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 90 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　「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，

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

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．

ヒント：確率の公理

> 事象 A の関数 $\mathbb{P}(\cdot)$ で，つぎの条件をみたすものを考える：
>
> (P1) すべての事象 A に対して，$0 \leq \mathbb{P}(A) \leq 1$
>
> (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[a]．
>
> [a](P3) のことを有限加法性というが，本来は完全加法性で確率は定義される：すなわち，事象列 $A_1, A_2, \ldots$ が互いに排反のとき，$\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

配点 (1) 5 (2) 5 (3) 5 (4) 5 計 20

Solutions

(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)

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

(4)

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

問題 2 　X を連続型確率変数とし，確率密度関数 $f_X(x)$ をもつとする．このとき，X の期待値と分散は存在し，以下のように定義される：

$$\begin{aligned}\mu &:= \mathbb{E}[X]=\int_{-\infty}^{\infty}xf_X(x)\,dx;\\ \mathbb{E}[aX+b] &= \int_{-\infty}^{\infty}(ax+b)f_X(x)\,dx;\\ \mathbb{VAR}[X] &= \int_{-\infty}^{\infty}(x-\mu)^2f_X(x)\,dx;\\ \mathbb{VAR}[aX+b] &= \int_{-\infty}^{\infty}(ax+b-\mathbb{E}[aX+b])^2f_X(x)\,dx;\\ \mathbb{E}[X^2] &= \int_{-\infty}^{\infty}x^2f_X(x)\,dx;\end{aligned}$$

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

ヒント：確率密度関数 f_X の性質を利用してよい．

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

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

配点 　(1) 5 (2) 5 (3) 5 計 15

Solutions

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

$$\mathbb{E}[aX+b] = \int_{-\infty}^{\infty}(ax+b)f_X(x)\,dx=a\int_{-\infty}^{\infty}xf_X(x)\,dx+b\int_{-\infty}^{\infty}f_X(x)\,dx=a\mathbb{E}[X]+b$$

(2)

$$\begin{aligned}
\mathbb{VAR}[aX+b] &= \int_{-\infty}^{\infty} (ax+b-\mathbb{E}[aX+b])^2 f_X(x)\,dx \\
&= \int_{-\infty}^{\infty} (ax+b-a\mathbb{E}[X]-b)^2 f_X(x)\,dx \\
&= \int_{-\infty}^{\infty} (ax-a\mathbb{E}[X])^2 f_X(x)\,dx \\
&= a^2 \int_{-\infty}^{\infty} (x-\mathbb{E}[X])^2 f_X(x)\,dx = a^2\mathbb{VAR}[X]
\end{aligned}$$

(3)

$$\begin{aligned}
\mathbb{VAR}[X] &= \int_{-\infty}^{\infty} (x-\mu) f_X(x)\,dx \\
&= \int_{-\infty}^{\infty} (x^2-2\mu x+\mu^2) f_X(x)\,dx \\
&= \int_{-\infty}^{\infty} x^2 f_X(x)\,dx - 2\mu \int_{-\infty}^{\infty} x f_X(x)\,dx + \mu^2 \int_{-\infty}^{\infty} f_X(x)\,dx \\
&= \mathbb{E}[X^2] - 2\mu\times\mu + \mu^2 = \mathbb{E}[X^2] - \mu^2 \\
&= \mathbb{E}[X^2] - \{\mathbb{E}[X]\}
\end{aligned}$$

問題 3　確率変数 Z は標準正規分布に従うとする．すなわち，任意の実数 $a, b\,(a<b)$ に対して，

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

である．また，$\mathbb{P}(Z\le b) = \lim_{a\to-\infty} \mathbb{P}(a<Z\le b)$ と $\mathbb{P}(Z>a) = \lim_{b\to\infty} \mathbb{P}(a<Z\le b)$ とする．

(1) 任意の正の実数 c に対して，

$$\begin{aligned}
\mathbb{P}(Z\le c) &= 1-\mathbb{P}(Z>c) \\
\mathbb{P}(Z\le -c) &= \mathbb{P}(Z>c)
\end{aligned}$$

