EXAMINATIONS OF THE HONG KONG STATISTICAL SOCIETY

HIGHER CERTIFICATE IN STATISTICS, 2015

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 8 printed pages.
This front cover is page 1.
Question 1 starts on page 2.

There are 4 questions altogether in the paper.

1. The random variables X and Y are jointly distributed. Define the *covariance* and the *correlation* of X and Y. Suppose that the correlation of X and Y is -1; what does this tell you about the relationship between X and Y?

(4)

In a certain population, the height of a father and the height of his eldest son have a bivariate Normal distribution with the father's height having mean 176.0 cm and standard deviation 7.0 cm, while his eldest son's height has mean 178.0 cm and standard deviation 7.2 cm. The correlation between the father's and eldest son's heights is 0.52.

(i) Show that the covariance between the father's and eldest son's heights is 26.208.

(2)

(ii) Find the probability that the mean height of a father and his eldest son exceeds 180 cm.

(7)

(iii) Find the probability that the height of an eldest son exceeds that of his father by at least 5%.

(7)

2. The daily number of road traffic accidents, Y, in a certain town can be modelled by a Poisson distribution with parameter $\lambda > 0$, which has probability mass function

$$P(Y = k) = \frac{e^{-\lambda}\lambda^k}{k!} \text{ for } k = 0, 1, 2, \ldots .$$

In a random sample of n days, the numbers of road traffic accidents on those days are found to be $Y_1, Y_2, \ldots, Y_n$.

(i) Show that the probability generating function (pgf) of Y is $e^{-\lambda(1-t)}$.

(3)

(ii) Use the pgf to show that $E(Y) = \text{Var}(Y) = \lambda$.

(5)

(iii) Use the pgf to find the distribution of $\sum_{i=1}^{n} Y_i$.

(3)

(iv) Find the maximum likelihood estimator, $\hat{\lambda}$, of λ.

(5)

(v) Show that $\hat{\lambda}$ is also the method of moments estimator of λ.

(2)

(vi) Use the central limit theorem to find an approximate 95% confidence interval for λ. (You may assume that n is large.)

(2)

3. The random sample $X_1, X_2, \ldots, X_n$ comes from a distribution with probability density function

$$f(x) = \frac{x^3 e^{-x/\sqrt{\alpha}}}{6\alpha^2} \quad \text{for } x > 0,$$

where $\alpha > 0$ is an unknown parameter. This distribution has moment generating function $m(t) = \left(1 - t\sqrt{\alpha}\right)^{-4}$ for $t < 1/\sqrt{\alpha}$.

(i) Use the moment generating function to find $E(X_1), E(X_1^2), E(X_1^3)$ and $E(X_1^4)$.

(8)

(ii) Show that

$$\hat{\alpha} = \frac{\sum_{i=1}^{n} X_i^2}{20n}$$

is an unbiased estimator of α.

(2)

(iii) Show that $\mathrm{Var}(\hat{\alpha}) = \dfrac{1.1\alpha^2}{n}$.

(3)

(iv) Find the second derivative of the log likelihood and hence find the efficiency of $\hat{\alpha}$.

(7)

4. The continuous random variables X and Y are jointly distributed and

$$f(x) = \begin{cases} 2x & \text{for } 0 < x < 1, \\ 0 & \text{otherwise,} \end{cases} \qquad f(y \mid x) = \begin{cases} x^{-1} & \text{for } 0 < y < x, \\ 0 & \text{otherwise.} \end{cases}$$

(i) Find the joint probability density function and sketch the region where it is non-zero.

(4)

(ii) Find the marginal probability density function of Y and show that $E(Y) = \frac{1}{3}$.

(5)

(iii) Find the conditional density function $f(x \mid y)$ and use it to evaluate $E\left(X \mid Y = \frac{1}{2}\right)$.

(6)

(iv) Evaluate $P(Y < \frac{1}{2}X)$.

(5)

BLANK PAGE

BLANK PAGE

BLANK PAGE