MATH20802

Two hours

To be supplied by the Examinations Office: Mathematical Formula Tables

THE UNIVERSITY OF MANCHESTER

STATISTICAL METHODS

21 June 2010

9:45 – 11:45

Answer any FOUR of the questions.

University-approved calculators may be used

---

1. This question explores the moment generating function of the normal distribution.

(i) Show that $\exp(t^2/2)$ is the moment generating function of the standard normal distribution. (5 marks)

(ii) Find the first four moments of the standard normal distribution by differentiating the moment generating function. (4 marks)

(iii) Suppose that $Z \sim N(0,1)$ and let $Y = \mu + \sigma Z$ for some scalar $\sigma > 0$. Find the moment generating function of Y. This is the form of the moment generating function of the $N(\mu, \sigma^2)$ distribution. (4 marks)

(iv) Let $X_1 \sim N(\mu_1, \sigma_1^2)$, let $X_2 \sim N(\mu_2, \sigma_2^2)$, and assume that X_1 and X_2 are independent. Find the moment generating function of $S = X_1 + X_2$. (4 marks)

(v) Use the linearity of expectation to find $E(S)$ and $Var(S)$. (4 marks)

(vi) Use the uniqueness of the moment generating function to determine the distribution of S. (4 marks)

2. Suppose $\widehat{\theta}$ is an estimator of θ. 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 bias$(\widehat{\theta})$); (2 marks)

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

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

Let X_i denote the time that it takes student i to complete a take-home exam, and suppose that $X_1, X_2, \ldots, X_n$ constitute a random sample from an exponential distribution with parameter β. Consider the following estimators for $\theta = 1/\beta$: $\widehat{\theta}_1 = c \min(X_1, X_2, \ldots, X_n)$ and $\widehat{\theta}_2 = 1/n \sum_{i=1}^{n} X_i$.

(i) Determine the value of c (perhaps as a function of n) for which $\widehat{\theta}_1$ is an unbiased estimator of θ. (4 marks)

(ii) Determine the variance and MSE of the estimator, $\widehat{\theta}_1$, from (i) as a function of the parameter θ. (4 marks)

(iii) Determine the bias, variance, and MSE of $\widehat{\theta}_2$ as a function of the parameter θ. (3 marks)

(iv) Which of the two estimators ($\widehat{\theta}_1$ or $\widehat{\theta}_2$) is better and why? (4 marks)

3. Let the random variable Y_i be the number of typographical errors on page i of a 400-page book (for $i = 1, 2, \ldots, 400$), and suppose that the Y_i's are independent and identically distributed according to a Poisson distribution with parameter λ. Let the random variable X be the number of pages of this book that contain at least one typographical error. Suppose that you are told the value of X but are not told anything about the values of the Y_i.

(i) Identify the probability distribution of X and write down its probability mass function in terms of λ. (5 marks)

(ii) Write down the likelihood function of λ. (5 marks)

(iii) Find the maximum likelihood estimator (MLE) of λ by maximizing the log likelihood function. (5 marks)

(iv) Suppose that the data reveal that 25 of the 400 pages contain a typographical error. What is the MLE of λ? (5 marks)

(v) Show that you could have used the invariance property of maximum likelihood estimators to determine the MLE of λ. (5 marks)

4. Suppose $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables with the common probability density function (pdf):

$$f(x) = \theta_2 x^{\theta_2 - 1} \theta_1^{-\theta_2}$$

for $0 < x < \theta_1$, $\theta_1 > 0$ and $\theta_2 > 0$. Both θ_1 and θ_2 are unknown.

(i) Calculate the cumulative distribution function, mean and variance corresponding to the given pdf. (5 marks)

(ii) Write down the joint likelihood function of θ_1 and θ_2. (5 marks)

(iii) Determine the maximum likelihood estimator (MLE) of θ_1. (5 marks)

(iv) Determine the MLE of θ_2. (5 marks)

(v) Show that the MLE, $\widehat{\theta_1}$, is a biased and consistent estimator for θ_1. (5 marks)

5. Suppose we wish to test $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$. Define what is meant by the following:

(i) the Type I error of the test. (2 marks)

(ii) the Type II error of the test. (2 marks)

(iii) the significance level of the test. (2 marks)

(iv) the power function of the test (denoted $\Pi(\theta)$). (2 marks)

Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a Bernoulli distribution with parameter p. State the rejection region for each of the following tests:

(i) $H_0 : p = p_0$ versus $H_1 : p \neq p_0$. (2 marks)

(ii) $H_0 : p = p_0$ versus $H_1 : p < p_0$. (2 marks)

(iii) $H_0 : p = p_0$ versus $H_1 : p > p_0$. (2 marks)

In each case, assume a significance level of α and that $\overline{X} = (X_1 + X_2 + \cdots + X_n)/n$ has an approximate normal distribution.

Under the same assumptions, find the power function, $\Pi(p)$, for each of the tests:

(i) $H_0 : p = p_0$ versus $H_1 : p \neq p_0$. (4 marks)

(ii) $H_0 : p = p_0$ versus $H_1 : p < p_0$. (4 marks)

(iii) $H_0 : p = p_0$ versus $H_1 : p > p_0$. (3 marks)

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

6. State 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 $f(x; \theta)$. (5 marks)

Let $X_1, X_2, \ldots, X_n$ be a random sample from a Bernoulli distribution with parameter p.

(i) Find the most powerful test at level α for $H_0 : p = p_0$ versus $p = p_1$, where $p_1 > p_0$ are constants. Show that the test rejects H_0 if and only if $\sum_{i=1}^{n} X_i > k$ for some k. (5 marks)

(ii) Determine the power function, $\Pi(p)$, of the test in part (i). (5 marks)

(iii) Find the value of k when $\alpha = 0.05$, $n = 5$, $p_0 = 0.5$ and $p_1 = 0.6$. (5 marks)

(iv) Find $\beta = \Pr$ (Type II error) when $n = 5$, $p_0 = 0.5$ and $p_1 = 0.6$. (5 marks)

END OF EXAMINATION PAPER