今野　(2007 年 7 月 18 日)

統計解析・演習の試験問題（試験時間は 120 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．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　離散型確率変数 X の確率関数

$$f_X(x|\,\theta) = \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|}, \qquad x = -1, 0, 1$$

とする．ただし，$0 < \theta < 1$ とする．確率変数

$$S = S(X) = |X|$$

を考える．

(1) 確率 S の確率関数 $f_S(s)$ と分布関数 $F_S(s) = \mathbb{P}(S \le s)\,(s \in \mathbb{R})$ を求めよ[1]．ただし，$F_S(s)$ は $\mathbb{R}$ 上の関数であることがわかるように答えよ．

(2) 確率 S の期待値 $\mathbb{E}[S]$ と分散 $\mathbb{VAR}[S]$ を求めよ．

(3) $S = s$ が与えられたときの X の条件付確率分布を求めよ[2]．ただし，$s = 0, 1$ である．

解答と配点　(1) 10 (2) 10 (3) 5 計 25

(1) S の取り得る値は 0, 1 である．

$$\{S = 1\} \Longleftrightarrow \{X = 1\} \cup \{X = -1\}$$

と事象 $\{X = 1\}$ と $\{X = -1\}$ は排反なので，

$$\mathbb{P}(S=1) = \mathbb{P}(\{X=1\} \cup \{X=-1\}) = \mathbb{P}(X=1) + \mathbb{P}(X=-1) = \left(\frac{\theta}{2}\right)^{|-1|} (1-\theta)^{1-|-1|} + \left(\frac{\theta}{2}\right)^{|1|} (1-\theta)^{1-|1|} = \theta$$

つぎに，

$$\{S = 0\} \Longleftrightarrow \{X = 0\}$$

より

$$\mathbb{P}(S=0) = \mathbb{P}(X=0) = \left(\frac{\theta}{2}\right)^{|0|} (1-\theta)^{1-|0|} = 1-\theta$$

[1] 確率関数の代わりに確率分布表で答えてもよい．ここで，確率変数 S の取りうる範囲に注意せよ．

[2] 条件付き確率関数を求めるか，S が与えられた値ごとの確率分布表を求めればよい．

したがって，

$$f_S(s) = \mathbb{P}(S=s) = \begin{cases} 1-\theta & (s=0) \\ \theta & (s=1) \\ 0 & (\text{その他}) \end{cases}$$

となる．

分布関数は

$$F_S(s) = \mathbb{P}(S \le s) = \begin{cases} 0 & (s<0) \\ 1-\theta & (0 \le s < 1) \\ 1 & (s \ge 1) \end{cases}$$

(2) S の期待値は

$$\mathbb{E}[S] \quad = \quad \sum_{s=0,\,1} s f_S(s) = 0 \times (1-\theta) + 1 \times \theta = \theta$$

[別解]

$$\mathbb{E}[S] = \mathbb{E}[|X|] \quad = \quad \sum_{x=-1,\,0,\,1} |x| f_X(x) = |-1| \times f_X(-1) + 0 \times f_X(0) + |1| \times f_X(1) = \frac{\theta}{2} + \frac{\theta}{2} = \theta.$$

S の分散を求めるために，

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

から

$$\mathbb{E}[S^2] \quad = \quad \sum_{s=0,\,1} s^2 f_S(s) = 0^2 \times (1-\theta) + 1^2 \times \theta = \theta$$

[別解]

$$\mathbb{E}[S^2] = \mathbb{E}[|X|^2] \quad = \quad \sum_{x=-1,\,0,\,1} |x|^2 f_X(x) = |-1|^2 \times f_X(-1) + 0^2 \times f_X(0) + |1|^2 \times f_X(1) = \frac{\theta}{2} + \frac{\theta}{2} = \theta.$$

よって，

$$\mathbb{VAR}[S] = \mathbb{E}[S^2] - \{\mathbb{E}[S]\}^2 = \theta - \theta^2$$

(3) $S=0$ のとき，

$$\mathbb{P}(X=0|\,S=0) = \frac{\mathbb{P}(\{X=0\} \cap \{S=0\})}{\mathbb{P}(S=0)} = \frac{\mathbb{P}(X=0)}{\mathbb{P}(S=0)} = \frac{\left(\frac{\theta}{2}\right)^{|0|}(1-\theta)^{1-|0|}}{1-\theta} = 1$$

よって，$\{S=0\} \Longleftrightarrow \{X=0\}$ なので，

$$f_{X|\,S}(x|S=0) = \mathbb{P}(X=x|\,S=0) = \begin{cases} 1 & (x=0) \\ 0 & (\text{その他}) \end{cases}$$

$S=1$ のとき，

$$\mathbb{P}(X=1|\,S=1) = \frac{\mathbb{P}(\{X=1\} \cap \{S=1\})}{\mathbb{P}(S=1)} = \frac{\mathbb{P}(X=1)}{\mathbb{P}(S=1)} = \frac{\left(\frac{\theta}{2}\right)^{|1|}(1-\theta)^{1-|1|}}{\theta} = \frac{1}{2}$$

