Write your name here

Surname	Other names

Pearson Edexcel GCE

Centre Number | | | | | |

Candidate Number | | | | |

Statistics S3

Advanced/Advanced Subsidiary

Wednesday 23 May 2018 – Morning
Time: 1 hour 30 minutes

Paper Reference
6691/01

You must have:
Mathematical Formulae and Statistical Tables (Pink)

Total Marks

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulae stored in them.

Instructions

- Use **black** ink or ball-point pen.
- If pencil is used for diagrams/sketches/graphs it must be dark (HB or B). Coloured pencils and highlighter pens must not be used.
- **Fill in the boxes** at the top of this page with your name, centre number and candidate number.
- Answer **all** questions and ensure that your answers to parts of questions are clearly labelled.
- Answer the questions in the spaces provided – *there may be more space than you need.*
- You should show sufficient working to make your methods clear. Answers without working may not gain full credit.
- Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information

- The total mark for this paper is 75.
- The marks for **each** question are shown in brackets – *use this as a guide as to how much time to spend on each question.*

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- Check your answers if you have time at the end.

Turn over ▶

P51573A

©2018 Pearson Education Ltd.

1/1/1/

P 5 1 5 7 3 A 0 1 2 4

Pearson

Leave blank

1. Phil measures the concentration of a radioactive element, c, and the amount of dissolved solids, a, of 8 random samples of groundwater. His results are shown in the table below.

Sample	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
c	625	700	650	645	720	600	825	665
a	1.28	1.30	1.00	1.20	1.55	1.15	1.40	1.45

Given that

$$S_{cc} = 34\,787.5 \qquad S_{aa} = 0.217\,287\,5 \qquad S_{ca} = 47.7625$$

(a) calculate, to 3 decimal places, the product moment correlation coefficient between the concentration of the radioactive element and the amount of dissolved solids for these groundwater samples.

(1)

(b) Use your value of the product moment correlation coefficient to test whether or not there is evidence of a positive correlation between the concentration of this radioactive element and the amount of dissolved solids in groundwater. Use a 5% significance level. State your hypotheses clearly.

(3)

(c) Calculate, to 3 decimal places, Spearman's rank correlation coefficient between the concentration of the radioactive element and the amount of dissolved solids.

(5)

(d) Use your value of Spearman's rank correlation coefficient to test for evidence of a positive correlation between the concentration of the radioactive element and the amount of dissolved solids. Use a 5% significance level. State your hypotheses clearly.

(3)

(e) Using your conclusions in part (b) and part (d), comment on the possible relationship between these variables.

(1)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 1 continued

Leave blank

Question 1 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 1 continued

(Total 13 marks)

Q1

Leave blank

2. Merchandise is sold at concerts. The manager of a concert claims that the mean value of merchandise sold to premium ticket holders is more than £6 greater than the mean value of merchandise sold to standard ticket holders.

(a) Given that all the tickets for the next concert have been sold, describe how a stratified sample should be taken at the concert.

(3)

The mean value of merchandise sold to a random sample of 60 standard ticket holders at the concert is £15 with a standard deviation of £10.

The mean value of merchandise sold to a random sample of 55 premium ticket holders at the concert is £23 with a standard deviation of £8.

(b) Test the manager's claim at the 5% level of significance. State your hypotheses clearly.

(8)

(c) For the test in part (b), state whether or not it is necessary to assume that values of merchandise sold have normal distributions. Give a reason for your answer.

(2)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 2 continued

Leave blank

Question 2 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 2 continued

(Total 13 marks)

Q2

Leave blank

3. A random sample of repair times, in hours, was taken for an electronic component. The 4 observed times are shown below.

1.3 1.7 1.4 1.8

(a) Calculate unbiased estimates of the mean and the variance of the population of repair times for this electronic component.

(4)

The population standard deviation of the repair times for this electronic component is known to be 0.5 hours.

An estimate of the population mean is required to be within 0.1 hours of its true value with a probability of at least 0.99

(b) Find the minimum sample size required.

(6)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 3 continued

(Total 10 marks)

Q3

Leave blank

4. The waiting times, in minutes, of patients at a doctor's surgery follows a normal distribution with unknown mean μ and known standard deviation σ

A random sample of 120 patients was taken.

(a) Find, in the form $k\sigma$, the width of a 99% confidence interval for μ based on this sample. Give the value of k to 2 decimal places.

(3)

A further random sample of 100 patients from the surgery gave a 90% confidence interval for μ of (5.14, 6.25)

(b) Use this confidence interval to determine whether or not it provides evidence that $\mu = 6$

State the hypotheses being tested here and write down the significance level being used. You do not need to carry out any further calculations.

(3)

(c) Find the value of σ

(3)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 4 continued

Leave blank

Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 4 continued

Q4

(Total 9 marks)

Leave blank

5. The weights, in kg, of cars may be assumed to follow the normal distribution $N(1000, 250^2)$. The weights, in kg, of lorries may be assumed to follow the normal distribution $N(2800, 650^2)$.

A lorry and a car are chosen at random.

(a) Find the probability that the lorry weighs more than 3 times the weight of the car.

(6)

A ferry carries vehicles across a river. The ferry is designed to carry a maximum weight of 20 000 kg.

(b) One morning, 8 cars and 3 lorries drive on to the ferry. Find the probability that their total weight will exceed the recommended maximum weight of 20 000 kg.

(5)

(c) State a necessary assumption needed for the calculation in part (b).

(1)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 5 continued

Leave blank

Question 5 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 5 continued

(Total 12 marks)

Q5

Leave blank

6. David carries out an experiment with 4 identical dice, each with faces numbered 1 to 6. He rolls the 4 dice and counts the number of dice showing an even number on the uppermost face. He repeats this 150 times. The results are summarised in the table below.

No. of dice showing an even number	0	1	2	3	4
Frequency	12	45	36	39	18

David defines the random variable C as the number of dice showing an even number on the uppermost face when the four dice are thrown.

David claims that $C \sim B(4, 0.5)$

(a) Stating your hypotheses clearly and using a 1% level of significance, test David's claim. Show your working clearly.

(9)

John claims that $C \sim B(4, p)$

(b) Calculate an estimate of the value of p from the summary of the results of David's experiment. Show your working clearly.

(2)

John decides to test his claim. He calculates expected frequencies using the results of David's experiment and obtains the following table.

No. of dice showing an even number	0	1	2	3	4
Expected frequency	8.65	36.00	d	39.00	e

(c) Calculate, to 2 decimal places, the value of d and the value of e

(3)

(d) State suitable hypotheses to test John's claim.

(1)

John obtained a test statistic of 16.9 and carries out a test at the 1% level of significance.

(e) State what conclusion John should make about his claim.

(3)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 6 continued

Leave blank

Question 6 continued

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 6 continued

Leave blank

Question 6 continued

Q6

(Total 18 marks)

TOTAL FOR PAPER: 75 MARKS

END

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA