EXAMINATIONS OF THE HONG KONG STATISTICAL SOCIETY

HIGHER CERTIFICATE IN STATISTICS, 2014

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.

© RSS 2014

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

$$f(x, y) = kx^2 \quad (0 < x < 1,\ 0 < y < x)$$

and zero elsewhere, where k is a constant.

(i) Sketch the region where the joint density is non-zero.

(4)

(ii) Use integration to show that $k = 4$.

(3)

(iii) Find the marginal probability density functions of X and Y.

(4)

(iv) Evaluate $\mathrm{Cov}(X, Y)$.

(9)

2. (i) Z_1 and Z_2 are independent random variables, each with the N(0, 1) distribution. The random variable U is defined by $U = a_0 + a_1Z_1 + a_2Z_2$, where a_0, a_1 and a_2 are constants. Find the moment generating function of U and hence deduce the distribution of U.

[You may use the result that the moment generating function of the $\mathrm{N}(\mu, \sigma^2)$ distribution is $m(t) = \exp(\mu t + \frac{1}{2}\sigma^2 t^2)$.]

(7)

(ii) Suppose that Z_1, Z_2 and U are as defined in part (i) and $V = b_0 + b_1Z_1 + b_2Z_2$, where b_0, b_1 and b_2 are constants. Give the distribution of V and show that $\mathrm{Cov}(U, V) = a_1b_1 + a_2b_2$. Assuming that the joint distribution of (U, V) is bivariate Normal, write down the five parameters of that distribution in terms of a_0, a_1, a_2, b_0, b_1, b_2.

(7)

(iii) A computer program allows you to generate pseudo-random numbers which behave like independent values from the N(0, 1) distribution. You wish to generate pairs of values (X_i, Y_i) $(i = 1, 2, 3, \ldots)$ which behave like observations from a bivariate Normal distribution with $E(X_i) = 5$, $E(Y_i) = 8$, $\mathrm{Var}(X_i) = 4$, $\mathrm{Var}(Y_i) = 16$ and $\mathrm{Cov}(X_i, Y_i) = -4$. From the results of part (ii), find a method of generating the pairs (X_i, Y_i) using the pseudo-random N(0, 1) numbers.

[There are multiple solutions; choose any solution that you find convenient.]

(6)

3. The time in seconds, T, between hits at a webpage has the exponential distribution with probability density function

$$f(t) = \lambda e^{-\lambda t} \quad \text{for } t > 0$$

where $\lambda > 0$ is an unknown parameter.

(i) Independent observations $T_1, T_2, \ldots, T_n$ have been made of T. Find the maximum likelihood estimator $\hat{\lambda}$ of λ.

(6)

(ii) Find the asymptotic variance of $\hat{\lambda}$.

(2)

(iii) Suppose now that we can only observe $X_1, X_2, \ldots, X_n$, defined by

$$X_i = \begin{cases} 0 & \text{if } T_i \le c \\ 1 & \text{if } T_i > c \end{cases}$$

for $i = 1, 2, \ldots, n$ where c (> 0) is a known constant. Let $Y = \sum_{i=1}^{n} X_i$. Show that $P(X_i = 1) = e^{-\lambda c}$ and hence find the distribution of Y. Using this result, find $\tilde{\lambda}$, the maximum likelihood estimator of λ based on Y.

(8)

(iv) Show that the asymptotic variance of $\tilde{\lambda}$ is $\dfrac{e^{\lambda c} - 1}{nc^2}$.

(4)

4. (a) The observations $X_1, X_2, \ldots, X_n$ form a random sample from a distribution with mean μ and variance σ^2, and $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Show that

$$\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

is an unbiased estimator of σ^2.

[You may assume the results $E(\bar{X}) = \mu$ and $\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}$.]

(7)

(b) The continuous random variable Y has probability density function

$$f(y) = \theta^{-2} y e^{-y/\theta}$$

for $y > 0$, where $\theta > 0$ is an unknown parameter, and $Y_1, Y_2, \ldots, Y_n$ are independent observations of Y.

(i) Find $E(Y)$, and use it to find the method of moments estimator $\hat{\theta}$ of θ based on $Y_1, Y_2, \ldots, Y_n$.

(6)

(ii) Show that $E(Y^2) = 6\theta^2$ and hence that the variance of $\hat{\theta}$ is $\frac{\theta^2}{2n}$.

(5)

(iii) Use the result of part (a) to find an unbiased estimator of the variance of $\hat{\theta}$ based on $Y_1, Y_2, \ldots, Y_n$.

(2)