今野　(2011 年 01 月 07 日)

確率統計と情報処理・演習の試験問題（試験時間は 80 分）: 第一部

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　$n \geq 2$ とする．n 個のデータの値を $x_1, x_2, \ldots, x_n$ とする．このとき，以下の問いに答えよ．

(1) 実数 $a \neq 0$ と b に対し，$y_i = ax_i + b\,(i = 1, 2, \ldots, n)$ とし，$\bar{y}_n = \sum_{i=1}^n y_i/n$, $\bar{x}_n = \sum_{i=1}^n x_i/n$ と定める．$\bar{y}_n$ を $\bar{x}_n$, a および b を用いて表せ．答えだけでなく，途中計算も書くこと．

(2) 実数 c に対して，

$$g(c) = \sum_{i=1}^n (x_i - c)^2,$$

と定める． c に関する $g(c)$ の導関数を $g'(c)$ としたとき，$g'(c) = 0$ を解け．ただし，解は $\bar{x}_n, \bar{y}_n, a$ および b を用いて表せ (すべてを用いる必要はない)．答えだけでなく，途中計算も書くこと．

(3) 不等式

$$g(c) \geq \sum_{i=1}^n (x_i - \bar{x}_n)^2$$

を示せ．

(4) $\sum_{i=1}^n (x_i - \bar{x}_n)^2$ と $\sum_{i=1}^n (y_i - \bar{y}_n)^2$ の間に成立する関係式を求め，それを示せ．

配点 (1) 5 (2) 5 (3) 5 (4) 10 計 25

解答

(1) $\sum_{i=1}^n x_i = n\bar{x}_n$ に注意して，

$$\bar{y}_n = \sum_{i=1}^n y_i \quad = \quad \frac{1}{n}\sum_{i=1}^n (ax_i + b) = \frac{1}{n}\left\{ a\sum_{i=1}^n x_i + b\sum_{i=1}^n 1 \right\} = \frac{1}{n}(an\bar{x}_n + nb) = a\bar{x}_n + b$$

(2)

$$\begin{aligned} g'(c) \quad &= \quad -\sum_{i=1}^n 2(x_i - c) = -2\sum_{i=1}^n x_I + 2nc = 0 \\ &\Longleftrightarrow \sum_{i=1}^n x_i = nc \Longleftrightarrow n\bar{x}_n = c \end{aligned}$$

よって，$c = \bar{x}_n$.

(3) $z = g(c)$ は c の 2 次関数で 2 次の係数が正なので，$c = \bar{x}_n$ で最大値をとるので，

$$g(c) \quad \geq \quad g(\bar{x}_n) = \sum_{i=1}^{n}(x_i - \bar{x}_n)^2$$

を得る．
(4) $\bar{y}_n = a\bar{x}_n + b$ に注意して，

$$\sum_{i=1}^{n}(y_i - \bar{y}_n)^2 \quad = \quad \sum_{i=1}^{n}(ax_i + b - (a\bar{x}_n + b))^2 = a^2\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$$

問題 2　連続型確率変数 X は閉区間 $[0, 2]$ 上の一様分布にしたがっているとする．すなわち，任意の実数 a, b に対し，

$$\mathbb{P}(a < X \leq b) = \int_a^b f_X(x)\,dx, \qquad f_X(x) = \begin{cases} \dfrac{1}{2} & (0 \leq x \leq 2), \\ 0 & (\text{その他}) \end{cases}$$

となる．このとき，以下の問いに答えよ．ただし，積分の計算においては，下のヒントを証明なしで用いてよい．

(1) X の期待値 $\mathbb{E}[X]$ を計算せよ．

(2) X の分散 $\mathbb{VAR}[X]$ を計算せよ．

(3) 分散 $\mathbb{VAR}[2X - 1]$ を計算せよ．

(4) 実数上で定義された関数である X の分布関数

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$$

を求め，$F_X(x)$ のグラフを作図せよ．

配点 (1) 5 (2) 5 (3) 5 (4) 10 計 25

ヒント：以下の積分の性質や公式は証明なしに用いてよい．

- X を連続型確率変数で，確率密度関数 $f_X(x)$ を持つとする．$\mathbb{R}$ 上で定義された実数値関数 $g(x)$ に対して，$g(X)$ の期待値を次で定める：

