OCR
Oxford Cambridge and RSA

Thursday 14 October 2021 – Afternoon

A Level Further Mathematics A

Y542/01 Statistics

Time allowed: 1 hour 30 minutes

7821455783

You must have:
- the Printed Answer Booklet
- the Formulae Booklet for A Level Further Mathematics A
- a scientific or graphical calculator

INSTRUCTIONS
- Use black ink. You can use an HB pencil, but only for graphs and diagrams.
- Write your answer to each question in the space provided in the **Printed Answer Booklet**. If you need extra space use the lined pages at the end of the Printed Answer Booklet. The question numbers must be clearly shown.
- Answer **all** the questions.
- Where appropriate, your answer should be supported with working. Marks might be given for using a correct method, even if your answer is wrong.
- Give non-exact numerical answers correct to **3** significant figures unless a different degree of accuracy is specified in the question.
- The acceleration due to gravity is denoted by $g\,\text{m}\,\text{s}^{-2}$. When a numerical value is needed use $g = 9.8$ unless a different value is specified in the question.
- Do **not** send this Question Paper for marking. Keep it in the centre of recycle it.

INFORMATION
- The total mark for this paper is **75**.
- The marks for each question are shown in brackets **[ ]**.
- This document has **8** pages.

ADVICE
- Read each question carefully before you start your answer.

© OCR 2021 [J/508/5510]
DC (CE) 307446/4
OCR is an exempt Charity

Answer **all** the questions.

1 At a seaside resort the number X of ice-creams sold and the temperature Y °F were recorded on 20 randomly chosen summer days. The data can be summarised as follows.

$\Sigma x = 1506 \quad \Sigma x^2 = 127\,542 \quad \Sigma y = 1431 \quad \Sigma y^2 = 104\,451 \quad \Sigma xy = 111\,297$

(a) Calculate the equation of the least squares regression line of y on x, giving your answer in the form $y = a + bx$. **[3]**

(b) Explain the significance for the regression line of the quantity $\sum[y_i - (ax_i + b)]^2$. **[1]**

(c) It is decided to measure the temperature in degrees Centigrade instead of degrees Fahrenheit. If the same temperature is measured both as f° Fahrenheit and c° Centigrade, the relationship between f and c is $c = \frac{5}{9}(f - 32)$.

Find the equation of the new regression line. **[2]**

2 A discrete random variable D has the following probability distribution, where a is a constant.

d	0	2	4	6
$\mathrm{P}(D = d)$	a	0.1	0.3	0.2

Determine the value of $\mathrm{Var}(3D + 4)$. **[7]**

3 In a large collection of coloured marbles of identical size, the proportion of green marbles is p. One marble is chosen randomly, its colour is noted, and it is then replaced. This process is repeated until a green marble is chosen.

The first green marble chosen is the Xth marble chosen.

(a) You are given that $p = 0.3$.

(i) Find $\mathrm{P}(5 \leqslant X \leqslant 10)$. **[2]**

(ii) Determine the smallest value of n for which $\mathrm{P}(X = n) < 0.1$. **[2]**

(b) You are given instead that $\mathrm{Var}(X) = 42$.

Determine the value of $\mathrm{E}(X)$. **[5]**

© OCR 2021

4 A random sample of 160 observations of a random variable X is selected. The sample can be summarised as follows.

$n = 160 \qquad \Sigma x = 2688 \qquad \Sigma x^2 = 48\,398$

(a) Calculate unbiased estimates of the following.

(i) $\mathrm{E}(X)$ **[1]**

(ii) $\mathrm{Var}(X)$ **[3]**

(b) Find a 99% confidence interval for $\mathrm{E}(X)$, giving the end-points of the interval correct to **4** significant figures. **[3]**

(c) Explain whether it was necessary to use the Central Limit Theorem in answering

(i) part **(a)**, **[1]**

(ii) part **(b)**. **[1]**

5 The numbers of each of 9 items sold in two different supermarkets in a week are given in the following table.

Item	1	2	3	4	5	6	7	8	9
Supermarket *A*	17	28	41	43	62	69	75	93	115
Supermarket *B*	24	7	18	12	47	29	58	42	37

A researcher wants to test whether there is association between the numbers of these items sold in the two supermarkets.

