今野　(2004 年 7 月 21 日)

統計解析・演習の試験問題の解答例

問題 1　連続型確率変数 X は確率密度関数

$$f_X(x) = \begin{cases} c(1-x^2) & (-1 < x < 1) \\ 0 & (\text{その他}) \end{cases}$$

をもつとする．ただし，c は正の定数とする．このとき，以下の問いに答えよ．

(a) $f_X(x)$ が確率密度関数になるように c を定めよ．

(b) X の期待値と分散 $\mathbb{E}(X)$, $\mathbb{VAR}(X)$ を求めよ．ただし，c に (a) で得た値を代入せずに求めた答えのみを採点の対象とする．

(c) X の分布関数 $F_X(x) = \mathbb{P}(X \leq x)$ を求めよ．ただし，$-\infty < x < \infty$ である．

(d) チェビシェフの不等式[1]を用いて確率 $\mathbb{P}(-0.75 < X < 0.75)$ の下限を求めよ．

解答　(a) $f_X(x)$ が確率密度関数であるためには

- すべての x に対して，$f_X(x) \geq 0$
- $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

とならなければいけない．2 番目の条件から

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} f_X(x)\,dx = c\int_{-1}^{1}(1-x^2)\,dx = c\left[x - \frac{x^3}{3}\right]_{-1}^{1} \\ &= c\left\{\left(1-\frac{1}{3}\right) - \left((-1) - \frac{(-1)^3}{3}\right)\right\} = \frac{4}{3}c \end{aligned}$$

よって，

$$c = \frac{3}{4}$$

(b) 連続型確率変数に対する期待値の定義から

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{\infty} x f_X(x)\,dx = c\int_{-1}^{1} x(1-x^2)\,dx = 0 \qquad (x(1-x^2) \text{ は奇関数なので}) \\ \mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2 f_X(x)\,dx = c\int_{-1}^{1} x^2(1-x^2)\,dx \\ &= 2c\int_{0}^{1} x^2(1-x^2)\,dx \qquad (x^2(1-x^2) \text{ は偶関数なので}) \\ &= 2c\left[\frac{x^3}{3} - \frac{x^5}{5}\right]_0^1 = 2c\left(\frac{1}{3} - \frac{1}{5}\right) = \frac{4}{15}c \end{aligned}$$

[1] $\mathbb{E}(X^2) < \infty$ のときに，任意の正の数 a に対して

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{\mathbb{VAR}(X)}{a^2}$$

これより
$$\mathbb{VAR}(X)=\mathbb{E}(X^2)-\{\mathbb{E}(X)\}^2=\frac{4}{15}c$$
よって
$$\mathbb{E}(X)=0,\qquad \mathbb{VAR}(X)=\frac{1}{5}$$
(c) $\{x\in\mathbb{R}:\ f_X(x)>0\}=(-1,1)$ に注意して場合わけを行う．
$x\le -1$ の場合（$f_X(t)=0$）
$$F_X(x)=\int_{-\infty}^{x}f_X(t)\,dt=0$$
$-1<x<1$ の場合（$t\le -1$ ならば，$f_X(t)=0$，$-1<t\le x$ の場合 $f_X(t)=c(1-t^2)$ ）
$$\begin{aligned}F_X(x) &= \int_{-\infty}^{-1}f_X(t)\,dt+\int_{-1}^{x}f_X(t)\,dt=c\left[t-\frac{t^3}{3}\right]_{-1}^{x}\\ &= c\left\{\left(x-\frac{x^3}{3}\right)-\left((-1)-\frac{-1}{3}\right)\right\}\\ &= c\left(x-\frac{x^3}{3}+\frac{2}{3}\right)\end{aligned}$$
$x\ge 1$ の場合（$t\le -1$ ならば，$f_X(t)=0$，$-1<t<1$ の場合 $f_X(t)=c(1-t^2)$，$1\le t\le x$ の場合 $f_X(t)=0$）
$$F_X(x)=\int_{-\infty}^{-1}f_X(t)\,dt+\int_{-1}^{1}f_X(t)\,dt+\int_{1}^{x}f_X(t)\,dt=c\left[t-\frac{t^3}{3}\right]_{-1}^{1}=\frac{4}{3}c$$
したがって，$c=\dfrac{3}{4}$ を代入すれば，
$$F_X(x)=\begin{cases}0 & (-\infty<x\le -1)\\ \dfrac{3}{4}\left(x-\frac{x^3}{3}+\frac{2}{3}\right) & (-1<x<1)\\ 1 & (x\ge 1)\end{cases}$$
(d) $\mathbb{E}(X)=0$, $\mathbb{VAR}(X)=1/16$ に注意して，チェビシェフの不等式を用いる：
$$\mathbb{P}\left(-\frac{3}{4}<X<\frac{3}{4}\right)=\mathbb{P}\left(|X|<\frac{3}{4}\right)=1-\mathbb{P}\left(|X-0|\ge\frac{3}{4}\right)\ge 1-\frac{1/5}{(3/4)^2}=1-\frac{16}{45}=\frac{29}{45}$$
となる．

