2008 年 01 月 11 日（今野）

確率統計と情報処理・演習（金曜後期 5–8 限，2 単位 + 1 単位）

試験について

- 実施日：掲示を参照.
- 場所：掲示を参照.
- 持ち込めるもの：筆記用具のみ.
- 答案の返却日時：2008 年度統計解析の第 1 回目（受講しないものは第 1 回目の講義後にに研究室に取りに来てください）.

講義内容要点

- 1 変量のデータの分析
  - ヒストグラム
  - 平均と中央値
  - 箱ひげ図
  - ばらつきの尺度 — 分散と標準偏差など.
- 2 変量のデータの分析
  - 散布図について
  - 回帰直線と最小 2 乗法について
  - 相関係数について
- 確率の定義と性質
- 確率関数とその分布
  - 離散型確率変数と確率関数
  - 連続型確率変数と確率密度関数
  - 累積分布関数
- 期待値と分散
  - 二項分布の定義とその性質
  - ポアソン分布の定義とその性質
  - 正規分布の定義とその性質
  - 一様分布の定義
  - 指数分布の定義とその性質
- 標本分布について
  - 有限母集団の標本分布

  - カイ自乗分布の確率密度関数
  - 自由度 1 のカイ自乗分布と標準正規分布の関係
  - カイ自乗分布の再帰性について
- カイ自乗分布と正規母集団からの標本との関係について
- t 分布の定義とその確率密度関数について
- 独立な標準正規分布とカイ自乗分布から t 分布の導出について
- F 分布の定義とその確率密度関数について
- 独立なふたつのカイ自乗分布から F 分布の導出について
- 一様分布と中心極点定理
- さまざまな分布と中心極限定理
- 母集団と標本
- 母数（母平均・母分散）
- 推定量と推定値
- 点推定法：推定量の性質について
  - 不偏推定量
  - 有効推定量：最良線形不偏推定量
- 区間推定法
  - 信頼係数と信頼区間（ここからは来週）
  - 正規母集団に対する母平均の区間推定
    * 母分散が既知の場合の母平均の区間推定
    * 母分散が未知の場合の母平均の区間推定

今野　(2007 年 1 月 12 日)

確率統計と情報処理・演習の試験問題（試験時間は 90 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．

答案は来年度開講の「統計解析」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　関数 $\mathbb{P}(\cdot)$ で，つぎの条件をみたすものを考える[1]：

(P1) すべての事象 A に対して，$0 \le \mathbb{P}(A) \le 1$

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

(P3) 事象 A と B が互いに排反のとき，$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ が成立

このとき，以下の問いに答えよ．ただし，A と B は事象とし，A^c は A の補事象である．

(1) 事象 A に対し，$\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$ を示せ．

(2) $\mathbb{P}(A \cup B) \le \mathbb{P}(A) + \mathbb{P}(B)$

解答と配点　(1) 10 (2) 10 計 20

(1) 補事象の定義から $A \cap A^c = \emptyset$ となるので，A と A^c は互いに排反．さらに，$A \cup A^c = \Omega$ より

$$\begin{aligned} 1 &= \mathbb{P}(\Omega) && \text{（確率の性質 (P2) より）} \\ &= \mathbb{P}(A \cup A^c) = \mathbb{P}(A) + \mathbb{P}(A^c) && \text{（確率の性質 (P3) より）} \end{aligned}$$

(2) $A \cap B^c$, $A \cap B$, $A^c \cap B$ は互いに排反で，$(A \cap B^c) \cup (A \cap B) \cup (A^c \cap B) = A \cup B$ に注意して，P3 を用いれば，

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

しかし，確率の性質 (P1) より $\mathbb{P}(A \cap B) \ge 0$ なので，

$$\mathbb{P}(A \cup B) \le \mathbb{P}(A) + \mathbb{P}(B)$$

となる．

問題 2　「2 人の子供がいる家庭について，子供の男女の性別を調べる」という試行を考える．女児を g，男児を b で表す．たとえば，生まれた順が，女，男ならば，gb と書く．標本空間は

$$\Omega = \{gg, gb, bg, bb\}$$

となる．男女の出生比率は $\frac{1}{2}$ とし，

$$\mathbb{P}(gg) = \mathbb{P}(gb) = \mathbb{P}(bg) = \mathbb{P}(bb) = \frac{1}{4}$$

