Queen Mary
University of London

Main Examination period 2017

MTH5122
Statistical Methods

Duration: 2 hours

Apart from this page, you are not permitted to read the contents of this question paper until instructed to do so by an invigilator.

You should attempt ALL questions. Marks available are shown next to the questions.

Statistical functions provided by the calculator may be used provided that you state clearly where you have used them.
The New Cambridge Statistical Tables are provided.
Only non-programmable calculators that have been approved from the college list of non-programmable calculators are permitted in this examination. Please state on your answer book the name and type of machine used. Statistical functions provided by the calculator may be used provided that you state clearly where you have used them.

Complete all rough work in the answer book and cross through any work that is not to be assessed.

Possession of unauthorised material at any time when under examination conditions is an assessment offence and can lead to expulsion from QMUL. Check now to ensure you do not have any notes, mobile phones, smartwatches or unauthorised electronic devices on your person. If you do, raise your hand and give them to an invigilator immediately.

It is also an offence to have any writing of any kind on your person, including on your body. If you are found to have hidden unauthorised material elsewhere, including toilets and cloakrooms, it shall be treated as being found in your possession. Unauthorised material found on your mobile phone or other electronic device will be considered the same as being in possession of paper notes. A mobile phone that causes a disruption in the exam is also an assessment offence.

Exam papers must not be removed from the examination room.

Examiners: A. Gnedin

Question 1. [25 marks] Measurements of the ear length on a sample of 15 plants of wheat yielded the sample mean $\bar{x} = 11$ (in centimetres) and the sample variance $s^2 = 24$. Assuming that the distribution of the wheat ear length is normal $\mathcal{N}(\mu, \sigma^2)$

(a) construct a 90% confidence interval for μ, **[10]**

(b) test the hypothesis that $\mu = 10$ (against alternative $\mu \neq 10$). Use a 10% significance level. **[5]**

(c) construct a 90% confidence interval for σ^2. **[10]**

Question 2. [15 marks] Let $T = X_1 + X_2 + \ldots + X_{12}$, where for $i = 1, 2, \ldots, 12$ the random variables X_i are independent, with $X_i \sim \chi^2_2$ (chi-square distribution with two degrees of freedom). Recall that $\mathbb{E}[X_i] = 2$ and $\mathrm{Var}[X_i] = 4$. This question is concerned with the probability

$$p = \mathbb{P}(|T - 24| > 14).$$

(a) Use Chebyshev's inequality to obtain an upper bound on p. **[5]**

(b) Use the normal approximation to estimate p. **[5]**

(c) Name the distribution of T and evaluate p using Cambridge Statistical Tables. **[5]**

Question 3. [20 marks] Let the random variable Y be the number of defects in a printed circuit board. A sample of 280 circuit boards was taken and the number of defects was recorded:

Number of defects	0	1	2	3	4	5+
Frequency	144	91	32	11	2	0

Test at a significance level of $\alpha = 0.1$ that the distribution of Y is Poisson. Remember to group the data appropriately, if necessary. Show all your working. **[20]**

(continued on the next page)

Question 4. [25 marks] Let $U \sim \mathcal{N}(0,1)$ and $V \sim \mathcal{N}(0,1)$ be independent normal random variables.

(a) Let $S = e^U$.

(i) Determine the probability density function of S. **[6]**

(ii) Find $\mathbb{E}[S]$ and $\mathrm{Var}[S]$ using the formula $M_U(t) = e^{t^2/2}$ for the moment generating function of the normal distribution. **[4]**

(b) Let $X = U + 2V,\ Y = 2U + V$.

(i) Calculate the covariance $\mathrm{Cov}(X, Y)$. **[5]**

(ii) Determine the joint probability density function $f_{X,Y}(x, y)$ of random variables X and Y. **[5]**

(iii) Determine the conditional probability density function $f_{X|Y=2}(x)$. **[5]**

Question 5. [15 marks] Two alternative treatments were tested on 110 patients divided in two groups. Treatment A on a group of 50 patients showed side effects in 12 cases. Treatment B on a group of the other 60 patients showed side effects in 11 cases. Decide if there is sufficient evidence that treatment A causes side effects more frequently than treatment B. **[15]**

End of Paper.