今野 (2010 年 7 月 28 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．

問題 1　連続型確率変数 X は確率密度関数

$$f_X(x) = \begin{cases} \dfrac{x^2}{9} & 0 < x < 3; \\ 0 & \text{その他} \end{cases}$$

を持つとする．このとき，以下の問いに答えよ．

(1) 確率 $\mathbb{P}(1 < X \leq 2)$ を求めよ．

(2) 確率変数 X の累積分布関数 $F_X(x) = \mathbb{P}(X \leq x)\,(x \in \mathbb{R})$ を求めよ．

(3) 確率変数 X の期待値 $\mathbb{E}[X]$ を求めよ．

(4) 確率変数 X の分散 $\mathbb{VAR}[X]$ を求めよ．

解答と配点　(1) 5 点 (2) 10 点 (3) 5 点 (4) 10 点　計 30 点

(1) 確率密度関数はつぎをみたす関数である：任意の $a, b \in \mathbb{R}(a < b)$ に対して，

$$\mathbb{P}(a < X \leq b) = \int_a^b f_X(x)\,dx$$

である．したがって，

$$\mathbb{P}(1 < X \leq 2) = \int_1^2 f_X(x)\,dx = \int_1^2 \frac{x^2}{9}\,dx = \left[\frac{x^3}{27}\right]_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}.$$

(2) $x < 0$ のとき，$f_X(x) = 0$ より $F_X(x) = 0$．$0 \leq x < 3$ のとき，

$$F_X(x) = \int_{-\infty}^x f_X(t)\,dt = \int_0^x \frac{t^2}{9}\,dt = \frac{x^3}{27}.$$

$x \geq 3$ のとき，

$$F_X(x) = \int_{-\infty}^3 f_X(t)\,dt + \int_3^x f_X(t)\,dt = \int_0^3 \frac{t^2}{9}\,dt + \int_3^x 0\,dt = 1.$$

よって，

$$F_X(x) = \begin{cases} 0 & x < 0; \\ \dfrac{x^3}{27} & 0 \leq x < 3; \\ 1 & x \geq 3. \end{cases}$$

(3) $\mathbb{E}[X] = \displaystyle\int_{-\infty}^{\infty} x f_X(x)\,dx = \int_0^3 \frac{x^3}{9}\,dx = \left[\frac{x^4}{36}\right]_0^3 = \frac{81}{36} = \frac{9}{4}$.

(4) $\mathbb{E}[X^2]$ を求める：$\mathbb{E}[X^2] = \displaystyle\int_{-\infty}^{\infty} x^2 f_X(x)\,dx = \int_0^3 \frac{x^4}{9}\,dx = \left[\frac{x^5}{45}\right]_0^3 = \frac{3^5}{45} = \frac{27}{5}$．分散公式より

$$\mathbb{VAR}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2 = \frac{27}{5} - \left(\frac{9}{4}\right)^2 = \frac{27}{5} - \frac{81}{16} = \frac{27\times 16 - 81\times 5}{80}$$

$\boxed{\text{問題 2}}$　連続型確率変数 X, Y は同時確率密度関数

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

を持つとする．このとき，以下の問いに答えよ．

(1) 確率変数 X の周辺確率密度関数 $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,\,y)\,dy$ を求めよ．

(2) 確率変数 Y の周辺確率密度関数 $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,\,y)\,dx$ を求めよ．

(3) $Y = y\,(0<y<1)$ を与えたときの X の条件付確率密度関数 $f_{X|Y}(x|\,y)$ を求め，$\int_{-\infty}^{\infty} f_{X|Y}(x|\,y)\,dx = 1$ をみたすことを確認せよ．

(4) $\displaystyle\iint_{\mathbb{R}^2} f_{X,Y}(x,\,y)\,dxdy = 1$ を確認せよ．

$\boxed{\text{解答と配点}}$　(1) 5 点 (2) 5 点 (3) 10 点 (4) 10 点　計 30 点

(1) $0<x<1$ に対して，

$$f_X(x) = \int_0^x f_{X,Y}(x,\,y)\,dy = 8x\int_0^x y\,dy = 8x\left[\frac{y^2}{2}\right]_0^x = 4x^3.$$

したがって，

$$f_X(x) = \begin{cases} 4x^3 & 0<x<1; \\ 0 & (\text{その他}) \end{cases}$$

(2) $0<y<1$ に対して，

$$f_Y(y) = \int_y^1 f_{X,Y}(x,\,y)\,dy = 8y\int_y^1 x\,dx = 8y\left[\frac{x^2}{2}\right]_y^1 = 4y(1-y^2).$$

したがって，

$$f_Y(y) = \begin{cases} 4y(1-y^2) & 0<y<1; \\ 0 & (\text{その他}) \end{cases}$$

(3) $0<y<1$ に対して，

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

さらに，$0<y<1$ に対して，

$$\int_{-\infty}^{\infty} f_{X|Y}(x|\,y)\,dx = \int_y^1 \frac{2x}{1-y^2}\,dx = \frac{1}{1-y^2}\int_y^1 2x\,dx = \frac{1}{1-y^2}\left[x^2\right]_y^1 = \frac{1}{1-y^2}\times(1-y^2) = 1.$$

