THE ROYAL STATISTICAL SOCIETY

2008 EXAMINATIONS – SOLUTIONS

HIGHER CERTIFICATE

(MODULAR FORMAT)

MODULE 2

PROBABILITY MODELS

Note. In accordance with the convention used in the Society's examination papers, the notation log denotes logarithm to base *e*. Logarithms to any other base are explicitly identified, e.g. $\log_{10}$.

Higher Certificate, Module 2, 2008. Question 1

(i) (a) The number of PINs with four different digits is $10\times9\times8\times7 = 5040$.

(b) We require exactly three different digits. We can choose the face value of the pair in 10 ways. We can then choose two other different digits in $\binom{9}{2} =$ 36 ways. The number of distinguishable linear arrangements of two like and two unlike objects is $\dfrac{4!}{2!\times1!\times1!} = 12$, so the total number of 4-digit PINs with exactly three different digits is $12\times10\times36 = 4320$.

(c) We require two different digits, each occurring twice. We can choose the face values of the two pairs in $\binom{10}{2} = 45$ ways. The number of distinguishable linear arrangements of two (different) pairs of like objects is $\dfrac{4!}{2!\times2!} = 6$, so the total number of 4-digit PINs with two pairs of (different) like digits is $6\times45 = 270$.

(d) We require exactly three digits the same. We can choose the face values of the triple and of the singleton in 10×9 ways (note that *aaab* and *bbba* are different PINs). The number of distinguishable linear arrangements is 4 (corresponding to 4 different places for the singleton), hence there are $4\times904 = 360$ possible PINs.

(ii) (a) There are altogether $\binom{10}{4} = 210$ ways of choosing the 4 digits of the second PIN, each being equally likely with probability 1/210.

Now consider the number of ways of choosing 4 digits for the second PIN such that k of them (for $k = 0, 1, 2, 3, 4$) are in common with digits in an arbitrary given PIN of four different digits. There are $\binom{4}{k}$ ways of choosing the k digits that are in common and $\binom{6}{4-k}$ ways of choosing the $4-k$ digits that are not in common. So the total number of ways is $\binom{4}{k}\binom{6}{4-k}$.

So the required probability is $\dfrac{\binom{4}{k}\binom{6}{4-k}}{\binom{10}{4}}$.

Solution continued on next page

(b) $$P(X=0)=\frac{\binom{4}{0}\binom{6}{4}}{\binom{10}{4}}=\frac{15}{210}=\frac{1}{14};$$

$$P(X=1)=\frac{\binom{4}{1}\binom{6}{3}}{\binom{10}{4}}=\frac{80}{210}=\frac{8}{21};$$

$$P(X=2)=\frac{\binom{4}{2}\binom{6}{2}}{\binom{10}{4}}=\frac{90}{210}=\frac{3}{7};$$

$$P(X=3)=\frac{\binom{4}{3}\binom{6}{1}}{\binom{10}{4}}=\frac{24}{210}=\frac{4}{35};$$

$$P(X=4)=\frac{\binom{4}{4}\binom{6}{0}}{\binom{10}{4}}=\frac{1}{210}.$$

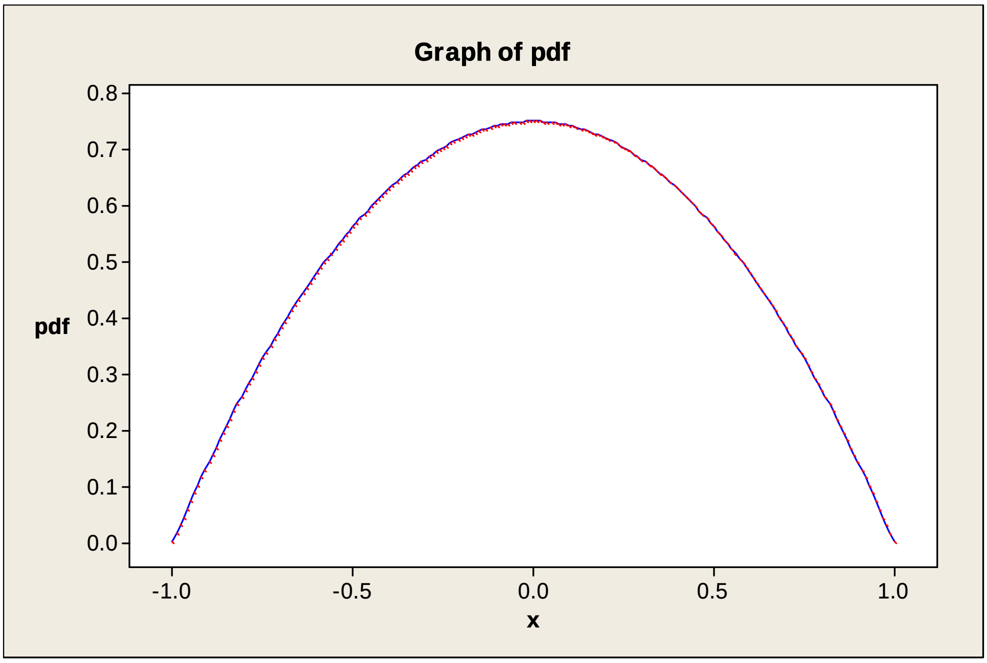
$$E(X)=\left(0\times\frac{15}{210}\right)+\left(1\times\frac{80}{210}\right)+\left(3\times\frac{90}{210}\right)+\left(3\times\frac{24}{210}\right)+\left(4\times\frac{1}{210}\right)$$

