EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY

HIGHER CERTIFICATE IN STATISTICS, 2017

MODULE 5 : Further probability and inference

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* **e**.
Logarithms to any other base are explicitly identified, e.g. $\log_{10}$.

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

This examination paper consists of 4 printed pages.
This front cover is page 1.
Question 1 starts on page 2.

There are 4 questions altogether in the paper.

1. The continuous random variables X and Y are jointly distributed with joint probability density function

$$f(x, y) = k(x^2 + xy), \quad 0 \le x \le 1, \ 0 \le y \le 2.$$

(i) Show that $k = \frac{3}{5}$.

(2)

(ii) Find the marginal probability density function of Y and use it to show that $P(Y \le 1) = \frac{7}{20}$.

(5)

(iii) Find $f(x \mid y)$, the conditional probability density function of X given $Y = y$. Use it to find the mean and variance of X given that $Y = 1$.

(10)

(iv) Use the joint probability density function to show that $P(Y < X) = \frac{9}{40}$.

(3)

2. Let $x_1, x_2, ..., x_n$ be independent observations of a discrete random variable X whose probability mass function is given by

$$f(x) = (1-p)p^x, \quad x = 0, 1, 2, \ldots$$

where $0 < p < 1$.

(i) Find the moment generating function of X and use it to show that

$$E(X) = \frac{p}{1-p} \quad \text{and} \quad \mathrm{Var}(X) = \frac{p}{(1-p)^2}.$$

(7)

(ii) Show that the method of moments and maximum likelihood estimators of p are identical.

(5)

(iii) By substituting the common estimator for p found in part (ii) into the expression for the variance in part (i), show that the maximum likelihood estimator of $\mathrm{Var}(X)$ is $\bar{X}(\bar{X}+1)$.

(2)

(iv) Show that $E\left[\bar{X}(\bar{X}+1)\right] = \dfrac{(n+1)p}{n(1-p)^2}$ and hence suggest an unbiased estimator of $\mathrm{Var}(X)$.

(6)

3. (a) Define the terms *bias* and *relative efficiency* of potential estimators of a population parameter. Explain briefly how these may be useful when deciding between two or more estimators.

(4)

(b) Two forecasters make forecasts of an economic parameter θ. The forecasters are known to make independent, unbiased predictions. The prediction of forecaster 1 is the random variable X_1, which follows a $N(\theta, \sigma^2)$ distribution, whilst that of forecaster 2 is X_2, which follows a $N(\theta, 2\sigma^2)$ distribution.

A pooled prediction of θ from both forecasters is $\hat{\theta} = a_1 X_1 + a_2 X_2$ where a_1 and a_2 are constants.

(i) State an equation which a_1 and a_2 must satisfy in order for the pooled forecast to be unbiased for θ.

(2)

(ii) Find the values of a_1 and a_2 which minimise the variance of an unbiased estimator $\hat{\theta}$ and find this minimum variance.

(8)

(iii) Comment briefly on why the values for a_1 and a_2 seem reasonable.

(1)

(iv) It is suggested that $\tilde{\theta}$, the mean of X_1 and X_2, should be used instead to estimate θ. Show that $\tilde{\theta}$ is also unbiased and calculate the relative efficiency of $\tilde{\theta}$ compared with $\hat{\theta}$.

(5)

4. Two fair tetrahedral dice are thrown, each having sides numbered 1 to 4. The random variable X_1 denotes the smaller of the two downturned numbers whilst X_2 is the larger of these.

(i) Find the joint probability distribution of X_1 and X_2, and display it in the form of a two-way table.

(4)

(ii) Find the marginal probability distribution of X_1 and hence find $E(X_1)$ and $\text{Var}(X_1)$. Deduce the values of $E(X_2)$ and $\text{Var}(X_2)$ without finding the marginal distribution of X_2.

(7)

(iii) Find the conditional probability distribution of X_1 given that $X_2 = 4$ and find $E(X_1 \mid X_2 = 4)$. Would you expect $E(X_1 \mid X_2 = 3)$ to be higher or lower than this value? Explain your answer.

(5)

(iv) Show that $\text{Cov}(X_1, X_2) = \frac{25}{64}$.

(4)