(4)

$$\iint_{\mathbb{R}^2} f_{X,Y}(x,\,y)\,dxdy = \int_0^1\left(\int_0^x f_{X,Y}(x,\,y)\,dy\right)dx = \int_0^1 f_X(x)\,dx = \int_0^1 3x^3\,dx = \left[x^4\right]_0^1 = 1.$$

$\boxed{\text{問題 3}}$ 離散型確率変数 $X_1, X_2, \ldots, X_n\,(n \geq 2)$ は独立同一の分布に従い，$i = 1, 2, \ldots, n$ に対して，

$$f_{X_i}(x) = \mathbb{P}(X_i = x) = \begin{cases} \dfrac{1}{2} & x = 1; \\ \dfrac{1}{3} & x = 2; \\ \dfrac{1}{6} & x = -1; \\ 0 & (\text{その他}) \end{cases}$$

とする．すわわち，f_{X_i} は X_i の確率関数である．このとき，以下の問いに答えよ．

(1) 期待値 $\mathbb{E}[X_i]$ を求めよ．ただし，$i = 1, 2, \ldots, n$ である．

(2) 期待値 $\mathbb{E}[X_1 + X_2 + \cdots + X_n]$ を求めよ．

(3) 分散 $\mathbb{VAR}[X_i]$ を求めよ．ただし，$i = 1, 2, \ldots, n$ である．

(4) $i = 1, \ldots, n, j = 1, \ldots, n$ とする．$i \neq j$ のとき，共分散 $\mathbb{COV}[X_i, X_j]$ を求めよ．

(5) 分散 $\mathbb{VAR}[X_1 + X_2 + \cdots + X_n]$ を求めよ．

(6) $\bar{X}_n = \dfrac{1}{n}(X_1 + X_2 + \cdots + X_n)$ としたとき，$\mathbb{VAR}[\bar{X}_n]$ を求めよ．

(7) $\sqrt{n}(\bar{X}_n - 1)$ の期待値 $\mathbb{E}[\sqrt{n}(\bar{X}_n - 1)]$ を求めよ．

(8) $\sqrt{n}(\bar{X}_n - 1)$ の分散 $\mathbb{VAR}[\sqrt{n}(\bar{X}_n - 1)]$ を求めよ．

$\boxed{\text{解答と配点}}$ (1) 5 点 (2) 5 点 (3) 5 点 (4) 5 点 (5) 5 点 (6) 5 点 (7) 5 点 (8) 5 点　計 40 点

(1)

$$\mathbb{E}[X_i] = \sum_{x=-1,1,2} x\mathbb{P}(X = x) = (-1) \times \frac{1}{6} + 1 \times \frac{1}{2} + 2 \times \frac{1}{3} = \frac{(-1) + 3 + 4}{6} = 1$$

(2) 期待値の性質より

$$\mathbb{E}[X_1 + X_2 + \cdots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n] = n$$

(3) まず，

$$\mathbb{E}[X_i^2] = \sum_{x=-1,1,2} x^2\mathbb{P}(X = x) = (-1)^2 \times \frac{1}{6} + 1^2 \times \frac{1}{2} + 2^2 \times \frac{1}{3} = \frac{1 + 3 + 8}{6} = 2$$

分散公式より

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

(4) X_i と X_j は独立なので，$\mathbb{COV}[X_i, X_j] = 0$.

(5) 分散の性質より

$$\mathbb{VAR}[X_1 + X_2 + \cdots + X_n] = \sum_{i=1}^{n} \mathbb{VAR}[X_i] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \mathbb{COV}[X_i, X_j].$$

しかし，$i \neq j$ のとき，X_i と X_j は独立なので，$\mathbb{COV}[X_i, X_j] = 0$ となるので，

$$\mathbb{VAR}[X_1 + X_2 + \cdots + X_n] = n$$

(6) 分散の性質より

$$\mathbb{VAR}[\bar{X}_n] = \mathbb{VAR}\left[\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right] = \frac{1}{n^2}\mathbb{VAR}[X_1 + X_2 + \cdots + X_n] = \frac{1}{n}$$

(7) 期待値の性質より

$$\begin{aligned} \mathbb{E}[\sqrt{n}(\bar{X}_n - 1)] &= \sqrt{n}(\mathbb{E}[\bar{X}_n] - 1) = \sqrt{n}\left(\mathbb{E}\left[\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right] - 1\right) \\ &= \sqrt{n}\left(\frac{1}{n}\mathbb{E}\left[(X_1 + X_2 + \cdots + X_n)\right] - 1\right) = \sqrt{n}\left(\frac{1}{n}n - 1\right) = 0. \end{aligned}$$

(8) 分散の性質より

$$\mathbb{VAR}[\sqrt{n}(\bar{X}_n - 1)] = (\sqrt{n})^2 \mathbb{VAR}[\bar{X}_n] = n \times \frac{1}{n} = 1.$$