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$$

  ただし，右辺の積分は有限の値を取ると仮定する．
- $\mathbb{VAR}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$
- 実数 a, b, c に対して，$\int_a^c f_Z(z)\,dz = \int_a^b f_Z(z)\,dz + \int_b^c f_Z(z)\,dz$

解答
(1)

$$\mathbb{E} \quad = \quad \int_{-\infty}^{\infty} x f_X(x)\,dx = \int_0^2 x\frac{1}{2}\,dx = \left[\frac{x^2}{4}\right]_0^2 = 1$$

(2) 上の問いの結果に注意すると

$$\mathbb{VAR}[X] \quad = \quad \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[(X-1)^2] = \int_{-\infty}^{\infty}(x-1)^2 f_X(x)\,dx = \int_0^2 (x-1)^2\frac{1}{2}\,dx = \left[\frac{(x-1)^3}{6}\right]_0^2 = \frac{1}{3}$$

(3) 上の問いの結果と分散の性質より

$$\mathbb{VAR}[2X - 1] = 4\mathbb{VAR}[X] = \frac{4}{3}$$

確率変数 X は閉区間 $[0, 2]$ 上の一様分布にしたがっているとする．すなわち，任意の実数 a, b に対し，

$$\mathbb{P}(a < X \leq b) = \int_a^b f_X(x)\,dx, \qquad f_X(x) = \begin{cases} \dfrac{1}{2} & (0 \leq x \leq 2), \\ 0 & (\text{その他}) \end{cases}$$

(4) $x < 0$ のとき，$f_X(x) = 0$ より，

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = 0$$

$0 \leq x < 2$ のとき，

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = \int_{-\infty}^{0} f_X(x)\,dx + \int_{0}^{x} f_X(x)\,dx = \frac{1}{2}\int_{0}^{x} dx = \frac{x}{2}$$

$x \geq 2$ のとき，

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = \int_{-\infty}^{2} f_X(x)\,dx + \int_{2}^{x} f_X(x)\,dx = F_X(2) = 1$$

よって，

$$F_X(x) = \begin{cases} 0 & (x < 0) \\ \dfrac{x}{2} & (0 \leq x < 2) \\ 1 & (x \geq 2) \end{cases}$$

問題 3　離散型確率変数 X は母数 $\lambda\,(\lambda > 0)$ のポアソン分布に従うとする．すなわち，X の確率関数は

$$f_X(x) = \begin{cases} \dfrac{\lambda^x e^{-\lambda}}{x!} & x = 0, 1, \ldots \\ 0 & (\text{その他}) \end{cases}$$

このとき，次の期待値を $\sum$ の記号 (和はどこからどこまで取るかを明示すること) および λ を用いて表現せよ．
(1) X の期待値 $\mathbb{E}[X]$
(2) $(X - \lambda)^2$ の期待値 $\mathbb{E}[(X - \lambda)^2]$

配点 (1) 5 (2) 5 計 10

ヒント：以下は証明なしに用いてよい．

- X を離散型確率変数で，確率関数 $f_X(x)$ を持つとする．X の取りうる値の集合を

$$\{x \in \mathbb{R} :\, f_X(x) > 0\} = \{x_1, x_2, \ldots\}$$

と書く．$\mathbb{R}$ 上で定義された実数値関数 $g(x)$ に対して，$g(X)$ の期待値を次で定める：

$$\mathbb{E}[g(X)] = \sum_{i=1}^{\infty} g(x_i) f_X(x_i)$$

ただし，右辺の和は有限の値を取ると仮定する．

解答　(1)

$$\mathbb{E} = \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!}$$

(2)

$$\mathbb{E}[(X - \lambda)^2] = \sum_{x=0}^{\infty} (x - \lambda)^2 \frac{\lambda^x e^{-\lambda}}{x!}$$

今野　(2010 年 01 月 08 日)

確率統計と情報処理・演習の試験問題（試験時間は 40 分）: 第二部

問題 4　以下のプログラムを入力し，問いのプログラムを作成したものを書け．ただし，出力された数値の結果を書く必要はない．

以下のプログラムを入力（有限母集団の生成）

```
> n.sample<-round(runif(1,0,100))+1
> x<-rchisq(n.sample,7)*3
> x0<-round(x*10)/10
```

