EXAMINATIONS OF THE HONG KONG STATISTICAL SOCIETY

HIGHER CERTIFICATE IN STATISTICS, 2009

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* ${}^nC_r$.

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

$$f(x,y)=\begin{cases}\dfrac{1}{3\log 2}\left(\dfrac{x}{y}+\dfrac{y}{x}\right) & 1\le x\le 2, 1\le y\le 2\\ 0 & \text{otherwise}\end{cases}.$$

(i) Find the marginal probability density function of X.

(4)

(ii) Show that $E(X)\approx 1.4991$ and $\text{Var}(X)\approx 0.0847$.

(5)

(iii) Show that $E(XY)\approx 2.2442$.

(4)

(iv) Find the covariance of X and Y.

(3)

(v) Find the conditional probability density function $f(y\,|\,x)$, for $1\le x\le 2, 1\le y\le 2$, and hence evaluate $P(Y<1.5\,|\,X=1)$.

(4)

2. The random variable X has the χ_k^2 distribution (k = 1, 2, 3, …), which has moment generating function (mgf) $m(t)=\left(1-2t\right)^{-k/2}$ for $t<\frac{1}{2}$.

(i) Using the mgf, find the mean and variance of X.

(6)

(ii) In the case k = 4, the probability density function is given by

$$f(x)=\tfrac{1}{4}xe^{-x/2} \qquad (x>0).$$

Using integration, confirm that the mgf of the χ_4^2 distribution is $m(t)=\left(1-2t\right)^{-2}$ (for $t<\frac{1}{2}$), as given by the above formula.

(5)

(iii) Show that if $Y_1, Y_2, \dots, Y_n$ are independent, each with a χ_1^2 distribution, then $V=\sum_{i=1}^{n}Y_i$ has a χ_n^2 distribution.

(4)

(iv) Use the previous results and the central limit theorem to find the approximate probability that $V\le 310$ when n = 300.

(5)

3. For a productive pair from a particular species of bird, the number X of eggs laid per season has the probability mass function

$$P(X = k) = C\frac{e^{-\lambda}\lambda^k}{k!} \quad (k = 1, 2, 3, \dots),$$

where C is a constant and λ (> 0) is an unknown parameter.

(i) Show that $C = \dfrac{1}{1-e^{-\lambda}}$.

(4)

(ii) The numbers of eggs in a random sample of n nests of productive pairs are X_1, X_2, …, X_n. Find $\ell(\lambda)$, the logarithm of the likelihood of λ based on this sample, and find an equation satisfied by $\hat{\lambda}$, the maximum likelihood estimator. (Do not attempt to solve this equation.)

(5)

(iii) Find the approximate variance of the maximum likelihood estimator of λ for the random sample $X_1, X_2, \dots, X_n$, when n is large.

(5)

(iv) Plot the first derivative of $\ell(\lambda)$ at λ = 2.0, 2.5, 3.0 and 3.5 for the case $n = 10$ and $\sum X_i = 30$. Using your diagram, find an approximate value of the maximum likelihood estimator.

(6)

Turn over

4. The random variable Y has the geometric distribution, parameter p $(0 < p < 1)$, i.e.

$$P(Y = y) = (1 - p)^y p \qquad \text{for } y = 0, 1, 2, \ldots .$$

This distribution has probability generating function

$$\pi(t) = \frac{p}{1-(1-p)t} \qquad \text{for } t < (1-p)^{-1} .$$

(i) Using the probability generating function, or otherwise, show that the mean of this distribution is $\frac{1-p}{p}$ and the variance is $\frac{1-p}{p^2}$.

(5)

(ii) The random variables $Y_1, Y_2, \ldots, Y_n$ constitute a random sample from this distribution.

Define $\bar{Y} = \sum Y_i / n$. Show that $\bar{Y}$ is a biased estimator of $1/p$. Hence find an unbiased estimator of $1/p$ and show that it is also a consistent estimator of $1/p$.

(6)

(iii) Find the method of moments estimator of p.

(4)

(iv) The random variable W is the number of the random variables $Y_1, Y_2, \ldots, Y_n$ that take the value zero. (For example, if $n = 5$, $Y_1 = 1$, $Y_2 = 0$, $Y_3 = 3$, $Y_4 = 0$ and $Y_5 = 2$, then $W = 2$.) State the distribution of W. Hence find an unbiased estimator of p based on W and give its variance. [The formulae for the mean and variance of standard distributions may be assumed.]

(5)