と

$$\mathbb{P}(X=-1|\,S=1) = \frac{\mathbb{P}(\{X=-1\} \cap \{S=1\})}{\mathbb{P}(S=1)} = \frac{\mathbb{P}(X=-1)}{\mathbb{P}(S=1)} = \frac{\left(\frac{\theta}{2}\right)^{|-1|}(1-\theta)^{1-|-1|}}{\theta} = \frac{1}{2}$$

となる．$\{S=1\} \Longleftrightarrow \{X=1\} \cup \{X=-1\}$ より，

$$f_{X|\,S}(x|S=1) = \mathbb{P}(X=x|\,S=1) = \begin{cases} \dfrac{1}{2} & (x=-1,\,1) \\ 0 & (\text{その他}) \end{cases}$$

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

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

をもつとする．このとき，以下の問いに答えよ．

(1) 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．

(2) $Y = -2\log X + 2\log 2$ 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．

(3) Y の期待値 $\mathbb{E}[Y]$ を求めよ．

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

解答と配点　(1) 10 (2) 10 (3) 5 計 25

(1) $x \in (0, 2)$ に対して，

$$F_X(x) = \int_0^x \frac{1}{2}\, dx = \frac{x}{2}$$

となる．明らかに，$x \le 0$ に対しては $F_X(x) = 0$，$x \ge 2$ に対しては $F_X(x) = 1$ となる．したがって，

$$F_X(x) = \begin{cases} 0 & (x \le 0), \\ \dfrac{x}{2} & (0 < x < 2), \\ 1 & (x \ge 0) \end{cases}$$

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

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

となる．

(2) $\mathbb{P}(0 < X < 2) = 1$ 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

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

となる．X は連続型確率変数なので，$\mathbb{P}(X = 2e^{-y/2}) = 0$ と確率の性質（事象 B に対して，$\mathbb{P}(B^c) = 1 - \mathbb{P}(B)$）より

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

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，

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

(3) Y の確率密度関数の台から Y の期待値は

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

となる．

問題 3　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．

(1) X の周辺確率密度関数 $f_X(x)$ を求め，積分 $\int_{-\infty}^{\infty} f_X(x)\,dx$ の値を計算せよ．

(2) $X = x\,(0 < x < 1)$ が与えたれたときの Y の条件付確率密度関数 $f_{Y|X}(y|\,x)$ と条件付期待値 $\mathbb{E}[Y|\,X = x]$ を求めよ．

(3) $U = X + Y$ 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．

(4) $f_{U,V}(u,\,v) > 0$ となる u と v の範囲に注意して，V の周辺確率密度関数 $f_V(v)$ を求めよ．さらに，積分 $\int_{-\infty}^{\infty} f_V(v)\,dv$ の値を計算せよ．

$\boxed{\text{解答と配点}}$ (1) 5 (2) 10 (3) 5 (4) 5 計 25

(1) $0 < x < 1$ 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．

　また，

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_0^1 \left(x + \frac{1}{2}\right)\,dx = \left[\frac{x^2}{2} + \frac{x}{2}\right]_0^1 = 1.$$

(2) $(0,\,1) = \{x \in \mathbb{R} :\, f_X(x) > 0\}$ 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．

　また，$0 < x < 1$ 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，

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

となる．

(3) 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．
(4) 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．
　また，

$$\int_{-\infty}^{\infty} f_V(v)\,dv=\int_{-1}^{0}(1+v)\,dv+\int_0^1(1-v)\,dv=\left[v+\frac{v^2}{2}\right]_{-1}^{0}+\left[v-\frac{v^2}{2}\right]_0^1=\left(1-\frac{1}{2}+1-\frac{1}{2}\right)=1.$$

問題 4　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	α	β	α
0	β	0	β
1	α	β	α

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

(1) 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．

(2) X の周辺分布関数 $F_X(x)=\mathbb{P}(X\le x)\,(x\in\mathbb{R})$ を求め，$F_X(x)$ のグラフを作図せよ．ただし，答えのみでよいが，$F_X(x)$ は $\mathbb{R}$ 上の関数であることがわかるように答えよ．

(3) $X,\,Y,\,XY$ の期待値 $\mathbb{E}(X),\,\mathbb{E}(Y),\,\mathbb{E}(XY)$ を計算せよ．

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

(5) X と Y は確率的に独立か従属かを調べよ．ただし，その理由を述べること．

解答と配点　(1) 5 (2) 5 (3) 5 (4) 5 (5) 5 計 25
(1)

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	$2\alpha+\beta$

(2)

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

(3) 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.
(4) したがって,

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

となる.
(5) 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

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

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.

解答と配点

成績について

得点	0 ∼ 14	15 ∼ 39	40 ∼ 64	65 ∼ 79	80 ∼ 100
成績	D	C	B	A	A^+