を示せ．(試験時間中にタイプミスは修正しました．)

(2) $P(-2<Z\le 1)$ の値を求めよ．

(3) $0<\alpha<1/2$ なる数 α に対して，実数 k_α は方程式

$$\int_{-\infty}^{k_\alpha} f_Z(z)\,dz = \alpha$$

をみたす点とする．このとき，

$$k_\alpha < 0$$

を示せ．

(4) $0<\alpha<1/2$ なる数 α に対して，実数 k_α は方程式

$$\int_{-\infty}^{k_\alpha} f_Z(z)\,dz = \alpha$$

をみたす点とする．このとき，確率

$$\mathbb{P}(k_\alpha < Z \le 0)$$

を求めよ．

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．

- $P(Z>1)=0.16$, $P(Z>2)=0.02$;
- $P(Z=c)=0$;
- 実数 a, b, c に対して，$\int_a^c f_Z(z)\,dz=\int_a^b f_Z(z)\,dz+\int_b^c f_Z(z)\,dz$;
- $\int_a^b f_Z(z)\,dz=-\int_b^a f_Z(z)\,dz$;
- $a<b$ とする．$g(x)>0\,(a\le x\le b)$ のとき，$\int_a^b g(x)\,dx>0$;
- 積分の変数変換の公式.

配点 (1) 5+5 (2) 5 (3) 5 (4) 5 計 25

Solutions

(1)

$$\begin{aligned}\mathbb{P}(Z\le c) &= \lim_{a\to-\infty} P(a<Z\le c)=\lim_{a\to-\infty}\int_a^c f_Z(z)\,dz\int_{-\infty}^c f_Z(z)\,dz=\int_{-\infty}^\infty f_Z(z)\,dz-\int_c^\infty f_Z(z)\,dz\\ &= 1-\mathbb{P}(Z>c).\end{aligned}$$

$w=-z$ とおけば，$f_Z(-z)=f_Z(w)$ に注意して，

$$\begin{aligned}\mathbb{P}(Z\le -c) &= \lim_{a\to-\infty}\mathbb{P}(a<Z\le -c)=\lim_{a\to-\infty}\int_a^{-c} f_Z(z)\,dz=\lim_{a\to-\infty}\int_{-a}^{c} f_Z(-w)\,(-dw)\\ &= \lim_{a\to-\infty}\int_{-a}^{c} -f_Z(w)\,dw=\lim_{a\to-\infty}\int_c^{-a} f_Z(w)\,dw=\lim_{b\to\infty}\int_c^b f_Z(w)\,dw=\lim_{b\to\infty}\mathbb{P}(c<Z\le b)\\ &= \mathbb{P}(c<Z)\end{aligned}$$

(2)

$$\begin{aligned}P(-2<Z\le 1) &= \int_{-2}^2 f_Z(z)\,dz=\int_{-\infty}^1 f_Z(z)\,dz-\int_{-\infty}^{-1} f_Z(z)\,dz\\ &= P(Z\le 1)-P(Z\le -2)=1-P(Z>1)-P(Z<-2)=1-P(Z>1)-P(Z>2)\\ &= 1-0.16-0.02=0.82.\end{aligned}$$

(3) $c=0$ とおけば，(1) より

$$\frac{1}{2}=\mathbb{P}(Z\le 0)=\int_{-\infty}^0 f_Z(z)\,dz$$

となる．さらに，$0<\alpha<1/2$ なので，

$$\int_{-\infty}^{k_\alpha} f_Z(z)\,dz<\frac{1}{2}=\int_{-\infty}^0 f_Z(z)\,dz$$

したがって，

$$0<\int_{-\infty}^0 f_Z(z)\,dz-\int_{-\infty}^{k_\alpha} f_Z(z)\,dz=\int_{k_\alpha}^0 f_Z(z)\,dz$$

