Candidate surname	Other names

**Pearson
Edexcel GCE**

Centre Number

Candidate Number

Wednesday 22 May 2019

Morning (Time: 1 hour 30 minutes)	Paper Reference **6691/01**

Statistics S3
Advanced/Advanced Subsidiary

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

Candidates may use any calculator allowed by Pearson regulations. 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 ▶

1. Universities *A*, *B*, *C*, *D*, *E*, *F*, *G*, *H* and *I* are ranked each year. A newspaper and a guide rank each university according to how it performs. The table below shows the rankings given by each publication starting from the university with the strongest performance.

Rank	1	2	3	4	5	6	7	8	9
Newspaper	*C*	*D*	*B*	*A*	*I*	*E*	*H*	*G*	*F*
Guide	*D*	*C*	*A*	*B*	*H*	*F*	*I*	*E*	*G*

(a) Calculate Spearman's rank correlation coefficient for these data.

(5)

University *F* claims that the publications do not generally agree with each other about the performance of the universities.

(b) Stating your hypotheses clearly, test at the 5% level of significance, whether or not these data provide support for university *F*'s claim.

(4)

Question 1 continued

(Total 9 marks)

2. Dave sells ice cream. Customers can choose one of four types of ice cream; a cone, a sandwich, a sundae or a float. He believes that the choice of type of ice cream is independent of gender.

Dave takes a random sample of 100 male and 120 female customers. He records the proportion of male customers that choose each type of ice cream. He also records the proportion of female customers that choose each type of ice cream. The results are summarised in the table below.

	Male	**Female**
Cone	$\frac{1}{2}$	$\frac{2}{3}$
Sandwich	$\frac{1}{4}$	$\frac{1}{6}$
Sundae	$\frac{1}{10}$	$\frac{1}{12}$
Float	$\frac{3}{20}$	$\frac{1}{12}$

Stating your hypotheses clearly, test Dave's belief using a 5% level of significance. Show your working clearly. State the critical value and degrees of freedom for this test.

(12)

Question 2 continued

P 5 5 8 6 0 A 0 5 2 4

Question 2 continued

Question 2 continued

(Total 12 marks)

3. A machine fills packets with seeds. The weight of seeds in a packet, W grams, is normally distributed with mean μ.

There is a requirement that the weight of seeds in a randomly selected packet is such that

$$P(W - \mu > 20) = 0.001$$

(a) Show that W has a standard deviation of $\sigma = 6.47$ to 3 significant figures for this requirement to be met.

(3)

The machine is set up so that this requirement is met.

A random sample of 12 packets filled by the machine is selected. The total weight of seeds in this sample is 5940 g.

(b) Stating your hypotheses clearly, test, at the 5% level of significance, whether or not the machine is filling packets with a mean weight of less than 500 g.

(7)

Question 3 continued

Question 3 continued

(Total 10 marks)

Q3

P55860A01124

4. The heights, in cm, of the population of *Seapron* are normally distributed with a variance of $38\,\text{cm}^2$. A random sample of 50 residents of *Seapron* has a mean height of $172\,\text{cm}$.

(a) Find the standard error of the mean.

(1)

(b) Use your answer to part (a) to find a 99% confidence interval for the mean height of the population of *Seapron*.

(4)

A second random sample is to be chosen from the population of *Seapron*. The width of the 95% confidence interval for the mean height is required to be at most $3\,\text{cm}$.

(c) Find the minimum sample size for this requirement to be met.

(4)

Question 4 continued

5. A group of veterinary care centres has 2520 male dogs and 1680 female dogs on its records that are taking part in a weight loss programme. A centre manager decides to take a stratified sample of size 60 from these records.

(a) Describe how to take this stratified sample.

(3)

The weights, w kg, of each of the 60 dogs in the sample are summarised below.

$$\sum w = 2700 \qquad \sum w^2 = 125\,500$$

(b) Find unbiased estimates of the mean and variance of the weights of all the dogs taking part in the weight loss programme.

(4)

Another random sample is taken from these records. This sample is made up of 20 male dogs and 40 female dogs. The mean weight loss of the male dogs in this sample is 0.8 kg. The mean weight loss of the female dogs in this sample is 2.1 kg. It is known that the population standard deviations for weight loss of male dogs and for weight loss of female dogs are 0.9 kg and 1.2 kg respectively. You may assume that the weight loss of male dogs and the weight loss of female dogs are each normally distributed.

The manager claims that the mean weight loss of the female dogs is greater than 1 kg more than the mean weight loss of the male dogs.

(c) Stating your hypotheses clearly test, at the 10% level of significance, the manager's claim.

(7)

Question 5 continued

P 5 5 8 6 0 A 0 1 5 2 4

Question 5 continued

P 5 5 8 6 0 A 0 1 6 2 4
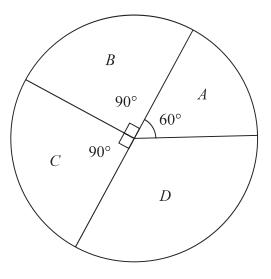
Question 5 continued

6.

Figure 1

A circular archery target is split into 4 sectors as shown in Figure 1. Paula records where 100 arrows hit the target. The results are shown in the table below.

Sector	A	B	C	D
Frequency	13	22	25	40

Paula claims that the probability that an arrow lands in a particular sector is directly proportional to the area of the sector.

Stating your hypotheses clearly and using a 1% level of significance, test whether or not the data supports Paula's claim. State the critical value and the degrees of freedom for this test.

(9)

Question 6 continued

Question 6 continued

Question 6 continued

Q6

(Total 9 marks)

P 5 5 8 6 0 A 0 2 1 2 4

7. Two independent random samples X_1, X_2, X_3, X_4 and Y_1, Y_2, Y_3, Y_4 are each taken from a normal population with mean μ and variance σ^2

(a) Find the distribution of the random variable $M = 4X_1 - 3X_2 - \overline{Y}$

(4)

(b) Hence show that $P(4X_1 < 3X_2 + \overline{Y} + \sigma) = 0.579$ to 3 significant figures.

(3)

A random sample W_1, W_2, W_3, W_4 is also taken from a normal population with mean μ and variance σ^2

(c) Find the distribution of the random variable $T = 4W_1 - 3W_2 - \overline{W}$

(5)

Question 7 continued

P 5 5 8 6 0 A 0 2 3 2 4

Question 7 continued

Q7

(Total 12 marks)

TOTAL FOR PAPER: 75 MARKS

END

P 5 5 8 6 0 A 0 2 4 2 4