RSS Higher Certificate in Statistics, Specimen B

Module2 : Probability Models

Solutions

Question 1

(i) (a) $E(X)=\sum_{x=0}^{\infty}x\frac{e^{-\lambda}\lambda^x}{x!}=\sum_{x=1}^{\infty}\frac{\lambda e^{-\lambda}\lambda^{x-1}}{(x-1)!}=\lambda\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^x}{x!}=\lambda\times 1=\lambda$.

分散は $\mathrm{Var}(X)=E(X^2)-\{E(X)\}^2$ で得られるが $E(X^2)=E[X(X-1)]+E(X)$ を用いた方が簡単で, $\mathrm{Var}(X)=E[X(X-1)]+E(X)-\{E(X)\}^2$ となる.

$$E[X(X-1)]=\sum_{x=0}^{\infty}x(x-1)\frac{e^{-\lambda}\lambda^x}{x!}=\sum_{x=2}^{\infty}\frac{\lambda^2e^{-\lambda}\lambda^x}{(x-2)!}=\lambda^2\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^x}{x!}=\lambda^2\times 1=\lambda^2$$

であるから, $\mathrm{Var}(X)=\lambda^2+\lambda-\lambda^2=\lambda$ が得られる.

（確率母関数やモーメント母関数を用いることもできる．この方法はModule 5で扱う.）

(b) もし n が大きく, p が小さければ, パラメーター n と p を持つ二項分布はパラメーター np のポアソン分布で近似できる.「rule of thumb（経験に基づく法則）」によると, $1/2\le np\le 10$ が n をどれくらい大きくすべきか, また p をどれくらい小さくすべきかの指標になる.（もし $np>10$ ならば, 二項分布の正規近似の方がよい.）

(ii) X を誤りが生じた数であるとすると, X は二項分布 B (200, 0.0075) に従う. よって求める確率は以下のようになる.

$$P(X=1)=\binom{200}{1}(0.0075)(0.9925)^{199}=200\times 0.0075\times 0.9925^{199}=0.33532\cdots=0.3353$$

$$P(X=4)=\binom{200}{4}(0.0075)^4(0.9925)^{196}$$

$$=\frac{200\times 199\times 198\times 197}{4\times 3\times 2\times 1}\times 0.0075^4\times 0.9925^{196}=0.04680\cdots=0.0468$$

(iii) X ~ Poisson (200 × 0.0075=1.5) を用いて近似する. これにより

$$P(X=1)=1.5e^{-1.5}=0.33470\cdots=0.3347$$

となり, 相対誤差は $\frac{100(0.33532-0.33470)}{0.33532}=0.18\%$ となる. また

$$P(X=4)=\frac{e^{-1.5}(1.5)^4}{4!}=0.04707\cdots=0.0470$$

となり, この相対誤差は $\frac{100(0.04707-0.04680)}{0.04680}=0.58\%$ である.

（注意： 二項分布とポアソン分布における確率計算をもっと高精度に行うならば，これらの相対誤差は少し異なる結果となるであろう．）

両者の近似はとても正確であり，相対誤差は 1% をはるかに下回る． $X=1$ のとき（誤りが 1 つのとき）の近似の方が $X=4$（誤りが 4 つのとき）の近似よりも正確である．また $X=1$ のときの近似は過小評価であり，$X=4$ のときは過大評価となっている．

Question 2

$H \sim \mathrm{N}\,(160,16)$ である.

(i) $P\left(156 < H \le 164\right) = P\left(\dfrac{156-160}{4} < Z \le \dfrac{164-160}{4}\right)$ （ここで $\mathrm{Z} \sim \mathrm{N}\,(0,1)$ とする.）

$= \mathrm{P}(-1 < \mathrm{Z} \le 1) = \Phi\,(1) - \Phi\,(-1) = \Phi\,(1) - \{1 - \Phi\,(-1)\}$ （対称性より）

$= 2\left(\Phi\left(1\right)\right) - 1 = 0.6826$. （$\Phi$ は標準正規分布の累積分布関数）

$P\left(H > 168\right) = P\left(Z > 2\right) = 1 - \Phi\left(2\right) = 0.0228$.

(ii) (a) APF の警官に関する部分は $H = 168$ から始まる．これは $Z = 2$ に対応する.

この部分における中央値 m は上側 $\dfrac{1}{2}\left(0.0228\right) = 0.0114$ の確率をとる値である.

