# EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY

**HIGHER CERTIFICATE IN STATISTICS, 2008**

**(Modular format)**

**MODULE 2 : Probability models**

**Time allowed: One and a half hours**

*Candidates should answer* **THREE** *questions.*

*Each question carries 20 marks.*
*The number of marks allotted for each part-question is shown in brackets.*

*Graph paper and Official tables are provided.*

*Candidates may use calculators in accordance with the regulations published in the Society's "Guide to Examinations" (document Ex1).*

*The notation* log *denotes logarithm to base* $\boldsymbol{e}$.
*Logarithms to any other base are explicitly identified, e.g.* $\log_{10}$.

*Note also that* $\binom{n}{r}$ *is the same as* ${}^nC_r$.



This examination paper consists of 4 printed pages **each printed on one side only**.
This front cover is page 1.
Question 1 starts on page 2.

There are 4 questions altogether in the paper.

1. A Personal Identification Number (PIN) consists of four digits in order, each of which may be any one of 0, 1, 2, …, 9.

(i) Find the number of PINs satisfying each of the following requirements.

(a) All four digits are different.

(b) There are exactly three different digits.

(c) There are two different digits, each of which occurs twice.

(d) There are exactly three digits the same.

(9)

(ii) Two PINs are chosen independently and at random, and you are given that each PIN consists of four different digits. Let $X$ be the random variable denoting the number of digits that the two PINs have in common.

(a) Explain clearly why $P(X = k) = \dfrac{\binom{4}{k}\binom{6}{4-k}}{\binom{10}{4}}$, for $k = 0, 1, 2, 3, 4$.

(4)

(b) Hence write down the values of the probability mass function of $X$, and find its mean.

(7)

2. The continuous random variable $X$ has probability density function given by

$$f_X(x) = c(1 - x^2), \quad -1 \le x \le 1,$$

where $c$ is a suitable constant.

(i) Show that $c = ¾$ and plot the graph of $f_X(x)$ against $x$.

(8)

(ii) Show that the cumulative distribution function of $X$ is given by

$$F_X(x) = \begin{cases} 0 & , \ x < -1, \\ \dfrac{2 + 3x - x^3}{4} & , \ -1 \le x \le 1, \\ 1 & , \ x > 1. \end{cases}$$

Also find $P(-\frac{1}{2} \le X \le \frac{1}{2})$.

(7)

(iii) Obtain the standard deviation of $X$, giving your answer correct to 3 significant figures.

(5)

3. Let $X$ and $Y$ be independent standard Normal random variables and let $\Phi(.)$ denote the cumulative distribution function of a standard Normal random variable.

(i) Find $P(3X > 4Y + 2)$, and write down $P(X \le x, Y \le x)$ in terms of $\Phi(x)$.

(7)

(ii) Let $W = \max(X, Y)$.

(a) Explain why the cumulative distribution function of $W$ is given by

$$F_W(w) = \left[\Phi(w)\right]^2, \ -\infty < w < \infty.$$

(3)

(b) Find $Q1$ and $Q3$, the lower and upper quartiles of $W$.

(5)

(iii) A random sample of 100 observations of $W$ is taken. Write down the distribution of the number $N$ of observations in the sample which lie outside the interval $(Q1, Q3)$. Use a suitable approximation to calculate $P(N \ge 58)$.

(5)



**Turn over**

4. The Poisson random variable $X$ with parameter $\lambda > 0$ has probability mass function

$$p_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots .$$

(i) Show that, for integer $x \geq 0$,

$$p_X(x+1) = \frac{\lambda}{x+1} p_X(x),$$

and outline how this could be used to calculate the probability mass function for a given value of $\lambda$.

(3)

(ii) Show that $E(X) = \text{Var}(X) = \lambda$.

(6)

(iii) Suppose that $Y$ has a Poisson distribution with mean $\mu$ independently of $X$, and that $W = X + Y$. Use the relation

$$P(W = w) = \sum_{x=0}^{w} P(X = x)P(Y = w - x)$$

to show that $W$ has a Poisson distribution, and write down its mean.

(4)

(iv) An office has two computer systems, one of Type $A$ and one of Type $B$. The numbers of breakdowns per day on these systems, $X$ and $Y$ say, have independent Poisson distributions with respective means 2 and 0.5.

(a) Find the conditional probability that if there is exactly one breakdown on a given day then it is the Type $A$ system that fails.

(2)

(b) Find the probability that on a given day there are more than 2 breakdowns.

(2)

(c) The office is one of 50 run by a large company. The offices are each equipped with one Type $A$ system and one Type $B$ system, which function independently in the way described above. Write down the distribution of $T$, the total number of breakdowns occurring in the 50 offices on any given day. Use a suitable approximation to estimate $T_{0.95}$, the number of breakdowns per day which will be exceeded on at most 5% of days.

(3)