MATH20802

Two hours

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

THE UNIVERSITY OF MANCHESTER

STATISTICAL METHODS

Answer any FOUR of the SIX questions.

University-approved calculators may be used

---

1. A random variable X is said to have the Gumbel distribution, written $X \sim \text{Gumbel}(\mu, \beta)$, if its probability density function is given by

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

for $-\infty < x < \infty$, $-\infty < \mu < \infty$ and $\beta > 0$.

(i) Show that the cumulative distribution function of X is

$$F_X(x) = \exp\left\{-\exp\left(-\frac{x-\mu}{\beta}\right)\right\}$$

for $-\infty < x < \infty$. (7 marks)

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

$$M_X(t) = \exp(\mu t)\Gamma(1 - \beta t)$$

for $t < 1/\beta$, where $\Gamma(\cdot)$ denotes the gamma function. (7 marks)

(iii) Show that

$$E(X) = \mu - \beta\Gamma^{'}(1),$$

where $\Gamma^{'}(\cdot)$ denotes the first derivative of $\Gamma(\cdot)$. (4 marks)

(iv) If $X_i \sim \text{Gumbel}(\mu, \beta)$, $i = 1, 2, \ldots, n$ are independent random variables then show that $\max(X_1, X_2, \ldots, X_n) \sim \text{Gumbel}(\mu + \beta \log n, \beta)$. (7 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 $\mathrm{bias}(\widehat{\theta})$); (2 marks)

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

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

(b) Suppose $X_1, X_2, \ldots, X_n$ is a random sample from the Exp (λ) distribution. Consider the following estimators for $\theta = 1/\lambda$: $\widehat{\theta_1} = (1/n) \sum_{i=1}^{n} X_i$ and $\widehat{\theta_2} = (1/(n+1)) \sum_{i=1}^{n} X_i$.

(i) Find the biases of $\widehat{\theta_1}$ and $\widehat{\theta_2}$. (4 marks)

(ii) Find the variances of $\widehat{\theta_1}$ and $\widehat{\theta_2}$. (4 marks)

(iii) Find the mean squared errors of $\widehat{\theta_1}$ and $\widehat{\theta_2}$. (3 marks)

(iv) Which of the two estimators ($\widehat{\theta_1}$ or $\widehat{\theta_2}$) is better and why? (4 marks)

[Total: 25 marks]

3. Consider the two independent random samples: $X_1, X_2, \ldots, X_n$ from $N(\mu_X, \sigma^2)$ and $Y_1, Y_2, \ldots, Y_m$ from $N(\mu_Y, \sigma^2)$, where σ^2 is assumed known. The parameters μ_X and μ_Y are assumed not known.

(i) Write down the joint likelihood function of μ_X and μ_Y. (5 marks)

(ii) Find the maximum likelihood estimators (mles) of μ_X and μ_Y. (10 marks)

(iii) Find the mle of $\Pr(X < Y)$, where $X \sim N(\mu_X, \sigma^2)$ and $Y \sim N(\mu_Y, \sigma^2)$ are independent random variables. (5 marks)

(iv) Show that the mle of μ_X in part (ii) is an unbiased and consistent estimator for μ_X. (3 marks)

(v) Show also that the mle of μ_Y in part (ii) is an unbiased and consistent estimator for μ_Y. (2 marks)

[Total: 25 marks]

4. Suppose $X_1, X_2, \ldots, X_n$ is a random sample from $N(\mu, \sigma^2)$, where both μ and σ^2 are unknown.

(i) Write down the joint likelihood function of μ and σ^2. (5 marks)

(ii) Show that the maximum likelihood estimator (mle) of μ is $\widehat{\mu} = \bar{X}$, where $\bar{X} = (1/n) \sum_{i=1}^{n} X_i$ is the sample mean. (5 marks)

(iii) Show that the mle of σ^2 is $\widehat{\sigma^2} = (1/n) \sum_{i=1}^{n} (X_i - \bar{X})^2$. (5 marks)

(iv) Show that the mle, $\widehat{\mu}$, is an unbiased and consistent estimator for μ. (5 marks)

(v) Show that the mle, $\widehat{\sigma^2}$, is a biased and consistent estimator for σ^2. (5 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) 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.

(c) Under the same assumptions as part (b), 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.

[Total: 25 marks]

6. (a) 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)

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from a Uniform$(0, \theta)$ distribution.

(i) Find the most powerful test at significance level α for $H_0 : \theta = \theta_1$ versus $H_1 : \theta = \theta_2$, where $\theta_2 > \theta_1$ are constants. Show that the test rejects H_0 if and only if $\max(X_1, X_2, \ldots, X_n) > k$ for some k. (5 marks)

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

(iii) Find the value of k when $\alpha = 0.05$, $n = 5$ and $\theta = \theta_1 = 0.5$. (5 marks)

(iv) Find $\beta = \Pr$ (Type II error) when $n = 5$, $\theta_1 = 0.5$ and $\theta = \theta_2 = 0.6$. (5 marks)

[Total: 25 marks]

END OF EXAMINATION PAPER