MATH20802

Two hours

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

THE UNIVERSITY OF MANCHESTER

STATISTICAL METHODS

23 May 2017

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. Let X denote a random variable with probability density function given by

$$f_X(x) = \frac{\exp\left(-\dfrac{x-\mu}{s}\right)}{s\left[1+\exp\left(-\dfrac{x-\mu}{s}\right)\right]^2}$$

for $-\infty < x < +\infty$, $-\infty < \mu < +\infty$ and $s > 0$. X is said to have a logistic distribution with parameters μ and s.

(a) Derive the cumulative distribution function of X. (6 marks)

(b) Show that the moment generating function of X can be expressed as

$$M_X(t) = E\left[\exp(tX)\right] = \exp\left(\mu t\right) B\left(1-st, 1+st\right)$$

for $-\frac{1}{s} < t < \frac{1}{s}$, where $B(a,b)$ denotes the beta function defined by

$$B\left(a,b\right) = \int_0^1 w^{a-1}(1-w)^{b-1} dw$$

for $a > 0$ and $b > 0$. (6 marks)

(c) Use your result in (b) to derive the first two moments of X. Express the moments in terms of the gamma function and its derivatives. (6 marks)

(d) Suppose Y and Z are independent unit exponential random variables (that is, $Y \sim \text{Exp}(1)$ and $Z \sim \text{Exp}(1)$). Derive the moment generating function of

$$W = -\log\left(\frac{Y}{Z}\right).$$

(5 marks)

(e) What is the distribution of W? (2 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 $\text{bias}(\widehat{\theta})$); (2 marks)

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

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

(b) Suppose X_1 and X_2 are independent $\text{Exp}(1/\lambda)$ random variables. Let $\widehat{\theta_1} = a\,(X_1 + X_2)$ and $\widehat{\theta_2} = b\sqrt{X_1 X_2}$ denote possible estimators of λ, where a and b are constants.

(i) Show that $a = 1/2$ if $\widehat{\theta_1}$ is to be an unbiased estimator of λ; (3 marks)

(ii) Show that $b = 4/\pi$ if $\widehat{\theta_2}$ is to be an unbiased estimator of λ; (3 marks)

(iii) Determine the variance of $\widehat{\theta_1}$; (3 marks)

(iv) Determine the variance of $\widehat{\theta_2}$; (3 marks)

(v) 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 the discrete uniform distribution on the set of integers $\{1, 2, \ldots, N\}$, where $N \geq 1$ is an unknown parameter.

(a) Write down the likelihood function of N. (3 marks)

(b) Show that the maximum likelihood estimator of N is $\max(X_1, X_2, \ldots, X_n)$. (3 marks)

(c) Derive the cumulative distribution function of $\max(X_1, X_2, \ldots, X_n)$ and hence show that the probability mass function of $\max(X_1, X_2, \ldots, X_n)$ is

$$\left(\frac{z}{N}\right)^n - \left(\frac{z-1}{N}\right)^n$$

for integer z.

(3 marks)

(d) Use part (c) to show that, as an estimator of N, $\max(X_1, X_2, \ldots, X_n)$ is biased but asymptotically unbiased. (8 marks)

(e) Show that $\max(X_1, X_2, \ldots, X_n)$ is a consistent estimator of N. (8 marks)

[Total: 25 marks]

4. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from a distribution specified by the probability density function

$$f_X(x) = \frac{1}{s} \exp\left(-\frac{x-\mu}{s}\right)$$

for $x \geq \mu$, $\mu > 0$ and $s > 0$. Assume both μ and s are unknown.

(a) Write down the joint likelihood function of μ and s. (3 marks)

(b) Show that the maximum likelihood estimator of μ is $\widehat{\mu} = \min(X_1, X_2, \ldots, X_n)$. (3 marks)

(c) Show that the maximum likelihood estimator of s is $\widehat{s} = \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) - \min(X_1, X_2, \ldots, X_n)$. (3 marks)

(d) Show that the estimator in (b) is biased for μ. (8 marks)

(e) Show that the estimator in (c) is also biased for s. (8 marks)

[Total: 25 marks]

5. (a) 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 a test. (2 marks)

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

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

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

(b) Let $X_1, X_2, \ldots, X_m$ be a random sample from a normal population with mean μ_X assumed unknown and variance σ_X^2 assumed known. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from a normal population with mean μ_Y assumed unknown and variance σ_Y^2 assumed known. Assume independence of the two samples. State the rejection region for each of the following tests:

(i) $H_0 : \mu_X = \mu_Y$ versus $H_1 : \mu_X \neq \mu_Y$. (2 marks)

(ii) $H_0 : \mu_X = \mu_Y$ versus $H_1 : \mu_X > \mu_Y$. (2 marks)

(iii) $H_0 : \mu_X = \mu_Y$ versus $H_1 : \mu_X < \mu_Y$. (2 marks)

In each case, assume a significance level of α.

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

(i) $H_0 : \mu_X = \mu_Y$ versus $H_1 : \mu_X \neq \mu_Y$. (5 marks)

(ii) $H_0 : \mu_X = \mu_Y$ versus $H_1 : \mu_X > \mu_Y$. (3 marks)

(iii) $H_0 : \mu_X = \mu_Y$ versus $H_1 : \mu_X < \mu_Y$. (3 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 form of 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

$$f(x) = \frac{1}{2\theta} \exp\left(-\frac{\mid x \mid}{\theta} \right)$$

for $-\infty < x < +\infty$ and $\theta > 0$, 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 rule for rejecting H_0 can be expressed as $\sum_{i=1}^{n} |X_i| < c$. (8 marks)

(ii) Show that the power function for the rejection rule in part (i) is

$$\Pi(\theta) = F_{\frac{1}{\theta}, n}(c),$$

where $F_{a,b}(\cdot)$ denotes the cdf of a gamma random variable with scale parameter a and shape parameter b (Hint: show that $\sum_{i=1}^{n} \mid X_i \mid$ is a gamma random variable with scale parameter $\frac{1}{\theta}$ and shape parameter n).

(8 marks)

(iii) Find the value of c in part (i) if $n = 1$, $\theta_1 = 1$ and the probability of a type I error is 0.05. (2 marks)

(iv) Find the probability of a type II error if $n = 1$, $\theta_1 = 1$, $\theta_2 = 0.5$ and the probability of a type I error is 0.05. (3 marks)

[Total: 25 marks]

END OF EXAMINATION PAPER