参考までに正確な確率を計算してみる．X は連続型確率変数なので，$\mathbb{P}(X=\pm\frac{3}{4})=0$ に注意する．(c) を利用すれば，
$$\begin{aligned}\mathbb{P}\left(-\frac{3}{4}<X<\frac{3}{4}\right) &= 1-2F_X\left(-\frac{3}{4}\right)=1-2\times\frac{3}{4}\left(-\frac{3}{4}+\left(\frac{3}{4}\right)^3\frac{1}{3}+\frac{2}{3}\right)\\ &= 1-\frac{2}{3}\times\left(-\frac{48}{64}+\frac{9}{64}+\frac{2}{3}\right)=1-\frac{2}{3}\times\frac{7}{192}=1-\frac{22}{243}=\frac{219}{243}\\ &= 0.9012346>0.6444444=\frac{29}{45}\end{aligned}$$

$\boxed{\text{問題 2}}$　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．
$$Z=\frac{Z-\mu}{\sigma}\qquad \text{タイプミスで正しくは } Z=\frac{X-\mu}{\sigma}$$

と定義したとき，以下の問いに答えよ．

(a) Z の分布関数

$$F_Z(z) = \mathbb{P}(Z \le z) \qquad (-\infty < z < \infty)$$

は μ と σ に依存しないことを示せ．

(b) Z の期待値と分散は

$$\mathbb{E}[Z] = 0, \qquad \mathbb{VAR}[Z] = 1$$

となることを期待値の定義から示せ．ただし，

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1 \quad \text{および} \quad \lim_{x\to\pm\infty} x\exp(-x^2/2) = 0$$

は証明なしで用いてよい．

(c) X の期待値と分散 $\mathbb{E}[X]$ と $\mathbb{VAR}[X]$ を求めよ．ただし，講義で述べた期待値と分散の性質と (b) の結果は証明なしで用いてよい．

$\boxed{\text{解答}}$ 以下はタイプミスの修正した後の想定された問題の解答です．得点は全員 20 点としました．(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，$F_Z(z)$ は μ と σ には依存しない．
(b) 前問の結果から Z の確率密度関数は

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

となることに注意する．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．したがって，

$$\mathbb{E}(Z) = 1, \qquad \mathbb{VAR}(Z) = \mathbb{E}(Z^2) - \{\mathbb{E}(Z)\}^2 = 1$$

を得る．
(c)

$$\begin{aligned}
\mathbb{E}(X) &= \mathbb{E}(\sigma Z + \mu) = \sigma\mathbb{E}(Z) + \mu = \mu \\
\mathbb{VAR}(X) &= \mathbb{VAR}(\sigma Z + \mu) = \sigma^2\mathbb{VAR}Z) = \sigma^2
\end{aligned}$$

$\boxed{\text{問題 3}}$ 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．

$x\setminus y$	-1	0	1
-1	α	β	α
0	β	0	β
1	α	β	α

ただし，α, β は定数で $\alpha > 0, \beta > 0$ かつ $\alpha + \beta = \dfrac{1}{4}$ である．

(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 $F_X(x) = \mathbb{P}(X \le x)$ 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, Y, XY の期待値 $\mathbb{E}(X), \mathbb{E}(Y), \mathbb{E}(XY)$ を計算せよ．

