MA1202
All candidates

UNIVERSITY OF LEICESTER

Midsummer Examinations 2018

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. (a) Let $\hat{\theta}$ be an estimator of an unknown parameter θ.

   i. Define the bias of $\hat{\theta}$, $\mathrm{bias}(\hat{\theta})$; [1 mark]

   ii. Define the mean squared error of $\hat{\theta}$, $\mathrm{MSE}(\hat{\theta})$; [1 mark]

   iii. Show that

$$\mathrm{MSE}(\hat{\theta}) = \mathrm{var}(\hat{\theta}) + (\mathrm{bias}(\hat{\theta}))^2.$$

[5 marks]

(b) A continuous random variable X has density function

$$f_X(x) = \begin{cases} 2\lambda^2 x \mathrm{e}^{-(\lambda x)^2}, & x \geq 0, \\ 0, & x < 0, \end{cases}$$

where $\lambda > 0$. Suppose $x_1, \ldots, x_n$ are observations of independent variables $X_1, \ldots, X_n$, respectively, all with the same distribution as X.

   i. Find the log-likelihood function $l(\mu)$ for this sample.
   [Note, μ is a variable from which the estimator $\hat{\lambda}$ of λ is selected.] [5 marks]

   ii. Hence, show that the maximum likelihood estimate $\hat{\lambda}$ of λ is

$$\hat{\lambda} = \left(\frac{n}{\sum_{i=1}^n x_i^2} \right)^{1/2}.$$

   [Hint: be careful to check that this is actually a maximum of $l(\mu)$.] [8 marks]

   iii. Now consider the random variable

$$L = \frac{\sum_{i=1}^n X_i^2}{n}.$$

   Assuming $\lambda^2 X^2$ has mean and variance 1, show that L is an unbiased estimator for λ^{-2} and hence, find $\mathrm{MSE}(L)$.
   Does it necessarily follow that $\hat{\lambda}$ is an unbiased estimator for λ? [5 marks]

Total: 25 marks

2. Let X be a continuous random variable with density function

$$f_X(x) = \begin{cases} \frac{2}{(x+1)^2}, & 0 \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the cumulative distribution $F_X(x)$ and hence, show that, if $a = \frac{1}{39}$ and $b = \frac{19}{21}$,

$$\mathrm{P}(a < X < b) = 0.9.$$

[8 marks]

(b) Let Y be a random variable such that, for unknown θ, the pivot

$$\frac{\theta - Y^2}{\theta}$$

has the same distribution as random variable X. Using the result from (a), find a 90% confidence interval for θ based on the single observation of Y, $y = -2$. [4 marks]

(c) For the random variable X, calculate the probability of X lying in each of the following intervals.

i. $[0, 0.2)$;
ii. $[0.2, 0.4)$;
iii. $[0.4, 0.6)$;
iv. $[0.6, 0.8)$;
v. $[0.8, 1]$.

[5 marks]

(d) A statistical experiment is performed in which an independent random sample of size 100 from a certain distribution is taken. The following table records how many observations lie in each of the same intervals from part (c):

Interval	$[0, 0.2)$	$[0.2, 0.4)$	$[0.4, 0.6)$	$[0.6, 0.8)$	$[0.8, 1]$
Frequency	40	31	12	10	7

Using the expected values from part (c), perform a χ^2 goodness of fit test of the hypothesis “the data is drawn from the X distribution”, at the 0.1 significance level. What is the conclusion? [8 marks]

Total: 25 marks

3. (a) Let X_1 and X_2 be two random variables with distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively. Assuming that X_1 and X_2 are independent and $X_1 - X_2 \sim N(\mu, \sigma^2)$, show that $\mu = \mu_1 - \mu_2$ and $\sigma^2 = \sigma_1^2 + \sigma_2^2$. [5 marks]

(b) Suppose we are carrying out a hypothesis test at the α-significance level. Let H_0 be the null hypothesis and H_1 be the alternative hypothesis. Define

i. the type I error;
ii. the type II error;
iii. the power of the test.

[3 marks]

(c) A statistics module has been running for many years and, in the past, it has been found that each year the number of students passing the exam has distribution $\mathrm{Bi}(n, 0.75)$, where n are the number of students taking the module that year.

A lecturer is teaching the module for the first time and 105 out of 150 students pass the exam. Perform a hypothesis test at the 0.05-significance level, where the null hypothesis is "The probability of a student passing the module is 0.75" and the alternative hypothesis is "The probability of a student passing the module is less than 0.75". What is the conclusion?

[Hint: Clearly state any assumptions made and recall the conditions under which a binomial distribution can be approximated by a normal distribution.] [10 marks]

(d) At another university, 300 students are taking a statistics module. Two lecturers A and B each teach 150 students. After the exam has been taken, 98 of lecturer A's students have passed, while 92 of lecturer B's students have passed.

Assuming that, for each lecturer, the number of students passing the exam has a binomial distribution, perform a hypothesis test at the 0.05-significance level to test the null hypothesis "Students taught by lecturer A or lecturer B have the same probability of passing" against the alternative hypothesis "Students taught by lecturer A have a *different* probability of passing than those taught by lecturer B". What is the conclusion?

[Hint: Recall the result from part (a) and again state clearly state any assumptions made]. [7 marks]

Total: 25 marks

END OF PAPER