MTH6136 / MTH6136P: Statistical Theory

Duration: 2 hours

Date and time: 31st May 2016, 10:00–12:00

Apart from this page, you are not permitted to read the contents of this question paper until instructed to do so by an invigilator.

You should attempt ALL questions. Marks awarded are shown next to the questions.

Calculators ARE permitted in this examination. The unauthorised use of material stored in pre-programmable memory constitutes an examination offence. Please state on your answer book the name and type of machine used.

Complete all rough workings in the answer book and cross through any work that is not to be assessed.

Possession of unauthorised material at any time when under examination conditions is an assessment offence and can lead to expulsion from QMUL. Check now to ensure you do not have any notes, mobile phones, smartwatches or unauthorised electronic devices on your person. If you do, raise your hand and give them to an invigilator immediately. It is also an offence to have any writing of any kind on your person, including on your body. If you are found to have hidden unauthorised material elsewhere, including toilets and cloakrooms it shall be treated as being found in your possession. Unauthorised material found on your mobile phone or other electronic device will be considered the same as being in possession of paper notes. A mobile phone that causes a disruption in the exam is also an assessment offence.

Exam papers must not be removed from the examination room.

Examiner(s): Dr Dudley Stark, Dr Joaquin Miguez

Question 1. Suppose that $Y_1, Y_2, \ldots, Y_n$ are independent geometric random variables such that $P(Y = y) = p(1-p)^{y-1}$ for $y = 1, 2, \ldots$. You may assume that the Y_i have expectation $E(Y_i) = 1/p$.

(a) Find the method of moments estimator for p. [**5**]

(b) Show that the distribution of the Y_i belongs to the exponential family of distributions. [**5**]

(c) Using (b) and a theorem stated in a lecture, show that $\sum_{i=1}^n Y_i$ is a complete, sufficient statistic for p and that $\overline{Y} = \frac{1}{n}\sum_{i=1}^n Y_i$ is a Minimum Variance Unbiased Estimator for $\phi(p) = 1/p$. [**10**]

Question 2. Suppose that $Y_1, Y_2, \ldots, Y_n$ are independent normal random variables with zero mean and variance θ, where $\theta > 0$.

(a) Use Neyman's factorisation theorem to show that $\sum_{i=1}^n Y_i^2$ is a sufficient statistic for θ. [**6**]

(b) Show that the Cramér-Rao lower bound for unbiased estimators of θ is given by $\text{CRLB}(\theta) = 2\theta^2/n$. [**8**]

(c) Prove that $T_n = (Y_2^2 + \cdots + Y_{n-1}^2)/(n-2)$ is an unbiased estimator of θ. Given that $\text{Var}(Y^2) = 2\theta^2$, find the efficiency of T_n and show that the sequence of estimators T_n is asymptotically efficient. [**6**]

Question 3. Suppose that $Y_1, Y_2, \ldots, Y_n$ are independent Weibull random variables with probability density function

$$f_Y(y) = \frac{3y^2}{\theta^3} \exp\left(-\frac{y^3}{\theta^3}\right), \quad y > 0,$$

where $\theta > 0$.

(a) Show that the maximum likelihood estimator of θ is

$$\hat{\theta} = \left(\frac{1}{n}\sum_{i=1}^{n} Y_i^3\right)^{1/3}.$$

[6]

(b) Given that $E(Y^3) = \theta^3$, show that the Fisher information $E\left(-\frac{\partial^2}{\partial\theta^2} \log L(\theta; Y_1, \ldots, Y_n)\right)$ of $Y_1, Y_2, \ldots, Y_n$ equals $9n/\theta^2$. **[4]**

(c) State a theorem about the limiting distribution (including its parameters) of maximum likelihood estimators as $n \to \infty$. **[4]**

(d) Use your results from parts (b) and (c) to obtain the asymptotic distribution of $\hat{\theta}$, and hence find an approximate $100(1-\alpha)\%$ confidence interval for θ. **[6]**

Question 4. Let $Y_1, \ldots, Y_n$ be independent Uniform$(0, \theta)$ distributed random variables.

(a) Define what is meant by a *pivot* for θ. **[6]**

(b) Let M be the maximum of $Y_1, \ldots, Y_n$. Show that $P(M \leq x) = (x/\theta)^n$ for $0 \leq x \leq \theta$. **[4]**

(c) Use (b) to show that M/θ is a pivot for θ. **[4]**

(d) Use part (c) to derive an exact 95% confidence interval for θ based on M. **[6]**

Question 5. Suppose that $Y_1, \ldots, Y_n$ are independent Pascal random variables with probability mass function

$$P(Y = y) = \binom{y + r - 1}{r - 1} \pi^r (1 - \pi)^y, \quad y = 0, 1, \ldots,$$

where $0 < \pi < 1$. Consider testing $H_0 : \pi = \pi_0$ against $H_1 : \pi \neq \pi_0$.

(a) Write down the likelihood, $L(\pi; \underline{y})$, and hence find the generalised likelihood ratio given by $\Lambda(\underline{y}) = L(\hat{\pi}_0; \underline{y})/L(\hat{\pi}; \underline{y})$, where $\hat{\pi}_0$ is the restricted maximum likelihood estimate of π under H_0 and $\hat{\pi}$ is the maximum likelihood estimate. [**9**]

(b) State the critical region of the generalised likelihood ratio test in terms of $\Lambda(\underline{y})$ and explain why this only depends on the data through a sufficient statistic. [**4**]

(c) Use Wilks' theorem to obtain the critical region of a test with approximate significance level α for large n. [**7**]

End of Paper.