Write your name here

Surname	Other names

Pearson Edexcel GCE

Centre Number	Candidate Number

Statistics S3

Advanced/Advanced Subsidiary

Wednesday 24 May 2017 – 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 ▶

P48792A

©2017 Pearson Education Ltd.

1/1/1/1/

P 4 8 7 9 2 A 0 1 2 4

Pearson

Leave blank

1. A company director decides to survey staff about changes to the company calendar. The company has staff in 4 different job roles

72 managers, 108 drivers, 180 administrators and 360 warehouse staff.

The director decides to take a stratified sample.

(a) Write down one advantage of using a stratified sample rather than a simple random sample for this survey.

(1)

(b) Find the number of staff in each job role that will be included in a stratified sample of 40 staff.

(3)

(c) Describe how to choose managers for the stratified sample.

(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 1 continued

Q1

(Total 6 marks)

Leave blank

2.

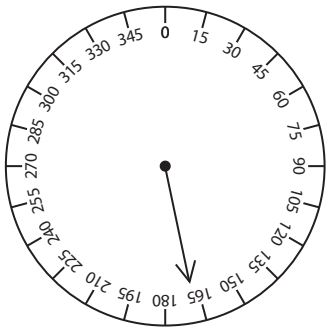


Figure 1

The pointer shown in Figure 1 is spun so that it comes to rest between 0 and 360 degrees.

Linda claims that it is equally likely to come to rest at any point between 0 and 360 degrees. She spins the pointer 100 times and her results are summarised in the table below. She calculates expected frequencies for some of the possible outcomes and these are also given in the table below.

Angle (degrees)	0–45	45–90	90–180	180–315	315–360
Frequency	18	16	18	29	19
Expected frequency	12.5	*a*	*b*	*c*	12.5

(a) Find the values of the missing expected frequencies a, b and c.

(2)

(b) Stating your hypotheses clearly and using a 5% level of significance, test whether or not Linda's claim is supported by these data.

(8)

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 10 marks)

Q2

Leave blank

3. A junior judge is being trained by a senior judge to learn how to assess ice skaters. After the training, the judges each assess 6 ice skaters *A*, *B*, *C*, *D*, *E* and *F*. They each list them in order of preference with the best ice skater first. The results are shown in the table below.

Rank	1	2	3	4	5	6
Senior Judge	*A*	*B*	*D*	*C*	*F*	*E*
Junior Judge	*B*	*D*	*A*	*F*	*C*	*E*

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

(5)

(b) Test, at the 5% level of significance, whether or not there is evidence of a positive correlation between the rankings of the junior judge and the senior judge. State your hypotheses clearly.

(4)

(c) Comment on the effectiveness of the training delivered by the senior judge.

(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 3 continued

Leave blank

Question 3 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 3 continued

(Total 10 marks)

Q3

Leave blank

4. A psychologist carries out a survey of the perceived body weight of 150 randomly chosen people. He asks them if they think they are underweight, about right or overweight. His results are summarised in the table below.

	Underweight	About right	Overweight
Male	20	22	30
Female	16	28	34

The psychologist calculates two of the expected frequencies, to 2 decimal places, for a test of independence between perceived body weight and gender. These results are shown in the table below.

	Underweight	About right	Overweight
Male	17.28		
Female	18.72		

(a) Complete the table of expected frequencies shown above.

(2)

(b) Test, at the 10% level of significance, whether or not perceived body weight is independent of gender. State your hypotheses clearly.

(7)

The psychologist now combines the male and female data to test whether or not body weight types are chosen equally.

(c) Find the smallest significance level, from the tables in the formula booklet, for which there is evidence of a preference.

(5)

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

(Total 14 marks)

Q4

Leave blank

5. Paul takes the company bus to work. According to the bus timetable he should arrive at work at 0831. Paul believes the bus is not reliable and often arrives late. Paul decides to test the arrival time of the bus and carries out a survey. He records the values of the random variable

X = number of minutes after 0831 when the bus arrives.

His results are summarised below.

$$n = 15 \qquad \sum x = 60 \qquad \sum x^2 = 1946$$

(a) Calculate unbiased estimates of the mean, μ, and the variance of X.

(3)

Using the mean of Paul's sample and given $X \sim N(\mu, 10^2)$

(b) (i) calculate a 95% confidence interval for the mean arrival time at work for this company bus.

(ii) State an assumption you made about the values in the sample obtained by Paul.

(5)

(c) Comment on Paul's belief. Justify 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 5 continued

Q5

(Total 10 marks)

Leave blank

6. An engineer has developed a new battery. She claims that the new battery will last more than 8 hours longer, on average, than the old battery. To test the claim, the engineer randomly selects a sample of 50 new batteries and 40 old batteries. She records how long each battery lasts, x hours for the new batteries and y hours for the old batteries. The results are summarised in the table below.

	n	Sample mean	s^2
New battery	50	$\bar{x} = 83$	7
Old battery	40	$\bar{y} = 74$	6

(a) Test, at the 5% level of significance, whether or not there is evidence to support the engineer's claim. State your hypotheses and show your working clearly.

(7)

(b) Explain the relevance of the Central Limit Theorem to the test in part (a).

(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 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

Q6

(Total 9 marks)

Leave blank

7. Sugar is packed into medium bags and large bags. The weights of the medium bags of sugar are normally distributed with mean 520 grams and standard deviation 10 grams. The weights of the large bags of sugar are normally distributed with mean 1510 grams and standard deviation 20 grams.

(a) Find the probability that a randomly chosen large bag of sugar weighs at least 15 grams more than the combined weight of 3 randomly chosen medium bags of sugar.

(6)

(b) Find the probability that a randomly chosen large bag of sugar weighs less than 3 times the weight of a randomly chosen medium bag of sugar.

(5)

A random sample of 5 medium bags of sugar is taken.

(c) Find the value of d so that the probability that all 5 bags of sugar each weigh more than 520 grams is equal to the probability that the mean weight of the 5 bags of sugar is more than d grams.

(5)

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 7 continued

Leave blank

Question 7 continued

Q7

(Total 16 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