(1) オブジェクト x0 の母平均偏差を計算するためのプログラムを書け．さらに，計算結果を書け．

(2) オブジェクト x0 の母標準偏差を計算するためのプログラムを書け．さらに，計算結果を書け．ただし，sd と var を用いないで作成せよ．

ヒント：$x_1, x_2, \ldots, x_n$ なる値からなる n 個体をもつでなる有限母集団の母平均，母平均偏差，および母標準偏差は，それぞれ $\bar{x}_n := (1/n)\sum_{i=1}^n x_i$, $\sum_{i=1}^n (1/n)|x_i - \bar{x}_n|$, $\sqrt{(1/n)\sum_{i=1}^n (x_i - \bar{x}_n)^2}$ で定義される．

配点 (1) 10 (2) 5 計 15

解答

(1) `> (1/length(x0))*sum(abs(x0-mean(x0)))`

(2) `> sqrt(sum((x0-mean(x0))**2)/length(x0))`

問題 5　連続型確率変数 X は平均 5，分散 3^2 の正規分布 $N(5, 3^3)$ に従うとする．このとき，以下の確率を求めよ．答えは小数第 3 位を四捨五入せよ．

(1) $\mathbb{P}(X \leq 2)$

(2) $\mathbb{P}(-1 < X \leq 3)$

(3) $\mathbb{P}\left(\frac{X-5}{3} \leq 1.645\right)$

配点 (1) 5 (2) 10 (3) 10 計 25

解答

(1)

```
> y1<-pnorm(2,5,3)
> y1
[1] 0.1586553
> round(y1*100)/100
[1] 0.16
```

(2)

```
> y2<-pnorm(3,5,3)-pnorm(-1,5,3)
> y2
[1] 0.2297424
> round(y2*100)/100
[1] 0.23
```

(3) 正規分布の性質より $(X-5)/3$ は標準正規分布（平均が 0，分散が 1^2 の正規分布）に従うので，

```
> y3<-pnorm(1.645,0,1)
> y3
[1] 0.950015
> round(y3*100)/100
[1] 0.95
```

解答

<u>得点分布</u>　平均点 $= 42.2$，中央値 $= 43.5$、標準偏差 $= 20.2$

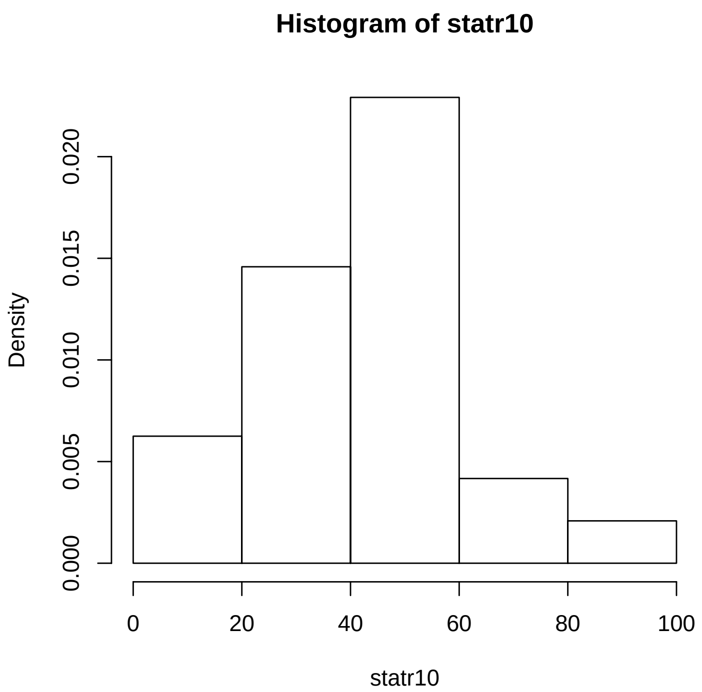

Figure 1: This is a figure

<u>成績について</u>

得点	$0 \sim 9$	$10 \sim 29$	$30 \sim 49$	$50 \sim 70$	$71 \sim 100$
成績	D	C	B	A	A^{+}