EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY

HIGHER CERTIFICATE IN STATISTICS, 2010

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 3 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. Two tennis players, A and B, are playing a match. Let X be the number of serves faster than 125 mph served by A in one of his service games and let Y be the number of these serves returned by B. The following probability model is proposed:

$$P(X = 0) = 0.4,\ P(X = 1) = 0.3,\ P(X = 2) = 0.2 \text{ and } P(X = 3) = 0.1.$$

The conditional distribution of Y (given that $X = x > 0$) is binomial with parameters x and 0.4, and $P(Y = 0 \mid X = 0) = 1$. Assume that this model is correct when answering the following questions.

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

(7)

(ii) Find the marginal distribution of Y and evaluate $E(Y)$.

(4)

(iii) Find $\mathrm{Cov}(X, Y)$.

(4)

(iv) Use your joint probability distribution table to find the probability distribution of the number of serves faster than 125 mph that are not returned by B in a game.

(5)

2. The joint probability density function of the random variables X and Y is

$$f(x, y) = \frac{1}{2\pi} \exp\left(-\tfrac{1}{4}(x-1)^2 - \left(y - \tfrac{1}{4}(1+x)\right)^2\right),\quad -\infty < x < \infty,\ -\infty < y < \infty.$$

(i) Use integration to show that X has the Normal distribution with mean 1 and variance 2.

(7)

(ii) Use integration to show that the moment generating function of X is $m_X(t) = \exp(t + t^2)$.

(7)

(iii) Use the moment generating function to find $E(X^3)$.

(6)

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

$$f(x) = \beta(1-x)^{\beta-1}, \quad 0 < x < 1,$$

where $\beta\,(> 0)$ is an unknown parameter.

(i) Find the maximum likelihood estimator, $\hat{\beta}$, of β.

(7)

(ii) Calculate the approximate variance of $\hat{\beta}$ and use it to determine an approximate 95% confidence interval for β when n is large.

(6)

(iii) Show that $P(X_i < 0.5) = 1 - 0.5^{\beta}$.

(3)

(iv) Suppose now that the values of $X_1, X_2, \ldots, X_n$ are not known, but you do know Y, the number of the X_i less than 0.5. State the distribution of Y, and write down the likelihood function of β based on Y.

(4)

4. (a) Define the *bias*, *relative efficiency* and *efficiency* of potential estimators of a population parameter, and explain briefly why these are useful when deciding between different estimators.

(6)

(b) The random variables $Y_1, Y_2, \ldots, Y_n$ constitute a random sample from a discrete distribution, with $P(Y_i = -2) = \frac{1}{2}(1-p)^2$, $P(Y_i = 0) = p(1-p)$, $P(Y_i = 1) = p$, $P(Y_i = 2) = \frac{1}{2}(1-p)^2$ and $P(Y_i = k) = 0$ for $k \neq -2, 0, 1$ or 2, where p $(0 < p < 1)$ is an unknown parameter.

(i) Find the method of moments estimator, $\tilde{p}$, of p. Describe one unsatisfactory feature that this estimator possesses.

(5)

(ii) Find $E(Y_i^2)$ and hence obtain the variance of $\tilde{p}$.

(5)

(iii) Making use of results found in answering parts (i) and (ii), or otherwise, find an unbiased estimator of p^2 based on ΣY_i and ΣY_i^2.

(4)