となる．$f_Z(z)>0$ より，$k_\alpha<0$ となる．

(4)

$$\mathbb{P}(k_\alpha<Z\le 0)=\int_{k_\alpha}^0 f_Z(z)\,dz=\int_{-\infty}^0 Z(z)\,dz-\int_{-\infty}^{k_\alpha} f_Z(z)\,dz=\frac{1}{2}-\alpha=\frac{1}{2}-\alpha$$

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<-rt(n.sample,4)*3
> x<-round(x*10)/10
```

(1) オブジェクト x 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 x 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

(1) > (1/length(x))*sum(abs(x-mean(x)))
(2) > sqrt(sum((x-mean(x))**2)/length(x))

問題 5

以下のプログラムを入力（乱数の作成プログラム）

```
> repp<-10000
> n.sample<-round(runif(1,1,5))+5
> psd<-3
> barx<-rep(0,repp)
> varx<-rep(0,repp)
> res<-rep(0,repp)
> for (i in 1:repp){
+ x<-rnorm(n.sample,0,psd)
+ barx[i]<-mean(x)
+ varx[i]<-var(x)
+ res[i]<-barx[i]/(sqrt(varx[i])/sqrt(n.sample))
}
> chi2<-varx*(n.sample-1)/psd**2
```

(1) n.sample はいくつかを述べよ．
(2) 上のプログラムにおける x<-rnorm(n.sample,0,psd) によって，どのようなデータ（乱数）が生成されるかを述べよ．
(3) オブジェクト barx のヒストグラムを作図し，それに適合する確率密度関数を上書きするプログラムを作成するためのプログラムを書け．さらに，その分布を述べよ．
(4) オブジェクト chi2 のヒストグラムを作図し，それに適合する確率密度関数を上書きするプログラムを作成するためのプログラムを書け．さらに，その分布を述べよ．
(5) オブジェクト res のヒストグラムを作図し，それに適合する確率密度関数を上書きするプログラムを作成するためのプログラムを書け．さらに，その分布を述べよ．

配点 (1) 5 (2) 5 (2) 5 (3) 5 (4) 5 計 25

```
(2) > hist(barx,freq=F)
    > curve(dnorm(x,0,psd/sqrt(n.sample)),add=T,col=2)
```

```
(3) > hist(chi2,freq=F)
    > curve(dchisq(x,n.sample-1),add=T,col=2)

(4) > hist(res,freq=F,ylim=c(0,0.4),nclass=100)
    > curve(dt(x,n.sample-1),add=T,col=2)
```

得点分布　平均点 $= 45.2$，中央値 $= 46.0$、標準偏差 $= 16.7$

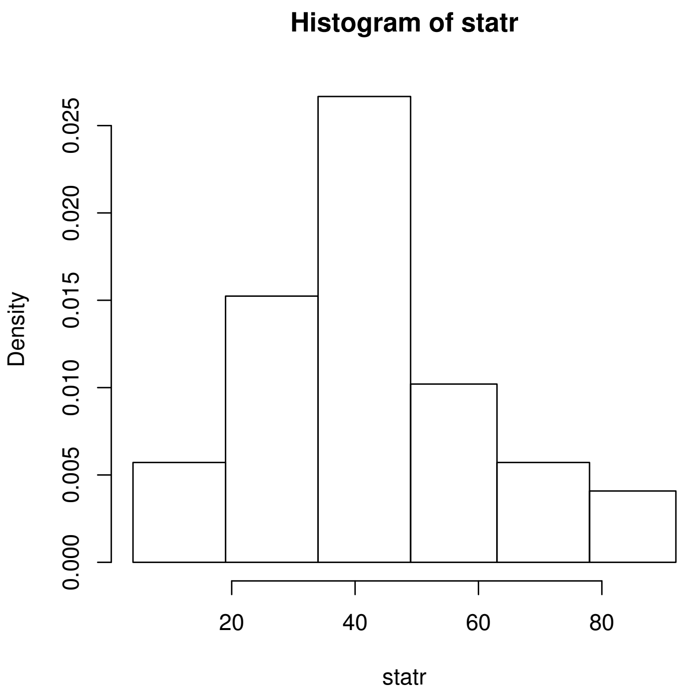

Figure 1: This is a figure