$$=\frac{336}{210}\ =\ 1.6.$$

Higher Certificate, Module 2, 2008. Question 2

(i) $1 = \int_{-1}^{1} c\left(1-x^2\right)dx = c\left[x-\frac{x^3}{3}\right]_{-1}^{1} = c\left(\left[1-\frac{1}{3}\right]-\left[(-1)-\frac{(-1)^3}{3}\right]\right) = \frac{4c}{3}$,

so $c = ¾$ or 0.75.



[The graph should of course be a smooth curve; due to the limits of electronic reproduction, it may not appear so. The maximum is at (0, 0.75), zeros at (±1,0).]

(ii) For $|x| \le 1$,

$$F_X(x) = \frac{3}{4}\int_{-1}^{x}\left(1-u^2\right)du = \frac{3}{4}\left[u-\frac{u^3}{3}\right]_{-1}^{x} = \frac{3}{4}\left[x-\frac{x^3}{3}-(-1)+\frac{(-1)^3}{3}\right] = \frac{2+3x-x^3}{4}.$$

For $x < -1$, $F_X(x) = 0$; for $x > 1$, $F_X(x) = 1$.

$$P\left(-\frac{1}{2} \le X \le \frac{1}{2}\right) = F_X\left(\frac{1}{2}\right) - F_X\left(-\frac{1}{2}\right) = \frac{2+1.5-0.125-2+1.5-0.125}{4} = \frac{11}{16}.$$

(iii) $E(X) = 0$ by symmetry (or by integration).

$$\therefore \ \mathrm{Var}(X) = E(X^2) = \frac{3}{4}\int_{-1}^{1} x^2\left(1-x^2\right)dx = \frac{3}{4}\left[\frac{x^3}{3}-\frac{x^5}{5}\right]_{-1}^{1} = \frac{3}{4}\times\frac{4}{15} = \frac{1}{5}$$

and $\mathrm{SD}(X) = 1/\sqrt{5} = 0.447$ to 3 significant figures.

Higher Certificate, Module 2, 2008. Question 3

$X \sim \mathrm{N}(0, 1)$, $Y \sim \mathrm{N}(0, 1)$; X and Y are independent

(i) $P(3X > 4Y + 2) = P(3X - 4Y > 2)$,

$= P(V > 2)$, where $V = 3X - 4Y \sim \mathrm{N}(0, 3^2 + 4^2 = 25)$.

$P(V > 2) = P(Z > \frac{2-0}{5} = 0.4$ [where $Z \sim \mathrm{N}(0, 1)$] $) = 1 - \Phi(0.4) = 0.3446$.

Since X and Y are independent, $P(X \le x, Y \le x) = P(X \le x).P(Y \le x) = [\Phi(x)]^2$.

(ii) (a) $\max(X, Y) \le w \Leftrightarrow (X \le w) \cap (Y \le w)$,

so $P(\max(X, Y) \le w) = [\Phi(w)]^2$ from above.

(b) $Q1$ satisfies $F_W(Q1) = \frac{1}{4}$, so $\Phi^2(Q1) = \frac{1}{4}$.

$\therefore \Phi(Q1) = \frac{1}{2}$, and $Q1 = \Phi^{-1}(0.5) = 0$.