したがって，$\Phi\left(m\right) = 1 - 0.0114 = 0.9886$ であり，$Z = 2.277$ に相当する．H の相当する値は $\mu + \sigma Z$ より，$160 + \left(4 \times 2.277\right) = 169.1\,\mathrm{cm}$ となる.

(b) $H = 170$ は $Z = \dfrac{10}{4} = 2.5$ に相当する．$\Phi\left(2.5\right) = 0.9938$ となるので，$1 - \Phi\left(2.5\right) = 0.0062$ となる．$H > 168$ は条件であるので，確率は $\dfrac{P\left(H > 170\right)}{P\left(H > 168\right)} = \dfrac{0.0062}{0.0228} = 0.272$ となる.

(iii) 25 人の平均身長は $N\left(169.5, \dfrac{1.352^2}{25}\right)$ に従うので

$$P(\mathrm{mean} > 170) = P\left(Z > \frac{170 - 169.5}{(1.352/5)}\right) = 1 - \Phi\left(\frac{0.5 \times 5}{1.352}\right) = 1 - 0.9677 = 0.0323$$

となる.

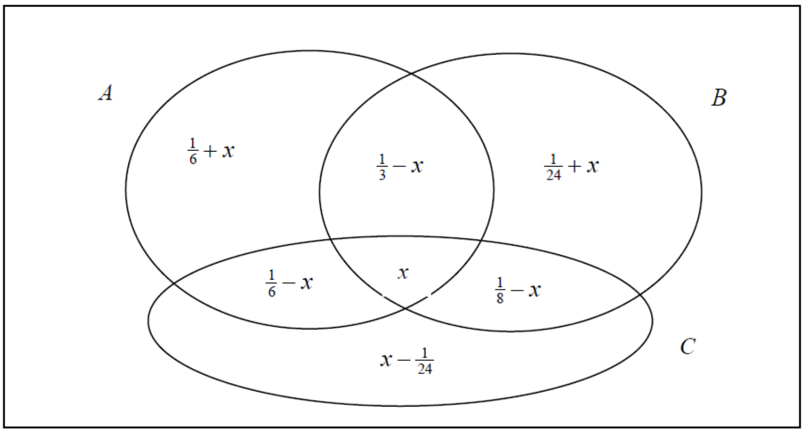
Question 3

$P(A)=\dfrac{2}{3}$ $\qquad$ $P(B)=\dfrac{1}{2}$ $\qquad$ $P(C)=\dfrac{1}{4}$

(i) (a) 独立性から $P\left(A\cap\bar{B}\cap\bar{C}\right)=P(A)P\left(\bar{B}\right)P\left(\bar{C}\right)=\dfrac{2}{3}\left(1-\dfrac{1}{2}\right)\left(1-\dfrac{1}{4}\right)=\dfrac{1}{4}$.

(b) $P(A\cap\bar{C}\mid A\cap\bar{B})=\dfrac{P((A\cap\bar{C})\cap(A\cap\bar{B}))}{P(A\cap\bar{B})}=\dfrac{P(A\cap\bar{B}\cap\bar{C})}{P(A\cap\bar{B})}=\dfrac{1/4}{2/3\times(1-1/2)}=\dfrac{3}{4}$.

(ii)



対ごとの独立性を用いると， $P\left(A\cap B\right)=P(A)P(B)$ などとなるので,

$$P(A\cap B\cap\bar{C})=1/3-x,\ P(\bar{A}\cap B\cap C)=1/8-x,\ P(A\cap B\cap\bar{C})=1/6-x,$$

が得られる．その他は $P(A)$， $P(B)$， $P(C)$ を用いて表される.

(a) 図から $P\left(A\cap\bar{B}\cap\bar{C}\right)=\dfrac{1}{6}+x$.