(d) X と Y の共分散 $\mathbb{COV}[X, Y]$ を求めよ．

(e) 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．

$\boxed{\text{解答}}$ (a)

x	-1	0	1
$f_X(x)$	$2\alpha+\beta$	2β	$2\alpha+\beta$

y	-1	0	1
$f_Y(y)$	$2\alpha+\beta$	2β	$2\alpha+\beta$

(b)

$$F_X(x) = \begin{cases} 0 & (x < -1) \\ 2\alpha+\beta & (-1 \le x < 0) \\ 2\alpha+3\beta & (0 \le x < 1) \\ 4(\alpha+\beta) = 1 & (x \ge 1) \end{cases}$$

(c) 離散型確率変数の期待値の定義より

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{x=-1}^{1} x f_X(x) = (-1)\times(2\alpha+\beta) + 0\times 2\beta + 1\times(2\alpha+\beta) = 0 \\
\mathbb{E}(Y) &= \sum_{x=-1}^{1} y f_Y(y) = (-1)\times(2\alpha+\beta) + 0\times 2\beta + 1\times(2\alpha+\beta) = 0 \\
\mathbb{E}(XY) &= \sum_{x,\,y=-1,\,0,\,1} xy f_{X,Y}(x, y) \\
&= (-1)\times(-1)\times f_{X,Y}(-1, -1) + 1\times 1\times f_{X,Y}(1, 1) = 0
\end{aligned}$$

となる．
(d) したがって，

$$\mathbb{COV}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0$$

となる．
(e) 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．

$$f_{X,Y}(1, 1) = \alpha \qquad f_X(1) f_Y(1) = (2\alpha+\beta)^2$$

したがって，これに $\alpha + \beta = 1/4$ を代入して整理すると独立であるために必要条件が

$$\alpha = (2\alpha+\beta)^2 \iff \left(a - \frac{1}{4}\right)^2 = 0$$

となり，独立であるための必要条件が $a = \dfrac{1}{4}$ であることがわかる．しかし，題意より $0 < \alpha < \dfrac{1}{4}$ なので，独立であるための必要条件を常にみたさないことがわかる．よって，X と Y は従属．なお，$x = 0,\, y = 0$ として，(1) を調べると簡単に従属性がわかる．

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

$$f_{X,Y}(x,\, y) = \begin{cases} 8xy & (0 < x < y < 1), \\ 0 & （その他） \end{cases}$$

を持つとする．

(a) $f_{X,Y}(x,\, y) > 0$ となる領域を図示した上で X と Y の周辺確率密度関数 $f_X(x)$ と $f_Y(y)$ を求めよ．なお，解答には $f_X(x) > 0$ と $f_Y(y) > 0$ となる範囲を明示すること．

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

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

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

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,\, y)\, dy \\ &= \int_{-\infty}^{x} f_{X,Y}(x,\, y)\, dy + \int_{x}^{1} f_{X,Y}(x,\, y)\, dy + \int_{1}^{\infty} f_{X,Y}(x,\, y)\, dy \\ &= \int_{x}^{1} 8xy\, dy = \left[4xy^2\right]_x^1 = 4x(1 - x^2) \end{aligned}$$

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

$$f_X(x) = \begin{cases} 0 & (x \leq 0) \\ 4x(1 - x^2) & (0 < x < 1) \\ 0 & (x \geq 1) \end{cases}$$

となる．

$0 < y < 1$ に対して

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,\, y)\, dx \\ &= \int_{-\infty}^{0} f_{X,Y}(x,\, y)\, dx + \int_{0}^{y} f_{X,Y}(x,\, y)\, dx + \int_{y}^{\infty} f_{X,Y}(x,\, y)\, dx \\ &= \int_{0}^{y} 8xy\, dy = \left[4x^2 y\right]_0^y = 4y^3 \end{aligned}$$

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

$$f_Y(y) = \begin{cases} 0 & (y \leq 0) \\ 4y^3 & (0 < y < 1) \\ 0 & (y \geq 1) \end{cases}$$