Similarly, $F_W(Q3) = \frac{3}{4} \Rightarrow \Phi(Q3) = \frac{\sqrt{3}}{2} = 0.866$, so $Q3 = \Phi^{-1}(0.866)$ $= 1.108$ using linear interpolation in the Society's *Statistical tables for use in examinations* [1.11 was allowed as the (3 s.f.) answer in the examination, being the nearest tabular entry].

(iii) $P(W$ outside $(Q1, Q3)) = 0.5$, so $N \sim \mathrm{B}(100, 0.5)$ which we approximate by $\mathrm{N}(50, 25)$. Hence

$$P(N \ge 58) \approx 1 - \Phi\left(\frac{57.5 - 50}{5}\right) = 1 - \Phi(1.5) = 1 - 0.9332 = 0.0668.$$

Higher Certificate, Module 2, 2008. Question 4

(i) $$p_X(x+1) = \frac{e^{-\lambda}\lambda^{x+1}}{(x+1)!} = \frac{\lambda}{x+1}\frac{e^{-\lambda}\lambda^{x}}{x!} = \frac{\lambda}{x+1}p_X(x).$$

This can be used recursively to find the probability mass function. Start with $p_X(0) = e^{-\lambda}$; then $p_X(1) = \lambda p_X(0) = \lambda e^{-\lambda}$, $p_X(2) = (\lambda/2)p_X(1) = (\lambda^2/2)e^{-\lambda}$, and so on.

(ii) $$E(X) = \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^{x}}{x!} = \sum_{x=1}^{\infty}\frac{e^{-\lambda}\lambda^{x}}{(x-1)!} = \lambda\sum_{x=1}^{\infty}\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} = \lambda\sum_{y=0}^{\infty}\frac{e^{-\lambda}\lambda^{y}}{y!} = \lambda,$$

putting $y = x - 1$ in the last summation and noticing that this re-creates the probability mass function. Similarly,

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

putting $y = x - 2$ in the last summation.

Hence $\text{Var}(X) = E[X(X-1)] + E(X) - \{E(X)\}^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$, as required.

(iii) $$P(W = w) = \sum_{x=0}^{w}\frac{e^{-\lambda}\lambda^{x}}{x!}\times\frac{e^{-\mu}\mu^{w-x}}{(w-x)!} = \frac{e^{-(\lambda+\mu)}}{w!}\sum_{x=0}^{w}\binom{w}{x}\lambda^{x}\mu^{w-x} = e^{-(\lambda+\mu)}\frac{(\lambda+\mu)^{w}}{w!},$$

confirming that $W \sim \text{Poisson}(\lambda + \mu)$. Since the general parameter λ has been shown in part (ii) to represent the mean, it follows that $E(W) = \lambda + \mu$.

(iv) (a) P(exactly one breakdown)

= P(A fails once, B does not fail) + P(B fails once, A does not fail)

= P(A fails once) × P(B does not fail)

+ P(B fails once) × P(A does not fail)

$= (\lambda e^{-\lambda}\times e^{-\mu}) + (\mu e^{-\mu}\times e^{-\lambda}) = (\lambda + \mu)e^{-(\lambda+\mu)}$.

$\therefore$ the required conditional probability is

$$\frac{P(A \text{ fails once, } B \text{ does not fail})}{(\lambda+\mu)e^{-(\lambda+\mu)}}$$

$$= \frac{(\lambda e^{-\lambda}\times e^{-\mu})}{(\lambda+\mu)e^{-(\lambda+\mu)}} = \frac{\lambda}{\lambda+\mu} = \frac{2}{2.5} = 0.8.$$

Solution continued on next page

(b) W = total number of breakdowns ~ Poisson(2.5).

$\therefore\ P(W > 2) = 1 - P(W \le 2)$

$$= 1 - e^{-2.5}\left(1 + 2.5 + (2.5^2/2)\right)$$

$$= 1 - 6.625e^{-2.5} \;=\; 1 - 0.5438 \;=\; 0.456$$

(alternatively, this can be obtained from the cumulative Poisson probabilities in the Society's *Statistical tables for use in examinations*).

(c) T ~ Poisson(50×2.5) or Poisson(125), which we approximate by N(125, 125).

The upper 5% point of N(125, 125) is $125 + 1.6449\sqrt{125} = 143.4$.

Since $T_{0.95}$ must be an integer and the question says "will be exceeded on at most 5% of days", we round up to $T_{0.95} = 144$.