u と v の範囲を明示すること．

(f) V の周辺確率密度関数 $f_V(v)$ を求めよ．ただし，$f_V(v)>0$ となる v の範囲を明示せよ．

$\boxed{\text{解答}}$ (a) $0<x<1$ に対して

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,\,y)\,dy \\ &= \int_{-\infty}^{0} f_{X,Y}(x,\,y)\,dy+\int_0^1 f_{X,Y}(x,\,y)\,dy+\int_1^{\infty} f_{X,Y}(x,\,y)\,dy \\ &= \int_0^1 (x+y)\,dy=\left[xy+\frac{y^2}{2}\right]_0^1=x+\frac{1}{2} \end{aligned}$$

となる．$x\notin(0,\,1)$ では $f_X(x)=0$ となる．したがって，

$$f_X(x)=\begin{cases} 0 & (x\le 0) \\ x+\frac{1}{2} & (0<x<1) \\ 0 & (x\ge 1) \end{cases}$$

となる．

(b) $f_X(x)>0$ となる x の範囲に注意すると X の期待値は

$$\mathbb{E}[X]=\int_0^1 xf_X(x)\,dx=\left[\frac{x^3}{3}+\frac{x^2}{4}\right]_0^1=\frac{7}{12}$$

となる．同様に，X^2 の期待値は

$$\mathbb{E}[X^2]=\int_0^1 x^2 f_X(x)\,dx=\left[\frac{x^4}{4}+\frac{x^3}{6}\right]_0^1=\frac{10}{24}=\frac{5}{12}$$

となる．したがって，分散公式より

$$\mathbb{VAR}[X]=\mathbb{E}[X^2]-\{\mathbb{E}[X]\}^2=\frac{5}{12}-\left(\frac{7}{12}\right)^2=\frac{11}{144}$$

XY の期待値は

$$\mathbb{E}[XY] = \int_0^1 \int_0^1 xy\,(x+y)\,dx\,dy = \int_0^1 y\left[\frac{x^3}{3} + y\frac{x^2}{2}\right]_0^1 dy = \int_0^1 y\left[\frac{1}{3} + y\frac{1}{2}\right]_0^1 dy = \left[\frac{y^2}{6} + \frac{y^3}{6}\right]_0^1 = \frac{1}{3}$$

同時確率密度関数 $f_{X,\,Y}(x,\,y)$ の x と y の対称性より

$$\mathbb{E}[Y] = \frac{7}{12}$$

がわかる．したがって，分散公式より

$$\mathbb{COV}(X,\,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{1}{3} - \left(\frac{7}{12}\right)^2 = -\frac{1}{144}$$

となる．
(c) $(0,\,1) = \{x \in \mathbb{R}:\, f_X(x) > 0\}$ から $0 < x < 1$ に対して $X = x$ が与えられたときの Y の条件付確率密度関数は定義され

$$f_{Y|\,X}(y|\,x) = \frac{f_{X,\,Y}(x,\,y)}{f_X(x)} = \begin{cases} \dfrac{x+y}{x+\frac{1}{2}} = \dfrac{2(x+y)}{2x+1} & (0 < y < 1), \\ 0 & (\text{その他}), \end{cases}$$

となる．
(d) $0 < x < 1$ に対して $X = x$ が与えられたときの Y の条件付期待値は定義され， $\{y:\, f_{Y|\,X}(y|\,x) > 0\} = (0,\,1)$ に注意すれば，

$$\mathbb{E}(Y|\,x) \;=\; \int_0^1 y f_{Y|\,X}(y|\,x)\,dy = \int_{-\infty}^0 \frac{2y(x+y)}{2x+1}\,dy = \frac{1}{2x+1}\left[2x\frac{y^2}{2} + \frac{2y^3}{3}\right]_0^1 = \frac{x+2}{3(2x+1)}$$

となる．
(e) U と V の定義から

$$\begin{cases} X = \dfrac{U+V}{2} \\ Y = \dfrac{U-V}{2} \end{cases}$$

となる．また，$\mathbb{P}(0 < X < 1,\, 0 < Y < 1) = 1$ より

$$\mathbb{P}\left(0 < \frac{U+V}{2} < 1,\, 0 < \frac{U-V}{2} < 1\right) = 1 \quad \Longleftrightarrow \quad \mathbb{P}\left(-U < V < 2-U,\, U-2 < V < U\right) = 1$$