となる．

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

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

となる．

(c) $0<y<1$ に対して $Y=y$ が与えられたときの X の条件付期待値は定義され，

$$\begin{aligned}\mathbb{E}(X|y) &= \int_{-\infty}^{\infty} x f_{X|Y}(x|y)\,dx \\ &= \int_{-\infty}^{0} x f_{X|Y}(x|y)\,dx + \int_{0}^{y} x f_{X|Y}(x|y)\,dx + \int_{y}^{\infty} x f_{X|Y}(x|y)\,dx \\ &= \int_{0}^{y} x f_{X|Y}(x|y)\,dx = \left[\frac{2x^3}{3y^2}\right]_0^y = \frac{2y}{3}\end{aligned}$$

$\boxed{\text{問題 5}}$　Ω を標本空間とし，$\mathcal{F}$ を Ω 上の完全加法族[2]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，$\mathbb{P}(A) \geq 0$

(P2) $\mathbb{P}(\Omega) = 1$

(P3) $i = 1, 2, \ldots$ に対して $A_i \in \mathcal{F}$ かつ $A_i \cap A_j = \emptyset\,(i \neq j)$ 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：$B_1, B_2 \in \mathcal{F}$ に対して

$$B_1 \cap B_2 = \emptyset \quad \text{ならば，} \qquad \mathbb{P}(B_1 \cup B_2) = \mathbb{P}(B_1) + \mathbb{P}(B_2) \tag{2}$$

となる．

(b) (P1)–(P3) および (3)（必要ならば）を用いてつぎのことを示せ：$B_3, B_4 \in \mathcal{F}$ に対して

$$\mathbb{P}(B_3 \cup B_4) \leq \mathbb{P}(B_3) + \mathbb{P}(B_4)$$

となる．

$\boxed{\text{例}}$　たとえば，

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

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

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

と

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

となることに注意すれば，

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

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，

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

[2] 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．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．

解答 (a) (P3) において $A_1 = B_1$, $A_2 = B_2$, $A_i = \emptyset\,(i \geq 3)$ とおけば，(P3) の仮定をみたすので，

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

がわかる．最後の等号は (3) からわかる．
(b) $B_3 \cup B_4 = (B_3 \cap B_4^c) \cup B_4$ と $(B_3 \cap B_4^c) \cap B_4 = \emptyset$ なることから $B_1 = B_3 \cap B_4^c$ と $B_2 = B_4$ として (a) を用いると

$$\mathbb{P}(B_3 \cup B_4) = \mathbb{P}(B_3 \cap B_4^c) + \mathbb{P}(B_4) \tag{7}$$

となる．さらに，$B_3 = (B_3 \cap B_4) \cup (B_3 \cap B_4^c)$ と $(B_3 \cap B_4) \cap (B_3 \cap B_4^c) = \emptyset$ となるので，$B_1 = B_3 \cap B_4$ と $B_2 = B_3 \cap B_4^c$ として (a) を用いると

$$\mathbb{P}(B_3) = \mathbb{P}(B_3 \cap B_4) + \mathbb{P}(B_3 \cap B_4^c)$$

を得る．これを (7) に代入すれば

$$\mathbb{P}(B_3 \cup B_4) = \mathbb{P}(B_3) + \mathbb{P}(B_4) - \mathbb{P}(B_3 \cap B_4) \leq \mathbb{P}(B_3) + \mathbb{P}(B_4)$$

を得る．最後の不等号は (P1) からわかある．

配点

問題 1 25 点 (a) 5 (b) 10 (c) 5 (d) 5　問題 2 20 点 (a) 5 (b) 10 (c) 5
問題 3 25 点 (a) 3+2 (b) 5 (c) 2+2+1 (d) 5 (e) 5　問題 4 20 点 (a) 10 (b) 5 (c) 5
問題 5 10 点 (a) 5 (b) 5

得点	$0 \sim 20$	$21 \sim 54$	$55 \sim 74$	$75 \sim 89$	$90 \sim 100$
成績	D	C	B	A	A^+