Write your name here

Surname	Other names

Pearson Edexcel GCE

Centre Number	Candidate Number

Statistics S3

Advanced/Advanced Subsidiary

Wednesday 25 May 2016 – 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 ▶

P46676A

1/1/1/1/

P 4 6 6 7 6 A 0 1 2 8

PEARSON

Leave blank

1. (a) State two reasons why stratified sampling might be a more suitable sampling method than simple random sampling.

(2)

(b) State two reasons why stratified sampling might be a more suitable sampling method than quota sampling.

(2)

Leave blank

Question 1 continued

Q1

(Total 4 marks)

Leave blank

2. A new drug to vaccinate against influenza was given to 110 randomly chosen volunteers. The volunteers were given the drug in one of 3 different concentrations, *A*, *B* and *C*, and then were monitored to see if they caught influenza. The results are shown in the table below.

	A	*B*	*C*
Influenza	12	29	9
No influenza	15	23	22

Test, at the 10% level of significance, whether or not there is an association between catching influenza and the concentration of the new drug. State your hypotheses and show your working clearly. You should state your expected frequencies to 2 decimal places.

(10)

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

Leave blank

Question 2 continued

(Total 10 marks)

Q2

Leave blank

3. (a) Describe when you would use Spearman's rank correlation coefficient rather than the product moment correlation coefficient to measure the strength of the relationship between two variables.

(1)

A shop sells sunglasses and ice cream. For one week in the summer the shopkeeper ranked the daily sales of ice cream and sunglasses. The ranks are shown in the table below.

	Sun	Mon	Tues	Weds	Thurs	Fri	Sat
Ice cream	6	4	7	5	3	2	1
Sunglasses	6	5	7	2	3	4	1

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

(3)

(c) Test, at the 5% level of significance, whether or not there is a positive correlation between sales of ice cream and sales of sunglasses. State your hypotheses clearly.

(4)

The shopkeeper calculates the product moment correlation coefficient from his raw data and finds $r = 0.65$

(d) Using this new coefficient, test, at the 5% level of significance, whether or not there is a positive correlation between sales of ice cream and sales of sunglasses.

(2)

(e) Using your answers to part (c) and part (d), comment on the nature of the relationship between sales of sunglasses and sales of ice cream.

(1)

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

Leave blank

Question 3 continued

Q3

(Total 11 marks)

Leave blank

4. The weights of eggs are normally distributed with mean 60 g and standard deviation 5 g

Sairah chooses 2 eggs at random.

(a) Find the probability that the difference in weight of these 2 eggs is more than 2 g

(5)

Sairah is packing eggs into cartons. The weight of an empty egg carton is normally distributed with mean 40 g and standard deviation 1.5 g

(b) Find the distribution of the total weight of a carton filled with 12 randomly chosen eggs.

(3)

(c) Find the probability that a randomly chosen carton, filled with 12 randomly chosen eggs, weighs more than 800 g

(2)

Leave blank

Question 4 continued

Leave blank

Question 4 continued

Leave blank

Question 4 continued

Q4

(Total 10 marks)

Leave blank

5. A doctor claims there is a higher mean lung capacity in people who exercise regularly compared to people who do not exercise regularly. He measures the lung capacity, x, of 35 people who exercise regularly and 42 people who do not exercise regularly. His results are summarised in the table below.

	n	$\bar{x}$	s^2
Exercise regularly	35	26.3	12.2
Do not exercise regularly	42	24.8	10.1

(a) Test, at the 5% level of significance, the doctor's claim. State your hypotheses clearly.

(6)

(b) State any assumptions you have made in testing the doctor's claim.

(2)

The doctor decides to add another person who exercises regularly to his data. He measures the person's lung capacity and finds $x = 31.7$

(c) Find the unbiased estimate of the variance for the sample of 36 people who exercise regularly. Give your answer to 3 significant figures.

(4)

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

Leave blank

Question 5 continued

Q5

(Total 12 marks)

Leave blank

6. An airport manager carries out a survey of families and their luggage. Each family is allowed to check in a maximum of 4 suitcases. She observes 50 families at the check-in desk and counts the total number of suitcases each family checks in. The data are summarised in the table below.

Number of suitcases	0	1	2	3	4
Frequency	6	25	12	6	1

The manager claims that the data can be modelled by a binomial distribution with $p = 0.3$

(a) Test the manager's claim at the 5% level of significance. State your hypotheses clearly.
Show your working clearly and give your expected frequencies to 2 decimal places.

(8)

The manager also carries out a survey of the time taken by passengers to check in. She records the number of passengers that check in during each of 100 five-minute intervals.

The manager makes a new claim that these data can be modelled by a Poisson distribution. She calculates the expected frequencies given in the table below.

Number of passengers	0	1	2	3	4	5 or more
Observed frequency	5	40	31	18	6	0
Expected frequency	16.53	29.75	r	s	7.23	3.64

(b) Find the value of r and the value of s giving your answers to 2 decimal places.

(3)

(c) Stating your hypotheses clearly, use a 1% level of significance to test the manager's new claim.

(6)

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

Leave blank

Question 6 continued

Q6

(Total 17 marks)

Leave blank

7. A restaurant states that its hamburgers contain 20% fat. Paul claims that the mean fat content of their hamburgers is less than 20%. Paul takes a random sample of 50 hamburgers from the restaurant and finds that they contain a mean fat content of 19.5% with a standard deviation of 1.5%

You may assume that the fat content of hamburgers is normally distributed.

(a) Find the 90% confidence interval for the mean fat content of hamburgers from the restaurant.

(4)

(b) State, with a reason, what action Paul should recommend the restaurant takes over the stated fat content of their hamburgers.

(2)

The restaurant changes the mean fat content of their hamburgers to μ% and adjusts the standard deviation to 2%. Paul takes a sample of size n from this new batch of hamburgers. He uses the sample mean $\bar{X}$ as an estimator of μ.

(c) Find the minimum value of n such that $P(|\bar{X} - \mu| < 0.5) \geqslant 0.9$

(5)

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

Leave blank

Question 7 continued

Leave blank

Question 7 continued

Q7

(Total 11 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