OCR
Oxford Cambridge and RSA

Wednesday 13 June 2018 – Morning

A2 GCE MATHEMATICS

4733/01 Probability & Statistics 2

QUESTION PAPER

Candidates answer on the Printed Answer Book.

OCR supplied materials:
- Printed Answer Book 4733/01
- List of Formulae (MF1)

Other materials required:
- Scientific or graphical calculator

Duration: 1 hour 30 minutes

7306571913

INSTRUCTIONS TO CANDIDATES

These instructions are the same on the Printed Answer Book and the Question Paper.

- The Question Paper will be found inside the Printed Answer Book.
- Write your name, centre number and candidate number in the spaces provided on the Printed Answer Book. Please write clearly and in capital letters.
- **Write your answer to each question in the space provided in the Printed Answer Book.** If additional space is required, you should use the lined page(s) at the end of the Printed Answer Book. The question number(s) must be clearly shown.
- Use black ink. HB pencil may be used for graphs and diagrams only.
- Answer **all** the questions.
- Read each question carefully. Make sure you know what you have to do before starting your answer.
- Do **not** write in the barcodes.
- You are permitted to use a scientific or graphical calculator in this paper.
- Give non-exact numerical answers correct to 3 significant figures unless a different degree of accuracy is specified in the question or is clearly appropriate.

INFORMATION FOR CANDIDATES

This information is the same on the Printed Answer Book and the Question Paper.

- The number of marks is given in brackets **[ ]** at the end of each question or part question on the Question Paper.
- **You are reminded of the need for clear presentation in your answers.**
- The total number of marks for this paper is **72**.
- The Printed Answer Book consists of **12** pages. The Question Paper consists of **4** pages. Any blank pages are indicated.

INSTRUCTION TO EXAMS OFFICER/INVIGILATOR

- Do not send this Question Paper for marking; it should be retained in the centre or recycled. Please contact OCR Copyright should you wish to re-use this document.

Answer **all** the questions.

1 The results of a random sample of 15 observations of the random variable W can be summarised by

$$n = 15, \qquad \Sigma w = 555, \qquad \Sigma w^2 = 20808.$$

Calculate an unbiased estimate of $\mathrm{E}(W)$ and an unbiased estimate of $\mathrm{Var}(W)$. **[4]**

2 **(i)** Explain an advantage of using random numbers in sampling. **[1]**

(ii) A random sample of 5 different letters of the alphabet is to be chosen using the following list of random numbers obtained from a calculator.

0.163 0.542 0.007 0.162 0.005 0.022 0.035 0.119 0.188

Determine which letters form the sample, making your method clear. (You do not need to use all the random numbers.) **[2]**

3 The number of dust particles in 1 cm^3 of air at a certain location is modelled by the distribution Po(2.7).

(i) Find the probability that, in a region of 1 cm^3 of air at this location, there are at least 5 dust particles. **[2]**

(ii) Use a binomial model to calculate the probability that, in 4 adjacent regions of 1 cm^3 of air at this location, there are at least 5 dust particles in exactly 2 of the regions. **[3]**

(iii) Use the formula for Poisson probabilities to write down an exact expression for the probability that, in a region of 4 cm^3 of air at this location, there are exactly 12 dust particles, and evaluate your expression. **[2]**

4 The discrete random variable Y has probability distribution given by

y	0	1	2	3
$\mathrm{P}(Y = y)$	0.4	0.2	0.3	0.1

$\overline{Y}$ denotes the mean of 50 random independent observations of Y.

(i) Find the approximate distribution of $\overline{Y}$, giving the value(s) of any parameter(s). **[5]**

(ii) State the possible values taken by $\overline{Y}$ in the range from 1.4 to 1.5 inclusive. **[1]**

5 A teacher knows from experience that on average she makes 6 mistakes per session when typing reports. She now adopts a new method of organising her report writing and she finds that in one session she has made 10 mistakes.

(i) Assume first that the number of mistakes she makes can be modelled by a Poisson distribution. Test at the 5% significance level whether the average number of mistakes per session has changed since she adopted the new method. **[7]**

(ii) It may be assumed that the teacher makes mistakes randomly and independently of one another.

(a) State another assumption needed if the number of mistakes in a session is to be modelled by a Poisson distribution. **[1]**

(b) Suggest a reason why this assumption might not hold in practice. **[1]**

(c) Explain why this means that further information might be needed to calculate the probability that the teacher makes a given number of mistakes in the second hour of a two-hour session. **[1]**

6 The working lifetime T hours of a computer monitor of a certain type is modelled by a continuous random variable with probability density function

$$\mathrm{f}(t) = \begin{cases} kt^{-4} & t \geqslant L, \\ 0 & \text{otherwise,} \end{cases}$$

where k and L are constants.

(i) The monitor is sold with a guarantee that its working lifetime is at least T_0 hours (if it fails before time T_0 then it is replaced). State what the model suggests about the value of T_0. **[1]**

(ii) Show that $k = 3L^3$. **[3]**

(iii) Find the variance of T in terms of L. **[6]**

(iv) Sketch the graph of the probability density function, and explain whether or not the distribution of T can be well approximated by a normal distribution. **[2]**

7 The maximum effective range (MER), X metres, of a long-established brand of wi-fi hub is a continuous random variable modelled by the distribution $N(\mu, \sigma^2)$. In a large random sample it is found that $P(X > 58) = 0.1587$ and $P(X < 40) = 0.3085$.

(i) Calculate the value of μ and the value of σ. **[6]**

(ii) The manufacturers of a new brand of hub claim that its mean MER is at least 10 metres greater than that of the long-established brand. A random sample of 200 values of the MER for this new brand were measured in a variety of environments. The mean MER for this new brand was found to be $(\mu + 8.8)$ metres, where μ is the value found in part **(i)**. A test is to be carried out at the 10% significance level of whether there is evidence that the manufacturer's claim is valid.

(a) State appropriate hypotheses for the test, explaining the meaning of any symbol you use (other than H_0 and H_1). **[3]**

(b) Assuming that the distribution of the MER of the new brand is normal with variance equal to that of the long-established brand, carry out the test. **[5]**

(iii) Explain whether it was necessary to use the Central Limit Theorem in the test in part **(ii)**. **[1]**

8 *In this question you should justify the use of any approximate distributions.*

The discrete random variable M has the distribution $B(60, p)$. A test of the hypotheses $H_0: p = 0.04$, $H_1: p > 0.04$ is carried out at a significance level as close to 5% as possible.

(i) Use an appropriate approximation to determine the critical region for the test. You should show the values of any relevant probabilities. **[7]**

(ii) Using the critical region found in part **(i)**, the hypothesis test is carried out 100 times. Each time the value of p is 0.04. State the expected number of times that the test will result in a Type I error. **[1]**

(iii) Use an appropriate approximation to find the probability of a Type II error when $p = 0.15$. **[7]**

END OF QUESTION PAPER