However, it is known that the collection of data in Supermarket *B* was done inaccurately and each of the numbers in the corresponding row of the table could have been in error by as much as 2 items greater or 2 items fewer.

(a) Explain why Spearman's rank correlation coefficient might be preferred to the use of Pearson's product-moment correlation coefficient in this context. **[2]**

(b) Carry out the test at the 5% significance level using Spearman's rank correlation coefficient. **[8]**

© OCR 2021

6 A practice examination paper is taken by 500 candidates, and the organiser wishes to know what continuous distribution could be used to model the actual time, X minutes, taken by candidates to complete the paper.

The organiser starts by carrying out a goodness-of-fit test for the distribution N(100, 15^2) at the 5% significance level. The grouped data and the results of some of the calculations are shown in the following table.

Time	$0 \leqslant X < 80$	$80 \leqslant X < 90$	$90 \leqslant X < 100$	$100 \leqslant X < 110$	$X \geqslant 110$
Observed frequency O	36	95	137	129	103
Expected frequency E	45.606	80.641	123.754	123.754	126.246
$\frac{(O-E)^2}{E}$	2.023	2.557	1.418	0.222	4.280

(a) State suitable hypotheses for the test. **[1]**

(b) Show how the figures 123.754 and 0.222 in the column for $100 \leqslant X < 110$ were obtained. **[3]**

(c) Carry out the test. **[4]**

The organiser now wants to suggest an improved model for the data.

(d) (i) Suggest an aspect of the data that the organiser should take into account in considering an improved model. **[1]**

(ii) The graph of the probability density function for the distribution N(100, 15^2) is shown in the diagram in the Printed Answer Booklet.

On the same diagram sketch the probability density function of an improved model that takes into account the aspect of the data in part **(d)(i)**. **[2]**

© OCR 2021

7 In a school opinion poll a random sample of 8 pupils were asked to rate school lunches on a scale of 0 to 20. The results were as follows.

0 1 2 3 4 10 11 13

After a new menu was introduced, the test was repeated with a different random sample of 8 pupils. The results were as follows.

7 8 9 14 15 17 19 20

(a) Carry out an appropriate Wilcoxon test at the 5% significance level to test whether pupils' opinions of school lunches have changed. **[8]**

A statistics student tells the organisers of the opinion poll that it would have been better to have asked the same 8 pupils both times.

(b) Explain why the statistics student's suggestion would produce a better test. **[1]**

(c) State which test should be used if the student's suggestion is followed. **[1]**

(d) You are given that there are 12 870 ways in which 8 different integers can be chosen from the integers 1 to 16 inclusive.

Estimate the number of ways of selecting 8 different digits between 1 and 16 inclusive that have a sum less than or equal to the critical value used in the test in part **(a)**. **[2]**

8 The continuous random variable Y has a uniform distribution on [0, 2].

(a) It is given that $\mathrm{E}[a\cos(aY)] = 0.3$, where a is a constant between 0 and 1, and aY is measured in radians.

Determine the value of the constant a. **[5]**

(b) Determine the 60^{th} percentile of Y^2. **[6]**

END OF QUESTION PAPER

© OCR 2021

BLANK PAGE

BLANK PAGE

OCR
Oxford Cambridge and RSA

Copyright Information

OCR is committed to seeking permission to reproduce all third-party content that it uses in its assessment materials. OCR has attempted to identify and contact all copyright holders whose work is used in this paper. To avoid the issue of disclosure of answer-related information to candidates, all copyright acknowledgements are reproduced in the OCR Copyright Acknowledgements Booklet. This is produced for each series of examinations and is freely available to download from our public website (www.ocr.org.uk) after the live examination series.

If OCR has unwittingly failed to correctly acknowledge or clear any third-party content in this assessment material, OCR will be happy to correct its mistake at the earliest possible opportunity.

For queries or further information please contact The OCR Copyright Team, The Triangle Building, Shaftesbury Road, Cambridge CB2 8EA.

OCR is part of the Cambridge Assessment Group; Cambridge Assessment is the brand name of University of Cambridge Local Examinations Syndicate (UCLES), which is itself a department of the University of Cambridge.

© OCR 2021