となる．したがって，

$$\{(u,\,v):\, f_{U,\,V}(u,\,v) > 0\} = \{(u,\,v):\, -u < v < 2-u,\, u-2 < v < u\} =: A$$

となる．また，

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-1}{2} \end{vmatrix} = -\frac{1}{2}$$

となるので，$(u,\,v) \in A$ に対して，

$$f_{U,\,V}(u,\,v) = f_{X,\,Y}\left(\frac{u+v}{2},\, \frac{u-v}{2}\right)\left|-\frac{1}{4}\right| = \frac{u}{2}$$

となる．したがって，

$$f_{U,\,V}(u,\,v) = \begin{cases} \dfrac{u}{2} & (u,\,v) \in A, \\ 0 & (\text{その他}), \end{cases}$$

となる．
(f) 上の問いの領域 A から $\{v: f_V(v)>0\}=(-1,1)$ となる．領域 A の形に注意して場合わけをする：$0\le v<1$ に対して，

$$f_V(v)=\int_v^{2-v} f_{U,V}(u,v)f_{U,V}(u,v)\,du=\left[\frac{u^2}{4}\right]_v^{2-v}=1-v$$

となる．また，$-1<v<0$ に対して，

$$f_V(v)=\int_{-v}^{v+2} f_{U,V}(u,v)f_{U,V}(u,v)\,du=\left[\frac{u^2}{4}\right]_{-v}^{v+2}=1+v$$

となる．したがって，

$$f_V(v)=\begin{cases}1-v & (0\le v<1),\\ 1+v & (-1<v<0),\\ 0 & (\text{その他}),\end{cases}$$

となる．

$\boxed{\text{問題 5}}$ Ω を標本空間とし，$\mathcal{F}$ を Ω 上の完全加法族[1]とする．確率 $\mathbb{P}$ は $\mathcal{F}$ 上で定義された実数値関数でつぎの条件をみたすものであった．

(P1) 任意の $A\in\mathcal{F}$ に対して，$\mathbb{P}(A)\ge 0$
(P2) $\mathbb{P}(\Omega)=1$
(P3) $i=1,2,\ldots$ に対して $A_i\in\mathcal{F}$ かつ $A_i\cap A_j=\emptyset\,(i\ne j)$ ならば，

$$\mathbb{P}(\cup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}\mathbb{P}(A_i)$$

(P1) から (P3) をどこでどのように使ったかを明示（例に倣って）して (a)，(b) を証明せよ．

(a) (P1)–(P3) および (3) を用いてつぎのことを示せ：$B\in\mathcal{F}$ に対して

$$\mathbb{P}(B^c)=1-\mathbb{P}(B) \tag{2}$$

となる．ただし，B^c は B の補集合（B に含まれない Ω の元を集めたもの）である．

(b) (P1)–(P3) および (3)（必要ならば）を用いてつぎのことを示せ：$B_1, B_2\in\mathcal{F}$ に対して

$$B_1\subset B_2 \quad \text{ならば，} \qquad \mathbb{P}(B_1)\le\mathbb{P}(B_2)$$

となる．

$\boxed{\text{例}}$ たとえば，

$$\mathbb{P}(\emptyset)=0 \tag{3}$$

をを示すには，$A_1=\Omega,\ A_i=\emptyset\,(i\ge 2)$ とおくと

$$\Omega=\cup_{i=1}^{\infty}A_i \tag{4}$$

と

$$A_i\cap A_j=\emptyset\,(i\ne j) \tag{5}$$

となることに注意すれば，

$$\mathbb{P}(\Omega)=\mathbb{P}(\cup_{i=1}^{\infty}A_i)=\sum_{i=1}^{\infty}\mathbb{P}(A_i)=\mathbb{P}(\Omega)+\sum_{i=2}^{\infty}\mathbb{P}(\emptyset)$$

[1] Ω の部分集合のなす集まりで任意の可算回の集合演算に関して閉じている．

となる．ただし，1 番目の等号は (4) からわかり，2 番目の等号は (5) と (P3) からわかる．したがって，

$$\sum_{i=2}^{\infty} \mathbb{P}(\emptyset) = 0 \tag{6}$$

となる．しかし，(P1) から $\mathbb{P}(\emptyset) \geq 0$ なることと (6) から $\mathbb{P}(\emptyset) = 0$ がわかる．

