EXAMINATIONS OF THE ROYAL STATISTICAL SOCIETY
(formerly the Examinations of the Institute of Statisticians)

GRADUATE DIPLOMA, 2006

Statistical Theory and Methods II

Time Allowed: Three Hours

Candidates should answer **FIVE** *questions.*

All questions carry equal 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 5 printed pages, **each printed on one side only**.
This front cover is page 1.
Question 1 starts on page 2.

There are 8 questions altogether in the paper.

1. Let $X_1, X_2, ..., X_n$ be a random sample from a Poisson population with mean θ.

(i) Show that $T = \sum X_i$ is a sufficient statistic for θ.

(6)

(ii) Find the maximum likelihood estimator (MLE) of θ. Hence deduce the MLE of $e^{-\theta}$.

(6)

(iii) Consider the estimator $\left(\frac{n-1}{n}\right)^T$. Show that it is an unbiased estimator of $e^{-\theta}$.

(5)

(iv) By using the result that $\frac{n-1}{n} < e^{-1/n}$, show that the expected value of the MLE of $e^{-\theta}$ exceeds $e^{-\theta}$.

(3)

2. Let $X_1, X_2, ..., X_n$ be a random sample from a continuous population that is uniformly distributed on the interval (0, θ), where the parameter θ is positive. You are given that this distribution has mean $\frac{\theta}{2}$ and variance $\frac{\theta^2}{12}$.

(i) Outline why the Cramér-Rao lower bound for the variance of unbiased estimators of θ does not apply in this case.

(4)

Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and let $Y = \max X_i$. Consider the three estimators of θ:

$$\hat{\theta}_1 = 2\bar{X} ; \qquad \hat{\theta}_2 = Y ; \qquad \hat{\theta}_3 = \left(\frac{n+1}{n}\right)Y .$$

(ii) Find the bias and variance of each of these estimators. [You are given that the probability density function of Y is $g(y) = \frac{ny^{n-1}}{\theta^n}$, $0 < y < \theta$.]

(7)

(iii) Hence discuss the consistency of each of the estimators.

(4)

(iv) Calculate the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_3$. What happens to the efficiency as n grows large?

(5)

Turn over

3. A random sample $X_1, X_2, ..., X_n$ is drawn from a discrete population with probability function

$$p(x;\theta) = \frac{(x+1)\theta^2}{(1+\theta)^{x+2}},$$

where $x = 0, 1, 2, \ldots$ and $\theta > 0$ is a parameter.

(i) Show that $E(X) = \dfrac{2}{\theta}$ and hence find the method of moments estimator of θ.

(6)

(ii) Obtain the maximum likelihood estimator $\hat{\theta}$ of θ.

(5)

(iii) Calculate the Fisher information for θ.

(5)

(iv) Suppose that $n = 100$ and $\hat{\theta} = 2.5$. Calculate an approximate 90% confidence interval for θ.

(4)

4. A random sample $x_1, x_2, ..., x_n$ is drawn from a population with probability density

$$f(x) = \begin{cases} A\exp(-x^2/2), & x \le 0, \\ A\exp(-x^2/2\sigma^2), & x > 0, \end{cases}$$

where $\sigma > 0$ is a parameter and $A = \dfrac{\sqrt{2}}{(1+\sigma)\sqrt{\pi}}$.

(i) What is the well-known distribution when $\sigma = 1$?

(2)

(ii) A test is required of the null hypothesis $H_0: \sigma = 1$ against the alternative $H_1: \sigma = \sigma^*$, where $\sigma^* > 1$ is a fixed value. Show that the Neyman-Pearson approach leads to the rejection of the null hypothesis when $\sum_+ x_i^2 > k$ for some suitable k, where $\sum_+ x_i^2$ is the sum of all x_i^2 values with $x_i > 0$.

(8)

(iii) Deduce that this test is uniformly most powerful against the alternative hypothesis $H_2: \sigma > 1$.

(4)

(iv) Perform the generalised likelihood ratio test of $H_0: \sigma = 1$ against the alternative $H_3: \sigma \ne 1$ when $n = 30$, $\sum_+ x_i^2 = 80.0$ and $\hat{\sigma} = 2.0$. Here $\hat{\sigma}$ is the unrestricted maximum likelihood estimator of σ, and you should use an approximate 5% significance level.

(6)

5. Large batches of widgets are produced and the following 3-stage sampling scheme is used to determine whether a batch should be accepted or rejected. At the first stage 30 widgets are randomly sampled. The batch is accepted if no widgets are defective, rejected if two or more are defective, and otherwise the stage 2 sample is taken. At stage 2, a further 30 widgets are randomly sampled. The batch is accepted if there are no defectives in this second sample, rejected if 3 or more in the second sample are defective, and otherwise the stage 3 sample is taken. At stage 3, 30 more widgets are randomly selected. The batch is accepted if, across all 90 widgets sampled, at most 3 are defective, and is rejected otherwise. The proportion of defective items in a batch is denoted by p, and $\lambda = 30p$. You should use the Poisson approximation to the binomial distribution in what follows.

(i) Calculate in terms of λ the expected number of widgets sampled per batch.

(8)

(ii) Find the value of p that maximises the probability that all three stages are needed for a batch.

(6)

(iii) Calculate in terms of λ the probability that the batch is accepted.

(6)

6. (a) Explain what is meant by a conjugate family of distributions.

(5)

(b) Let $X_1, X_2, ..., X_n$ be a random sample from a Poisson distribution with mean θ. Suppose that the prior distribution for θ is gamma with parameters $\alpha = 2$ and $\beta = ½$.

[You are given that a gamma random variable Y with parameters $\alpha > 0$ and $\beta > 0$ has probability density function

$$g(y) = \frac{\beta^\alpha y^{\alpha-1} e^{-\beta y}}{\Gamma(\alpha)}, \qquad y > 0,$$

where $\Gamma(.)$ denotes the gamma function. Its moment generating function is $\{\beta/(\beta - t)\}^\alpha$ for $t < \beta$ and, if α is a positive integer, then $2\beta Y$ has the $\chi^2_{2\alpha}$ distribution.]

(i) Obtain the posterior distribution of θ.

(5)

(ii) Assuming a squared error loss function, find the Bayes estimator of θ.

(5)

(iii) For the case when $n = 3$ and $\sum X_i = 8$, find a 95% Bayesian confidence interval for θ.

(5)

7. (a) Explain what is meant by a pivotal quantity. Describe how such a quantity may be used to obtain a confidence set for a parameter.

(6)

(b) A single observation X is taken from a distribution with probability density function

$$f(x) = \theta x^{\theta - 1}, \qquad 0 < x < 1,$$

where $\theta > 0$ is a parameter.

(i) Show that $-\theta \log X$ is a pivotal quantity.

(5)

(ii) Use this result to construct a 90% confidence interval for θ.

(5)

(iii) Hence calculate a 90% confidence interval for θ^{-1} when the observed value of X is 0.5.

(4)

8. Describe the various criteria that may be used for deciding between estimators. Discuss the importance of each of these criteria.

(20)