

OCR

Oxford Cambridge and RSA

Tuesday 9 June 2015 – Morning

AS GCE MEI STATISTICS

G243/01 Statistics 3 (Z3)

QUESTION PAPER

Candidates answer on the Printed Answer Book.

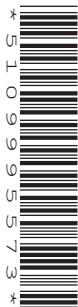
OCR supplied materials:

- Printed Answer Book G243/01
- MEI Examination Formulae and Tables (MF2)

Other materials required:

- Scientific or graphical calculator

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.** Additional paper may be used if necessary but you must clearly show your candidate number, centre number and question number(s).
- 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.

Section A (46 marks)

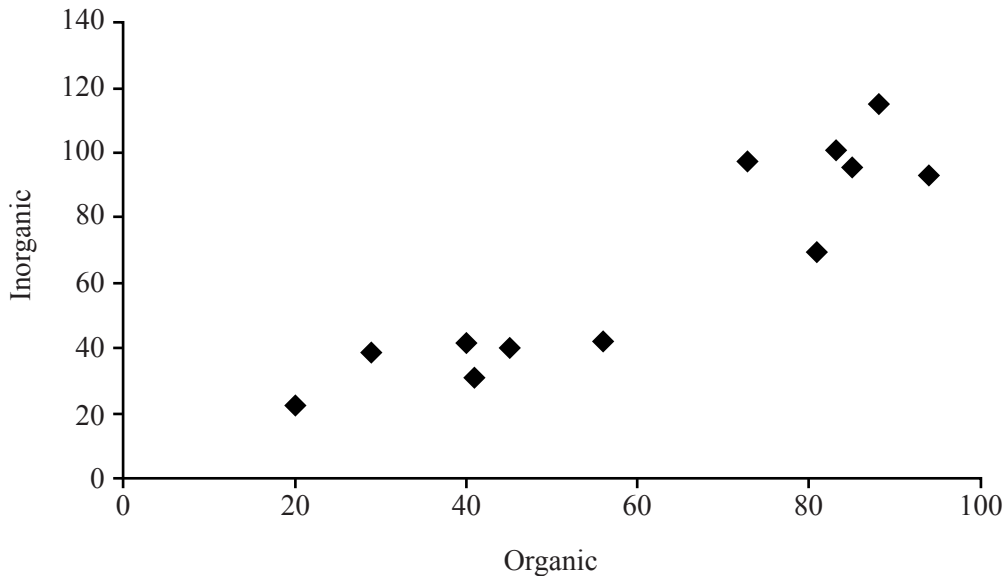
- 1 An airline consultant is investigating the punctuality of two airlines, A and B, which share the same route. She suspects that, on average, flights with airline A will arrive earlier than flights with airline B. She selects random samples of 50 flights for each airline. The difference in minutes between the scheduled arrival times and actual arrival times are denoted by x for airline A and y for airline B. A negative value indicates a late arrival. You are given that $\Sigma x = 275.8$, $\Sigma x^2 = 2910.5$ and the sample mean and sample standard deviation of y are 3.74 and 6.45 respectively.
- (i) Calculate the sample mean and sample standard deviation of x . [3]
- (ii) Carry out a test at the 10% significance level to investigate the consultant's suspicion. [11]
- (iii) Suppose that a 1% significance level is used instead of a 10% significance level in a hypothesis test such as this, and that the null hypothesis is true. Is the null hypothesis more likely, equally likely or less likely to be rejected? Briefly explain your answer. [2]
- 2 Two friends, Jenna and Ronin, are comparing the speeds of copying files from a USB drive to their computers. They wish to check if, on average, the copying speeds are different. 11 files are randomly selected and Jenna and Ronin each copy these files onto their computers. The copying speeds, in suitable units, are as follows.

File	A	B	C	D	E	F	G	H	I	J	K
Ronin	9.98	9.79	9.68	9.74	10.33	9.94	10.56	9.91	10.23	11.23	10.64
Jenna	10.20	11.36	8.68	10.50	10.97	7.86	8.84	9.26	10.46	11.41	10.73

- (i) Briefly explain why it is better to use the same files to copy onto both of the computers. [2]
- (ii) State the distributional assumption necessary for the use of a paired sample t test. [2]
- (iii) Use a paired sample t test to examine, at the 5% significance level, whether it appears that there is any difference between the average copying speeds for the two computers. [11]

- 3 An agricultural scientist thinks that there may be a relationship between levels of organic and inorganic phosphorus in the soil in a river basin. She takes samples of soil from 12 randomly chosen locations. The levels of both types of phosphorus in each of the samples, measured in suitable units, are given below. The data are also shown on a scatter diagram.

Location	1	2	3	4	5	6	7	8	9	10	11	12
Organic	85	83	73	20	29	88	94	45	41	81	56	40
Inorganic	95	100	97	22	39	115	93	40	31	69	42	41



- (i) Explain why, in view of the shape of the scatter diagram, the assumptions required to carry out a test based on the product moment correlation coefficient may not be satisfied. [3]
- (ii) Calculate the value of Spearman's rank correlation coefficient for these data. [5]
- (iii) Using your answer to part (ii) carry out a test, at the 1% level of significance, to determine whether it is reasonable to assume that levels of organic and inorganic phosphorus are positively associated. [7]

Question 4 begins on page 4.

Section B (26 marks)

- 4 A researcher is investigating the performances of various groups of people in an online typing test.
- (i) He first considers whether to use a census or a survey. Give one advantage of each approach. [2]
 - (ii) The researcher decides to use a survey to select 12 women who have taken the typing test. Explain why it would not be sensible to select the first 12 women who have taken the test on a particular day. [2]
 - (iii) He then considers selecting people by systematic sampling. Given that 960 women have taken the test, explain how he can select a systematic sample of 12 of them. [3]

The researcher wishes to investigate whether there is any difference between the average test results of women and men. He takes random samples of 12 women and 12 men. The results, measured in words per minute, rounded to the nearest integer, are as follows.

Women	18	21	26	27	29	33	36	37	41	46	65	78
Men	13	14	22	23	24	28	30	31	35	42	43	61

- (iv) State the assumptions which are required for a t test to examine whether there is any difference in the mean typing speeds of women and men. [2]
- (v) Draw a histogram for the data for women, using class intervals 9.5–19.5, 19.5–29.5, 29.5–39.5, 39.5–49.5, 49.5–79.5. Name the type of skewness shown by the histogram and explain why it would be unwise to assume an underlying Normal distribution. [5]
- (vi) Carry out an appropriate test, at the 5% significance level, to investigate whether it appears that the median performances of women and men are the same. [10]
- (vii) Explain why the researcher took a random sample of 12 women and 12 men, rather than randomly selecting just 1 woman and 1 man. [2]

END OF QUESTION PAPER

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 Copyright Team, First Floor, 9 Hills Road, Cambridge CB2 1GE.

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.