(b)

$$P(A\cap\bar{C}\mid A\cap\bar{B})=\frac{P((A\cap\bar{C})\cap(A\cap\bar{B}))}{P(A\cap\bar{B})}=\frac{P(A\cap\bar{B}\cap\bar{C})}{P(A\cap\bar{B})}$$
$$=\frac{1/6+x}{(1/6+x)+(1/6-x)}=3x+\frac{1}{2}$$

(c) $P(A\cup B\cup C)=\dfrac{19}{24}+x$

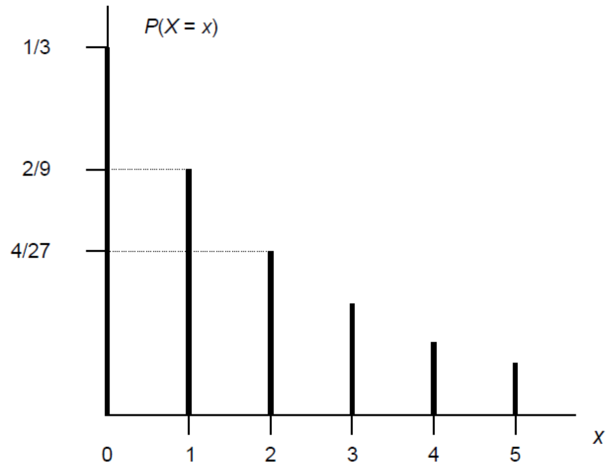
全ての確率が区間 [0, 1] に入らなければならないので， $x\geq\dfrac{1}{24}$ かつ $x\leq\dfrac{1}{8}$ ， すなわち

$\dfrac{1}{24}\leq x\leq\dfrac{1}{8}$ である.

Question 4

確率関数: $f(x)=(1-p)^x p \quad x=0,1,2,\ldots \quad 0<p<1$

(i)



(ii) 期待値は $E(X)=\sum_{x=0}^{\infty}x(1-p)^x p$ で表される．これはさまざまな方法で求めることができる．

ここで示す方法（幾何数列の使用）はそれほど効率的ではないが，理解しやすい．

$$\sum_{x=0}^{\infty}x(1-p)^x p = p\left\{0+(1-p)+2(1-p)^2+3(1-p)^3+\ldots\right\}$$

$$\begin{aligned}= p\{0+(1-p)+(1-p)^2+(1-p)^3+\ldots \\ +(1-p)^2+(1-p)^3+\ldots \\ +(1-p)^3+\ldots \\ +\ldots\}\end{aligned}$$

$$= p\left\{\frac{1-p}{1-(1-p)}+\frac{(1-p)^2}{1-(1-p)}+\frac{(1-p)^3}{1-(1-p)}+\ldots\right\}=\frac{p}{p}\left\{\frac{1-p}{1-(1-p)}\right\}=\frac{1-p}{p}$$

同様に $E(X^2)=\sum_{x=0}^{\infty}x^2(1-p)^x p=\frac{(1-p)^2+(1-p)}{p^2}$.

したがって，$\mathrm{Var}(X)=E(X^2)-\{E(X)\}^2=\frac{(1-p)^2+(1-p)}{p^2}-\left(\frac{1-p}{p}\right)^2=\frac{1-p}{p^2}$.

（確率母関数やモーメント母関数を用いることもできる．この方法は Module 5 で扱う.）

(iii) $P(X\geq x)=\sum_{r=x}^{\infty}(1-p)^r p=\frac{(1-p)^x p}{1-(1-p)}=(1-p)^x \quad (x=0,1,2,\ldots)$

ここで，$P(A|B)=\dfrac{P(A\cap B)}{P(B)}$ を用い，事象 A を「$X\ge l+m$」，B を「$X\ge l$」とすると，$A\cap B=A$ となる．したがって $P(X\ge l+m\,|\,X\ge l)=\dfrac{(1-p)^{l+m}}{(1-p)^{l}}=(1-p)^m=P(X\ge m)$ が成り立つ．これは幾何分布の「無記憶性」である．

(iv) 独立性より $P(Z\ge z)=P(X\ge z)P(Y\ge z)=(1-p)^z(1-\theta)^z$ であるから

$$
\begin{aligned}
P(Z=z)&=P(Z\ge z)-P(Z\ge z+1)=\{(1-p)(1-\theta)\}^z-\{(1-p)(1-\theta)\}^{z+1}\\
&=\{(1-p)(1-\theta)\}^z(1-1+p+\theta-p\theta)=\{(1-p)(1-\theta)\}^z(p+\theta-p\theta) \qquad z=0,1,\ldots
\end{aligned}
$$

となる．これは問題で与えられた幾何分布の確率関数において p を $p+\theta-p\theta$ に置き換えたものである．したがって (ii) から

$$
E(Z)=\frac{1-p-\theta+p\theta}{p+\theta-p\theta},\quad \mathrm{Var}(Z)=\frac{1-p-\theta+p\theta}{(p+\theta-p\theta)^2}
$$

が得られる．