ヒント

- X を連続型確率変数とし，$F_X(\cdot)$ をその分布関数とする．$\mathbb{R}$ 上の非負値関数 $f_X(\cdot)$ で任意の $x \in \mathbb{R}$ に対して $F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy$ をみたすものが存在するとき，$f_X(\cdot)$ を X の確率密度関数という．すると，

  (i) $f_X(x) \geq 0$;

  (ii) $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$;

  (iii) 任意の実数 $a, b \in \mathbb{R}(a < b)$ に対して，$\mathbb{P}(a < X \leq b) = \int_a^b f_X(x)\,dx$;

  をみたす．

- 確率変数 X は確率関数または密度関数 $f_X(\cdot)$ を持つとする．$\mathbb{R}$ 上のボレル可側関数[1] $g(\cdot)$ に対し，$g(X)$ の期待値を

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

  で定義する．ただし，離散型の場合は $\sum_x |g(x)| f_X(x) < \infty$ のとき，連続型の場合は $\int_{-\infty}^{\infty} |g(x)| f_X(x)\,dx < \infty$ のとき，$g(X)$ の期待値を定義することにする．

- (X, Y) は連続型確率ベクトルとし，同時確率密度関数 $f_{X,Y}(x, y)$ および周辺確率密度関数 $f_X(x)$ と $f_Y(y)$ を持つとする．このとき，$f_Y(y) > 0$ なる任意の y 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を $f_{X|Y}(x|y)$ で記し，

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

  で定める．

- つぎの期待値の性質は証明なしで用いてよい：

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

  (ii) $\mathbb{VAR}(aX_1 + bX_2) = a^2 \mathbb{VAR}(X_1) + b^2 \mathbb{VAR}(X_2) + 2ab\,\mathbb{COV}(X_1, X_2)$.

  (iii) $\mathbb{COV}(X_1, X_2) = \mathbb{E}[(X_1 - \mathbb{E}(X_1))(X_2 - \mathbb{E}(X_2))] = \mathbb{E}[X_1 X_2] - \mathbb{E}[X_1]\mathbb{E}[X_2]$.

  (iv) X_1, X_2 が独立のとき，$\mathbb{E}(X_1 X_2) = \mathbb{E}(X_1)\mathbb{E}(X_2)$.

  (v) X が非負値確率変数[2]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，$\mathbb{E}(c) = c$.

  (vii) $\mathbb{VAR}(X_1) = \mathbb{E}[(X_1 - \mathbb{E}(X_1))^2] = \mathbb{E}[X_1^2] - \{\mathbb{E}[X_1]\}^2$.

  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．

[1]任意のボレル集合 $B \in \mathcal{B}(\mathbb{R})$ に対し，$\{x \in \mathbb{R} : g(x) \in B\} \in \mathcal{B}(\mathbb{R})$ が成り立つような関数である．
[2]$\mathbb{P}(X \geq 0) = 1$ 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．

<u>成績について</u>

得点	$0 \sim 14$	$15 \sim 39$	$40 \sim 59$	$60 \sim 79$	$80 \sim 100$
成績	D	C	B	A	A^+

<u>得点分布</u>　平均点 $= 37.9$，中央値 $= 40$、標準偏差 $= 24.3$

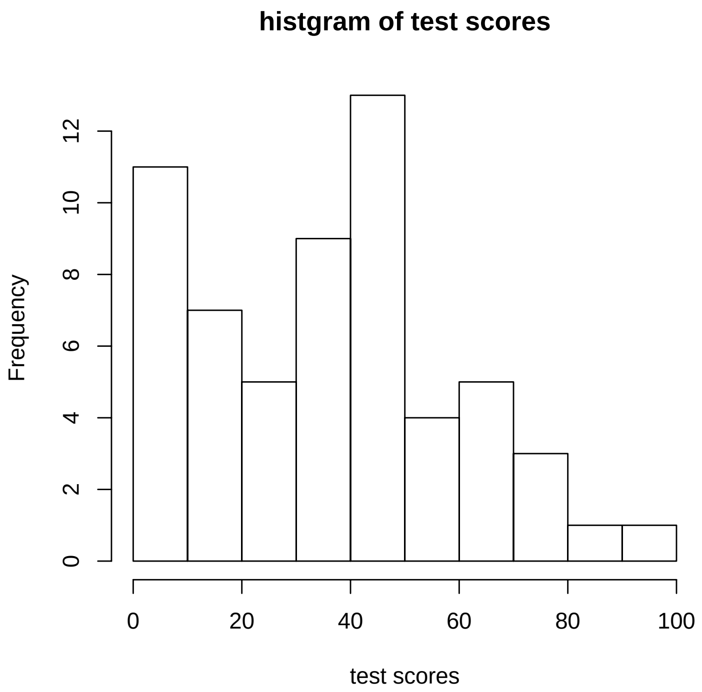

Figure 1: This is a figure