Queen Mary
University of London

Main Examination period 2019

MTH6136 / MTH6136P: Statistical Theory

Duration: 2 hours

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 available are shown next to the questions.

Only non-programmable calculators that have been approved from the college list of non-programmable calculators are permitted in this examination. Please state on your answer book the name and type of machine used.

Complete all rough work 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.

Examiners: Dr D. S. Stark, Dr D. S. Coad

Question 1. [20 marks] Suppose that $Y_1, \ldots, Y_n$ are independent binomial random variables such that

$$P(Y = k) = \binom{m}{k} p^k (1-p)^{m-k} \text{ for } k = 0, 1, \ldots,$$

where m is known. You may assume that the Y_i have expectation $E(Y_i) = mp$. Let $\overline{Y}$ denote

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i.$$

(a) Show that the method of moments estimator for p is $\overline{Y}/m$. **[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 $\overline{Y}/m$ is a Minimum Variance Unbiased Estimator for $\phi(p) = p$. **[10]**

Question 2. [**20 marks**] Suppose that $Y_1, \ldots, Y_n$ are independent inverse gamma random variables with probability density function

$$f_Y(y) = \frac{\beta^3}{2y^4} e^{-\frac{\beta}{y}}, \quad y > 0,$$

where $\beta > 0$.

(a) Show that the Cramér-Rao lower bound for unbiased estimators of β is

$$\frac{\beta^2}{3n}.$$

[**7**]

(b) Given that $E(Y) = \beta/2$ and $\text{var}(Y) = \beta^2/4$, show that the Mean Square Error of the estimator of β

$$T_n(\underline{Y}) = \frac{2}{n+1} \sum_{i=1}^{n} Y_i$$

is

$$\frac{\beta^2}{n+1}$$

and that the sequence of estimators $T_n(\underline{Y})$ is consistent. [**8**]

(c) Use Neyman's Factorisation Lemma to show that

$$\sum_{i=1}^{n} \frac{1}{Y_i}$$

is a sufficient statistic for β. [**5**]

Question 3. [**20 marks**] Suppose that $Y_1, \ldots, Y_n$ are independent gamma distributed random variables with mean 2θ and probability density function

$$f_Y(y) = \frac{y}{\theta^2} e^{-\frac{y}{\theta}}, \quad y > 0,$$

where $\theta > 0$.

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

$$\hat{\theta} = \frac{1}{2n} \sum_{i=1}^{n} Y_i$$

. [**7**]

(b) Obtain the asymptotic distribution of $\hat{\theta}$, and hence write down an approximate $100(1-\alpha)\%$ confidence interval for θ. [**8**]

(c) Given that the random variable $X = Y/\theta$ is $\Gamma(2, 1)$ with probability density function $f_X(x) = xe^{-x}$ for $x > 0$, explain why

$$\frac{1}{\theta} \sum_{i=1}^{n} Y_i$$

is an exact pivot for θ and derive an exact $100(1-\alpha)\%$ confidence interval for θ. You may use a fact about sums of independent gamma random variables stated in a lecture. [**5**]

Question 4. [20 marks]

(a) Suppose that $Y_1, \ldots, Y_{n_1}$ are $\mathrm{N}(\mu_1, \sigma^2)$ random variables and $Y_{n_1+1}, \ldots, Y_{n_1+n_2}$ are $\mathrm{N}(\mu_2, \sigma^2)$ random variables, all independent, where σ^2 is known.

(i) Show that the maximum likelihood estimators of μ_1 and μ_2 are

$$\hat{\mu}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} Y_i$$

and

$$\hat{\mu}_2 = \frac{1}{n_2} \sum_{i=n_1+1}^{n_1+n_2} Y_i.$$ **[7]**

(ii) State a pivot for $\mu_1 - \mu_2$ and give an exact $100(1-\alpha)\%$ confidence interval for $\mu_1 - \mu_2$. **[7]**

(b) Let $\underline{Y} = (Y_1, \ldots, Y_n)^{\mathrm{T}}$ where the Y_i are continuous random variables whose distribution has parameter θ and sample space which does not depend on θ and let $L(\theta; \underline{Y})$ be the associated likelihood function. Show that

$$\mathbb{E}\left(\frac{d \log L(\theta; \underline{Y})}{d\theta}\right) = 0.$$

[6]

Question 5. [20 marks] Suppose that $Y_1, \ldots, Y_n$ are independent Pascal random variables with probability mass function

$$P(Y = y) = \pi(1-\pi)^{y-1}, \quad y = 1, 2, \ldots,$$

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.