OCR
Oxford Cambridge and RSA

Wednesday 7 June 2017 – Morning

AS GCE MEI STATISTICS

G242/01 Statistics 2 (Z2)

QUESTION PAPER

Duration: 1 hour 30 minutes

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.
*   Read each question carefully. Make sure you know what you have to do before starting your answer.
*   Answer **all** the questions.
*   Do **not** write in the bar codes.
*   You are permitted to use a scientific or graphical calculator in this paper.
*   Final answers should be given to a degree of accuracy appropriate to the context.

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 advised that an answer may receive **no marks** unless you show sufficient detail of the working to indicate that a correct method is being used.
*   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.

OCR is an exempt Charity

Turn over

1   **(i)** On average 1.5% of a particular model of calculator are faulty. Faults occur independently and randomly. The number of faulty calculators in a random sample of 100 calculators of this model is denoted by X.

    (*A*) State the exact distribution of X. **[2]**

    (*B*) Explain why the exact distribution of X may be approximated by a Poisson distribution. Write down the mean of this distribution. **[3]**

    (*C*) Use this Poisson distribution to find the probability that a sample contains at least 1 faulty calculator. **[2]**

  (ii) The weight of bananas supplied to a supermarket chain is Normally distributed with mean 105 g and standard deviation 10 g.

    (*A*) Find the probability that the weight of a randomly selected banana is less than 100 g. **[3]**

    (*B*) Find the probability that the total weight of 5 randomly selected bananas is less than 550 g. **[5]**

2   A company that makes activity monitors is developing a wristband to measure walking distance. In one test of accuracy, the tester wears a wristband and walks a course of exactly 1500 m in length. The measurements, in metres, given by a random sample of 15 wristbands are as follows.

    1470 1482 1493 1465 1502 1471 1506 1512 1484 1509 1489 1501 1464 1521 1492

  (i) Use a Wilcoxon test to examine, at the 5% significance level, whether these data provide evidence that on average the wristbands give an accurate measurement of the 1500 m course. State any necessary assumption. **[12]**

  (ii) What further assumption would be necessary for a *t* test to be appropriate in this case? **[1]**

3 Hydroponics is a soil-free method used to grow plants, in which soil is replaced by material known as a 'growing medium'. A horticulturalist is investigating the possible association between growing medium and yield for a particular type of chilli plant. He decides to carry out a test. The growing medium and the yield of a random sample of 200 plants are recorded. Yield is classified as Low or High. The results are as follows.

		Yield	
		Low	High
Type of growing medium	Coco Fibre	19	39
	Perlite	24	23
	Rockwool	23	22
	Vermiculite	31	19

The following tables show some of the expected frequencies and contributions to the test statistic.

Expected frequencies		Yield	
		Low	High
Type of growing medium	Coco Fibre	28.13	29.87
	Perlite	22.795	
	Rockwool	21.825	
	Vermiculite	24.25	

Contributions to the test statistic		Yield	
		Low	High
Type of growing medium	Coco fibre	2.9633	2.7907
	Perlite	0.0637	
	Rockwool	0.0633	
	Vermiculite	1.8789	

(i) Calculate the remaining expected frequencies and contributions to the test statistic. Carry out the test at the 5% level of significance. **[11]**

(ii) For each type of growing medium, comment briefly on how the yield compares to what would be expected if there were no association. **[4]**

4 Pesticides are used widely in the agricultural industry. It is common for small amounts of pesticide to be present in items that are sold as food. A food distributor tests all of its products for pesticide content to ensure that health regulations are met.

One test involves sunflower seeds which are packaged in large bags. When testing sunflower seeds the food distributor selects a sample of bags and measures the amount of pesticide present in each bag. The results, in suitable units, for a random sample of 12 bags from a particular batch are as follows.

0.51	0.39	0.54	0.42	0.22	0.08
0.27	0.48	0.56	0.67	0.19	0.48

For a batch to meet health regulations the mean pesticide content of a bag should be below 0.5.

(i) Carry out a suitable test to examine whether the mean pesticide content of a bag for this batch meets health regulations. Use a 5% significance level. State any necessary assumption. **[13]**

A hypothesis test involving a different batch of sunflower seeds gave a p-value of 0.073. This test was also carried out at the 5% significance level.

(ii) State the conclusion of this test. Justify your answer. **[2]**

5 A golf equipment manufacturer uses rechargeable batteries in its golf trolleys. It is trialling a new type of battery which is claimed to have a greater capacity than the old battery. The old battery is known to have a capacity of 20 units on average. The capacities of a random sample of 50 of the new batteries are measured. The sample mean is 21.6 units and the sample standard deviation is 3.1 units.

(i) Write down estimates of the population mean and population variance for the new battery. **[2]**

(ii) The sample mean exceeds 20. Explain why it is not appropriate, based on this evidence alone, to conclude that the new battery has a greater capacity than the old battery. **[2]**

(iii) Write down an estimate of the standard error of the sample mean. **[1]**

(iv) Calculate a 95% confidence interval, based on the Normal distribution, for the population mean capacity of the new battery. **[5]**

(v) Does the interval found in part **(iv)** provide evidence to suggest that the new battery has greater mean capacity than the old? Justify your answer. **[2]**

(vi) Explain why it was appropriate to use the Normal distribution to construct the confidence interval in part **(iv)** rather than the t distribution. **[2]**

END OF QUESTION PAPER