$\boxed{\text{解答}}$ (a) (P3) において $A_1 = B$, $A_2 = B^c$, $A_i = \emptyset\,(i \geq 3)$ とおけば，(P3) の仮定（$A_i \cap A_j = \emptyset$, $i \neq j$）をみたす．また， $\Omega = B \cap B^c = \cup_{i=i} A_i$ となることに注意する．(P2) と (P3) を使えば，

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(B \cup B^c) = \mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) = \mathbb{P}(B_1) + \mathbb{P}(B_2) + \sum_{i=3}^{\infty} \mathbb{P}(\emptyset) = \mathbb{P}(B) + \mathbb{P}(B^c)$$

より示せた．最後の等号は (3) からわかる．
(b) $B_2 = B_1 \cup (B_1^c \cap B_2)$ と $B_1 \cap (B_1^c \cap B_2) = (B_1 \cap B_1^c) \cap B_2 = \emptyset$ に注意する．$A_1 = B_1$, $A_2 = B_1^c \cap B_2$, , $A_i = \emptyset$, $(i \geq 3)$ とおけば，

$$\cup_{i=1}^{\infty} A_i = B_2, \qquad A_i \cap A_j = \emptyset, \qquad (i \neq j),$$

となるので，(P3) を用いれば，

$$\mathbb{P}(B_2) = \mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) = \mathbb{P}(B_1) + \mathbb{P}(B_1^c \cap B_2) + \sum_{i=3}^{\infty} \mathbb{P}(A_i) = \mathbb{P}(B_1) + \mathbb{P}(B_1^c \cap B_2) \tag{7}$$

となる．最後の等号は (3) からわかる．最後に，(P1) を用いれば，$\mathbb{P}(B_1 \cap B_2) \geq 0$ から

$$\mathbb{P}(B_2) \geq \mathbb{P}(B_1)$$

を得る．

配点

$\boxed{\text{問題 1}}$ 25 点 (a) 5+5 (b) 5+5(c) 5 $\boxed{\text{問題 2}}$ 10 点 (a) 5 (b) 2+3
$\boxed{\text{問題 3}}$ 15 点 (a) 3+2 (b) 5 (c) 5 $\boxed{\text{問題 4}}$ 40 点 (a) 5 (b) 4+4+4+3 (c) 5 (d) 5 (e) 5 (f) 5 $\boxed{\text{問題 5}}$ 10 点 (a) 5 (b) 5

成績について

得点	$0 \sim 19$	$20 \sim 46$	$47 \sim 66$	$67 \sim 84$	$85 \sim 100$
成績	D	C	B	A	A^{+}

得点分布　平均点 $= 50.875$，中央値 $= 54.5$、標準偏差 $= 19.52045$

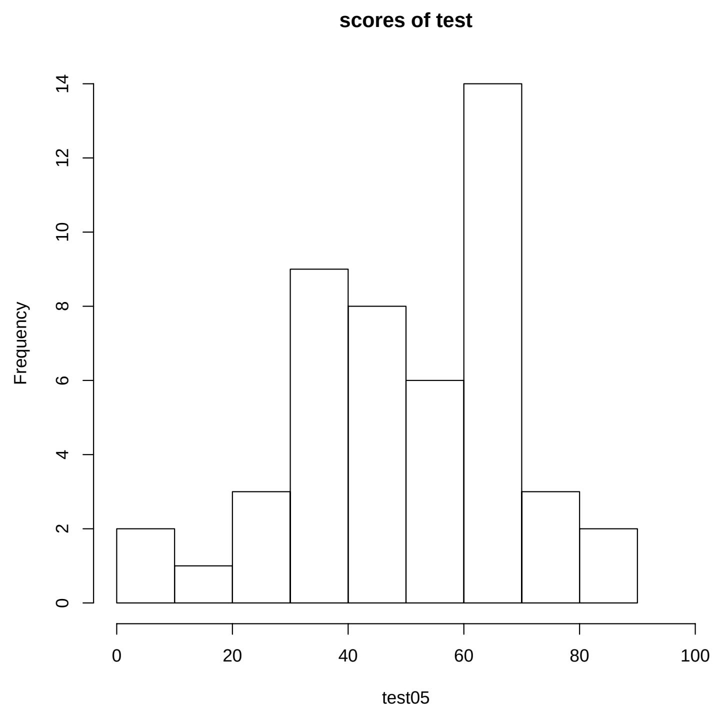

Figure 1: This is a figure