MA1202
All candidates

Midsummer Examinations 2017

DO NOT OPEN THE QUESTION PAPER UNTIL INSTRUCTED TO DO SO BY THE CHIEF INVIGILATOR	
Department	MATHEMATICS
Module Code	MA1202
Module Title	INTRODUCTORY STATISTICS
Exam Duration	Two hours
CHECK YOU HAVE THE CORRECT QUESTION PAPER	
Number of Pages	4
Number of Questions	3
Instructions to Candidates	Answer all questions. All marks gained will be counted. All questions carry equal weight.
FOR THIS EXAM YOU ARE ALLOWED TO USE THE FOLLOWING:	
Calculators	Approved calculators may be used.
Books/Statutes provided by the University	Statistical tables.
Are students permitted to bring their own Books/Statutes/Notes?	No
Additional Stationery	No

1. In this question, X is a continuous random variable with density function

$$f(x) = \begin{cases} \frac{1}{\alpha} x^{(\alpha^{-1})-1} & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where α is an unknown parameter which is *strictly positive*. You wish to estimate α using observations $x_1, \ldots, x_n$ of an independent random sample $X_1, \ldots, X_n$ from X.

i) Write down the likelihood function $\mathcal{L}(a)$, simplifying your answer as much as possible. [2 marks]

ii) Show that the derivative of the log likelihood function $l(a)$ is

$$-\frac{n}{a} - a^{-2} \sum_{i=1}^{n} \log x_i$$

[4 marks]

iii) Show that the derivative $\frac{dl}{da}$ of the log likelihood function is equal to zero if and only if $a = \frac{1}{n} \sum_{i=1}^{n} - \log x_i$. [4 marks]

You may assume that this critical point is a maximum, so that the estimator $\hat{\alpha}$ defined by

$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} - \log X_i.$$

is a maximum likelihood estimator of α.

You are further told that $-\log X_i$ has the exponential distribution with parameter α^{-1}, and that if $Y \sim \text{Exponential}(\lambda)$ then $E(Y) = \lambda^{-1}$ and $V(Y) = \lambda^{-2}$.

iv) Show that $\hat{\alpha}$ is unbiased as an estimator for the parameter α in (1). [4 marks]

v) Calculate the mean squared error of $\hat{\alpha}$. [6 marks]

Total: 20 marks

2. A continuous random variable X has the $\text{Exponential}(\lambda)$ distribution if its density function is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

i) Prove that if X has the $\text{Exponential}(1/2)$ distribution, $a = \log((40/39)^2)$, and $b = \log 1600$, then

$$P(a < X < b) = 0.95.$$

[4 marks]

ii) Suppose Y is a random variable, θ is an unknown parameter and

$$10(Y + \theta) \sim \text{Exponential}(1/2).$$

Using $10(Y+\theta)$ as a pivot, and the result of part (i), find a 95% confidence interval for θ based on an observation $y = 5$ of Y. [6 marks]

iii) Suppose $Z \sim \text{Exponential}(1/5)$. In an independent random sample of size 100 from Z, calculate the expected number of observations in each of the following intervals:

a) $[0,2)$
b) $[2,4)$,
c) $[4,6)$,
d) $[6,8)$,
e) $[8,\infty)$.

[5 marks]

iv) You perform a statistical experiment in which you take an independent random sample of size 100 from a certain distribution. The following table records how many observations lie in each of the above intervals:

Interval	$[0,2)$	$[2,4)$	$[4,6)$	$[6,8)$	$[8,\infty)$
Number of observations	30	20	18	11	21.

Perform a χ^2 goodness of fit test of the hypothesis that the data is drawn from an $\text{Exponential}(1/5)$ distribution, at the 0.05 significance level, using the expected values from part (iii). What is your conclusion? [5 marks]

Total: 20 marks

3. An experiment is performed in which 15 people have their reaction times measured before and after drinking a cup of coffee. We write b_i for the before-coffee time of the ith person, and a_i for their after-coffee time. We regard b_i as an observation of a random variable B_i and a_i as an observation of a random variable A_i.

   i) Would it be reasonable to assume A_i and B_i are independent? Explain your answer. [3 marks]

   Let $D_i = B_i - A_i$. We assume that there are numbers δ, σ^2 such that for all i, $D_i \sim N(\delta, \sigma^2)$, and that the random variables $D_1, \ldots, D_{15}$ are independent. Let

$$\bar{D} = \frac{1}{15}\sum_{i=1}^{15} D_i$$
$$S = \sqrt{\frac{1}{14}\sum_{i=1}^{15}(D_i - \bar{D})^2}$$

   ii) If coffee tends to make people react faster, will δ be positive or negative? Why? [2 marks]

   iii) State the distribution of the random variable

$$\frac{\bar{D} - \delta}{S/\sqrt{15}}.$$

   [2 marks]

   iv) When the experiment is carried out, the observed values of $\bar{D}$ and S^2 are $\bar{d} = 0.98$ and $s^2 = 0.24$. Test the null hypothesis that $\delta = 0$ against the alternative that $\delta > 0$, at the 0.05 significance level. [7 marks]

   v) The distribution of $14S^2/\sigma^2$ is χ^2_{14}. Use this to find a 95% confidence interval for σ^2 based on the observation of s^2 given in part iv). [6 marks]

Total: 20 marks