[1] $\mathbb{P}$ を教科書では Pr と書いている．

とする．

いま，女子の人数 X だけに注目する．標本点 ω が与えられれば，X の値を標本点の女子の人数に対応させる，その意味で X は ω の関数 $X = X(\omega)$ となる．このとき，以下の問いに答えよ．

(1) $\mathbb{P}(X = 0)$ を求めよ．

(2) $\mathbb{P}(X = 1)$ を求めよ．ただし，答えのみでなく確率の条件をどのように使い求めたかを明記せよ．

(3) 確率変数 X の確率関数 $f_X(x)$ を求めよ．

(4) 確率変数 X の累積分布確率関数 $F_X(x)$ を求め，そのグラフを作図せよ．

解答と配点　(1) 5 (2) 5 (3) 5 (4) 5 計 20

(1) $\mathbb{P}(X = 0) = \mathbb{P}(bb) = \frac{1}{4}$

(2) 事象 $\{gb\}$ と $\{bg\}$ は排反であるので，

$$\mathbb{P}(X = 1) = \mathbb{P}(\{gb\} \cup \{bg\}) = \mathbb{P}(\{gb\}) + \mathbb{P}(\{bg\}) = \frac{1}{2}$$

(3)

$$f_X(x) = \begin{cases} \dfrac{1}{4} & (x = 0) \\[2ex] \dfrac{1}{2} & (x = 1) \\[2ex] \dfrac{1}{4} & (x = 2) \end{cases}$$

(4)

$$F_X(x) = \begin{cases} 0 & (x < 0) \\ \dfrac{1}{4} & (0 \le x < 1) \\[2ex] \dfrac{3}{4} & (1 \le x < 2) \\[2ex] 1 & (x \ge 2) \end{cases}$$

問題 3　$n \ge 2$ を自然数とし，X を離散型確率変数とする．X は有限の点 $x_1, x_2, \ldots, x_n$ において正の確率を持つとする．すなわち，$\mathbb{P}(X = x_i) > 0\,(i = 1, 2, \ldots, n)$ である．このとき，X の期待値と分散は

$$\begin{aligned} \mu &:= \mathbb{E}[X] = \sum_{i=1}^{n} x_i f_X(x_i) \\ \mathbb{E}[aX + b] &= \sum_{i=1}^{n} (a x_i + b) f_X(x_i) \\ \mathbb{VAR}[X] &= \sum_{i=1}^{n} (x_i - \mu)^2 f_X(x_i) \\ \mathbb{VAR}[aX + b] &= \sum_{i=1}^{n} (a x_i + b - \mathbb{E}[aX + b])^2 f_X(x_i) \\ \mathbb{E}[X^2] &= \sum_{i=1}^{n} x_i^2 f_X(x_i) \end{aligned}$$

である．ただし，a と b を定数とした．このとき，以下のことを示せ．

(1) $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$

(2) $\mathbb{VAR}[aX + b] = a^2 \mathbb{VAR}[X]$

(3) $\mathbb{VAR}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2$

ただし，確率関数 f_X の性質を利用してよい：

(i) $\mathbb{P}(X = x_i) = f_X(x_i) > 0, \qquad i = 1,\, 2,\, \ldots,\, n,$

(ii) $\sum_{i=1}^{n} f_X(x_i) = 1$

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

(1) $\sum_{i=1}^{n} f_X(x_i) = 1$ より

$$\mathbb{E}[aX + b] \;=\; \sum_{i=1}^{n}(ax_i + b) f_X(x_i) = a^sum_{i=1}^{n} x_i f_X(x_i) + b\sum_{i=1}^{n} f_X(x_i) = a\mathbb{E}[X] + b$$

