MATH20802

Two hours

To be supplied by the Examinations Office: Mathematical Formula Tables and Statistical Tables

THE UNIVERSITY OF MANCHESTER

STATISTICAL METHODS

24 May 2016

14:00-16:00

Answer FOUR of the SIX questions. If more than FOUR questions are attempted, then credit will be given for the best FOUR answers.

University-approved calculators may be used

---

1. Suppose X is a random variable with its probability density function given by

$$f(x) = \alpha\lambda \exp(-\lambda x) + (1-\alpha)\mu \exp(-\mu x)$$

for $x > 0$, $0 < \alpha < 1$, $\lambda > 0$ and $\mu > 0$.

(a) Show that the moment generating function of X is

$$M_X(t) = E\left[\exp(tX)\right] = \frac{\alpha\lambda}{\lambda - t} + \frac{(1-\alpha)\mu}{\mu - t}.$$

(8 marks)

(b) Use your result in (a) to derive the first four moments of X. (8 marks)

(c) If X_i are independent and identical random variables and are distributed as X derive the moment generating function of $Y = X_1 + \cdots + X_n$. (3 marks)

(d) Derive the mean and variance of Y. (3 marks)

(e) What is the distribution of Y if $\lambda = \mu$? (3 marks)

[Total: 25 marks]

2. (a) Suppose $\widehat{\theta}$ is an estimator of θ based on a random sample of size n. Define what is meant by the following:

(i) $\widehat{\theta}$ is an unbiased estimator of θ; (2 marks)

(ii) $\widehat{\theta}$ is an asymptotically unbiased estimator of θ; (2 marks)

(iii) the bias of $\widehat{\theta}$ (written as $\text{bias}\left(\widehat{\theta}\right)$); (2 marks)

(iv) the mean squared error of $\widehat{\theta}$ (written as $\text{MSE}\left(\widehat{\theta}\right)$); (2 marks)

(v) $\widehat{\theta}$ is a consistent estimator of θ. (2 marks)

(b) Suppose $X_1, \ldots, X_n$ are independent Uniform$[0, \theta]$ random variables. Let $\widehat{\theta_1} = \frac{2(X_1+\cdots+X_n)}{n}$ and $\widehat{\theta_2} = \max\left(X_1, \ldots, X_n\right)$ denote possible estimators of θ.

(i) Derive the bias and mean squared error of $\widehat{\theta_1}$; (4 marks)

(ii) Derive the bias and mean squared error of $\widehat{\theta_2}$; (7 marks)

(iii) Which of the estimators ($\widehat{\theta_1}$ and $\widehat{\theta_2}$) is better with respect to bias and why? (1 marks)

(iv) Which of the estimators ($\widehat{\theta_1}$ and $\widehat{\theta_2}$) is better with respect to mean squared error and why? (3 marks)

[Total: 25 marks]

3. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a distribution specified by the probability density function $f(x) = \frac{1}{2a} \exp\left(-\frac{|x|}{a}\right)$, $-\infty < x < \infty$, where $a > 0$ is an unknown parameter.

(a) Write down the likelihood function of a. (5 marks)

(b) Show that the maximum likelihood estimator of a is $\widehat{a} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} |X_i|$. (6 marks)

(c) Derive the expected value of $\widehat{a}$ in part (b). (6 marks)

(d) Derive the variance of $\widehat{a}$ in part (b). (6 marks)

(e) Show that $\widehat{a}$ is an unbiased and consistent estimator for a. (2 marks)

[Total: 25 marks]

4. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$. Suppose $Y_1, Y_2, \ldots, Y_m$ is a random sample from $LN(\mu, \sigma^2)$ independent of $X_1, X_2, \ldots, X_n$. Assume both μ and σ^2 are unknown.

(a) Write down the joint likelihood function of μ and σ^2 based on all the data $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$. (5 marks)

(b) Show that the maximum likelihood estimator of μ is $\displaystyle\frac{1}{m+n}\left[\sum_{i=1}^{n} X_i + \sum_{i=1}^{m} \log Y_i\right]$. (5 marks)

(c) Show that the maximum likelihood estimator of σ^2 is $\displaystyle\frac{1}{m+n}\left[\sum_{i=1}^{n} \left(X_i - \widehat{\mu}\right)^2 + \sum_{i=1}^{m} \left(\log Y_i - \widehat{\mu}\right)^2\right]$. (5 marks)

(d) Show that the estimator in (b) is unbiased and consistent for μ. (5 marks)

(e) If $X \sim N(\mu, \sigma^2)$ and $Y \sim LN(\mu, \sigma^2)$ are independent random variables find the maximum likelihood estimator of $\Pr(2X < \log Y)$. (5 marks)

[Total: 25 marks]

5. (a) Suppose we wish to test $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$. Define what is meant by the following:

(i) the Type I error of a test; (1 marks)

(ii) the Type II error of a test; (1 marks)

(iii) the significance level of a test; (1 marks)

(iv) the power function of a test (denoted $\Pi(\mu)$). (1 marks)

(b) Suppose $X_1, X_2, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$, where σ^2 is assumed known. State the rejection region for each of the following tests:

(i) $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$; (2 marks)

(ii) $H_0 : \mu = \mu_0$ versus $H_1 : \mu < \mu_0$. (2 marks)

In each case, assume a significance level of α.

(c) Under the same assumptions as in part (b), find the power function, $\Pi(\mu)$, for each of the tests:

(i) $H_0 : \mu = \mu_0$ versus $H_1 : \mu \neq \mu_0$; (6 marks)

(ii) $H_0 : \mu = \mu_0$ versus $H_1 : \mu < \mu_0$. (3 marks)

In each case, you should express the power function, $\Pi(\mu)$, in terms of $\Phi(\cdot)$, the standard normal distribution function.

(d) Are the power functions in (c) most powerful in the sense of the Neyman-Pearson lemma? Justify your answer. (8 marks)

[Total: 25 marks]

6. (a) Describe the Neyman-Pearson test for $H_0 : \theta = \theta_1$ versus $H_1 : \theta = \theta_2$ based on a random sample $X_1, X_2, \ldots, X_n$ from a distribution with the probability density function $f(x; \theta)$. State both the test statistic and the rejection region. (4 marks)

(b) Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a distribution specified by the probability density function $f(x) = \exp(\theta - x)$, $x > \theta > 0$, where θ is unknown.

(i) Derive the Neyman-Pearson test for $H_0 : \theta = \theta_1$ versus $H_1 : \theta = \theta_2$, where $\theta_1 < \theta_2$. Show that the rejection region is either $\min(X_1, \ldots, X_n) > \theta_2$ or the empty set. (6 marks)

(ii) Another test for $H_0 : \theta = \theta_1$ has rejection region

$$\min(X_1, \ldots, X_n) > c.$$

Assuming $c > \theta_1$ is a constant, show that the power function is

$$\Pi(\theta) = \begin{cases} \exp(n\theta - nc), & \text{if } \theta < c, \\ 1, & \text{if } \theta \geq c. \end{cases}$$

(6 marks)

(iii) For the test in part (ii), find the value of c if $n = 10$, $\theta_1 = 1$ and the probability of type I error is 0.05. (4 marks)

(iv) For the test in part (ii), find the probability of a type II error if $n = 10$, $\theta_1 = 1$, $\theta_2 = 2$ and the probability of type I error is 0.05. (5 marks)

[Total: 25 marks]

END OF EXAMINATION PAPER