(2)

$$\begin{aligned}
\mathbb{VAR}[aX + b] &= \sum_{i=1}^{n}(ax_i + b - \mathbb{E}[aX + b])^2 f_X(x_i) \\
&= \sum_{i=1}^{n}(ax_i + b - a\mathbb{E}[X] - b)^2 f_X(x_i) \\
&= \sum_{i=1}^{n}(ax_i - a\mathbb{E}[X])^2 f_X(x_i) \\
&= a^2\sum_{i=1}^{n}(x_i - \mathbb{E}[X])^2 f_X(x_i) = a^2\mathbb{VAR}[X]
\end{aligned}$$

(3)

$$\begin{aligned}
\mathbb{VAR}[X] &= \sum_{i=1}^{n}(x_i - \mu) f_X(x_i) \\
&= \sum_{i=1}^{n}(x_i^2 - 2\mu x_i + \mu^2) f_X(x_i) \\
&= \sum_{i=1}^{n} x_i^2 f_X(x_i) - 2\mu\sum_{i=1}^{n} x_i f_X(x_i) + \mu^2\sum_{i=1}^{n} f_X(x_i) \\
&= \mathbb{E}[X^2] - 2\mu \times \mu + \mu^2 = \mathbb{E}[X^2] - \mu^2 \\
&= \mathbb{E}[X^2] - \{\mathbb{E}[X]\}
\end{aligned}$$

$\boxed{\text{問題 4}}$ 大小ふたつのサイコロを投げる場合には，標本空間は

$$\Omega = \{(i,\, j) :\, i,\, j = 1,\, 2,\, \ldots,\, 6\}$$

となる．ただし，i は大きいサイコロの目，j は小さいサイコロの目とする．このとき，

$$\mathbb{P}\{(i,\, j)\} = \frac{1}{36}, \qquad i,\, j = 1,\, 2,\, \ldots,\, 6$$

となる．確率変数 X を

$$X((i,\, j)) = i + j$$

で定義する．このとき，以下の問いに答えよ．ただし，(1) と (4) のみ途中経過を記述しなくともよい．

(1) $\mathbb{P}(X = x) > 0$ となる x の値をすべてもとめよ．

(2) $\mathbb{P}(X = 3)$ の確率を求めよ．

(3) $\mathbb{P}(X = 6)$ の確率を求めよ．

(4) X の確率分布表を完成させよ．

$\boxed{\text{解答と配点}}$ (1) 5 (2) 5 (3) 5 (4) 5 計 20

(1) 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

(2) $\mathbb{P}(X=3)$ の確率を求めよ．

$$\{i+j=3\}=\{(1,\,2)\}\cup\{(2,\,1)\}$$

に注意して，

$$\mathbb{P}(i+j=3)=\mathbb{P}(\{(1,\,2)\}\cup\{(2,\,1)\})=\mathbb{P}(\{1,\,2\})+\mathbb{P}(\{2,\,1\})=\frac{1}{36}+\frac{1}{36}=\frac{2}{36}$$

となる．

(3)

$$\{i+j=6\}=\{(1,\,5)\}\cup\{(2,\,4)\}\cup\{(3,\,3)\}\cup\{(4,\,2)\}\cup\{(5,\,1)\}$$

に注意して，

$$\begin{aligned}
\mathbb{P}(i+j=6) &= \mathbb{P}(\{(1,\,5)\}\cup\{(2,\,4)\}\cup\{(3,\,3)\}\cup\{(4,\,2)\}\cup\{(5,\,1)\}) \\
&= \mathbb{P}(\{1,\,5\})+\mathbb{P}(\{2,\,4\})++\mathbb{P}(\{3,\,3\})++\mathbb{P}(\{4,\,2\})++\mathbb{P}(\{5,\,1\}) \\
&= \frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36}+\frac{1}{36} \\
&= \frac{6}{36}
\end{aligned}$$

となる．

(4)

X の値	2	3	4	5	6	7	8	9	10	11	12	合計
確率	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	1

問題 5　確率変数 Z が標準正規分布 $N(0,\,1)$ に従うとする．すわわち，

$$\mathbb{P}(a<Z\le b)=\int_a^b \frac{1}{\sqrt{2\pi}}e^{-z^2/2}\,dz, \qquad (\text{任意の }-\infty<a<b<\infty\ )$$

このとき，

$$Y=Z^2$$

の確率密度関数が

$$f_Y(y)=\begin{cases} \dfrac{1}{\sqrt{2}\sqrt{\pi}}\dfrac{1}{\sqrt{y}}e^{-y/2} & (y>0) \\ 0 & (\text{その他}) \end{cases}$$

で与えられることを示せ．

解答と配点　計 10 点

Y の確率密度関数 $f_Y(y)$ を求めるために，$F_Y(y)$ を y について微分する：

$$\begin{aligned}
f_Y(y) &= \frac{d}{dy}F_Y(y)=\frac{d}{dy}\int_{-\sqrt{y}}^{\sqrt{y}}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\,dz \\
&= \frac{d}{dy}\left\{\int_0^{\sqrt{y}}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\,dz-\int_0^{-\sqrt{y}}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}\,dz\right\} \\
&= \frac{1}{\sqrt{2\pi}}e^{-y/2}\frac{d}{dy}\sqrt{y}-\frac{1}{\sqrt{2\pi}}e^{-y/2}\frac{d}{dy}(-\sqrt{y}) \\
&= \frac{1}{\sqrt{2\pi}}e^{-y/2}\frac{1}{2\sqrt{y}}-\frac{1}{\sqrt{2\pi}}e^{-y/2}\frac{-1}{2\sqrt{y}} \\
&= \frac{1}{\sqrt{2}\sqrt{\pi}}\frac{1}{\sqrt{y}}e^{-y/2} \\
&= \frac{1}{2^{1/2}\Gamma(1/2)}y^{\frac{1}{2}-1}e^{-y/2}
\end{aligned}$$

注意：

$\star$ 2 番目の等号は定積分の性質を使っている！

$$\begin{aligned}\int_a^b f(x)\,dx &= \int_a^c f(x)\,dx + \int_c^b f(x)\,dx, \\ \int_a^b f(x)\,dx &= -\int_b^a f(x)\,dx\end{aligned}$$

$\star$ 3 番目の等号は微積分の基本的定理

$$\frac{d}{dx}\int_0^x g(z)\,dz = g(x)$$

と合成関数の微分を使っている！

$\boxed{\text{解答と配点}}$

成績について

得点	0 ～ 44	45 ～ 59	60 ～ 79	80 ～ 89	90 ～ 100
成績	D	C	B	A	A^+

得点分布　平均点 = 69.2，中央値 = 69.0、標準偏差 = 13.6

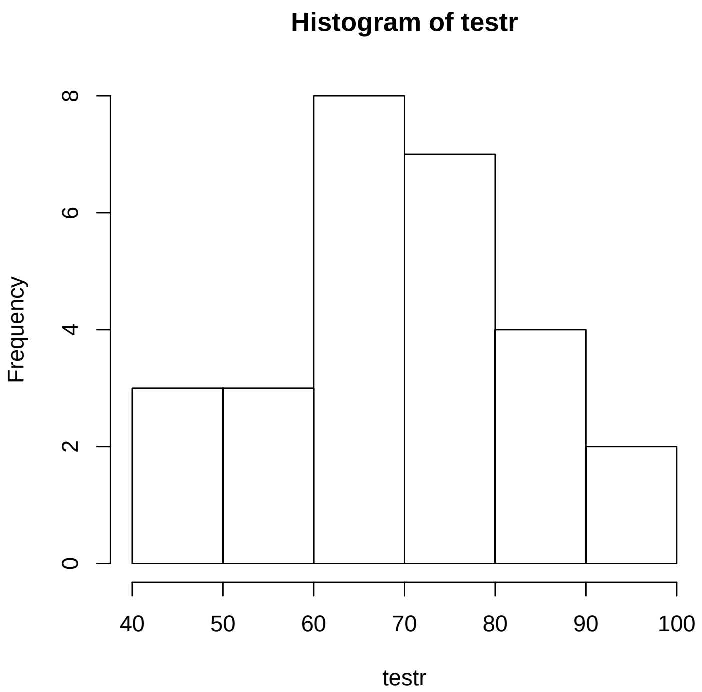

Figure 1: This is a figure