Pleasechecktheexaminationdetailsbelowbeforeenteringyourcandidateinformation

Candidatesurname	Othernames

Centre Number | Candidate Number

Pearson Edexcel International Advanced Level

Time 1 hour 30 minutes

Paper reference **WST03/01**

Mathematics

International Advanced Subsidiary/Advanced Level
Statistics S3

You must have:
Mathematical Formulae and Statistical Tables (Yellow), calculator

TotalMarks

Candidates may use any calculator permitted 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).
- **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. If a calculator is used instead of the tables, the value should be given to an equivalent degree of accuracy.
- Inexact answers should be given to three significant figures unless otherwise stated.

Information

- A booklet 'Mathematical Formulae and Statistical Tables' is provided.
- There are 7 questions in this question paper. 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.
- If you change your mind about an answer, cross it out and put your new answer and any working underneath.

Turn over ▶

P71287A

©2021 Pearson Education Ltd.

E:1/1/1/1/

P 7 1 2 8 7 A 0 1 2 4

Pearson

Leave blank

1. A machine makes screws with a mean length of 30 mm and a standard deviation of 2.5 mm.

A manager claims that, following some repairs, the machine is now making screws with a mean length of less than 30 mm. The manager takes a random sample of 80 screws and finds that they have a mean length of 29.5 mm.

Use a suitable test, at the 5% level of significance, to determine whether there is evidence to support the manager's claim. State your hypotheses clearly.

(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 1 continued

Q1

(Total 5 marks)

Leave blank

2. Andy has some apple trees. Over many years she has graded each apple from her trees as *A*, *B*, *C*, *D* or *E* according to the quality of the apple, with *A* being the highest quality and *E* being the lowest quality.

She knows that the **proportion** of apples in each grade produced by her trees is as follows.

Grade	*A*	*B*	*C*	*D*	*E*
Proportion	4%	28%	52%	10%	6%

Raj advises Andy to add potassium to the soil around her apple trees.

Andy believes that adding potassium will not affect the distribution of grades for the quality of the apples.

To test her belief Andy adds potassium to the soil around her apple trees. The following year she counts the number of apples in each grade. The number of apples in each grade is shown in the table below.

Grade	*A*	*B*	*C*	*D*	*E*
Frequency	9	71	136	21	3

Test Andy's belief using a 5% level of significance. Show your working clearly, stating your hypotheses, expected frequencies and degrees of freedom.

(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

Q2

(Total 8 marks)

Leave blank

3. A cafe owner wishes to know whether the price of strawberry jam is related to the taste of the jam. He finds a website that lists the price per 100 grams and a mark for the taste, out of 100, awarded by a judge, for 9 different strawberry jams *A*, *B*, *C*, *D*, *E*, *F*, *G*, *H* and *I*. He then ranks the marks for taste and the prices.

The ranks are shown in the table below.

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

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

(5)

(b) Test, at the 5% level of significance, whether or not there is a relationship between the price and the taste of these strawberry jams. State your hypotheses clearly.

(3)

A friend suggests that it would be better to use the price per 100 grams, *c*, and the mark for the taste, *m*, for each strawberry jam rather than rank them.

Given that

$$S_{cc} = 2.0455 \qquad S_{mm} = 243.5556 \qquad S_{cm} = 16.4943$$

(c) calculate the product moment correlation coefficient between the price and the mark for taste of these strawberry jams, giving your answer correct to 3 decimal places.

(2)

(d) Use your value of the product moment correlation coefficient to test, at the 5% level of significance, whether or not there is evidence of a positive correlation between the price and the mark for taste of these 9 strawberry jams.
State your hypotheses clearly.

(3)

(e) State which of the tests in parts (b) and (d) is more appropriate for the cafe owner to use. Give a reason for your answer.

(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 14 marks)

Q3

Leave blank

4. A local village radio station, *LSB*, decides to survey adults in its broadcasting area about the programmes it produces.
LSB broadcasts to 4 villages A, B, C and D.
The number of households in each of the villages is given below.

Village	Number of households
A	41
B	164
C	123
D	82

LSB decides to take a stratified sample of 200 households.

(a) Explain how to select the households for this stratified sample.

(3)

One of the questions in the survey related to the age group of each member of the household and whether they listen to *LSB*. The data received are shown below.

	Age group		
	18–49	50–69	Older than 69
Listen to *LSB*	130	162	65
Do not listen to *LSB*	78	98	62

The data are to be used to determine whether or not there is an association between the age group and whether they listen to *LSB*.

(b) Calculate the expected frequencies for the age group 50–69 that

(i) listen to *LSB*

(ii) do not listen to *LSB*

(2)

Given that for the **other 4** classes $\sum \frac{(O-E)^2}{E} = 4.657$ to 3 decimal places,

(c) test at the 5% level of significance, whether or not there is evidence of an association between age and listening to *LSB*. Show your working clearly, stating the degrees of freedom and the critical value.

(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 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 11 marks)

Leave blank

5. Assam produces bags of flour. The stated weight printed on the bags of flour is 3 kg. The weights of the bags of flour are normally distributed with standard deviation 0.015 kg.

Assam weighs a random sample of 9 bags of flour and finds their mean weight is 2.977 kg.

(a) Calculate the 99% confidence interval for the mean weight of a bag of flour. Give your limits to 3 decimal places.

(3)

Assam decides to increase the amount of flour put into the bags.

(b) Explain why the confidence interval has led Assam to take this action.

(1)

After the increase a random sample of n bags of flour is taken. The sample mean weight of these n bags is 2.995 kg. A 95% confidence interval for μ gave a lower limit of less than 2.991 kg.

(c) Find the maximum value of n.

(4)

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 8 marks)

Leave blank

6. Amala believes that the resting heart rate is lower in men who exercise regularly compared to men who do not exercise regularly. She measures the resting heart rate, h, of a random sample of 50 men who exercise regularly and a random sample of 40 men who do not exercise regularly. Her results are summarised in the table below.

	Sample size	$\sum h$	$\sum h^2$	**Unbiased estimate of the mean**	**Unbiased estimate of the variance**
Exercise regularly	50	3270	214 676	α	β
Do not exercise regularly	40	2832	201 660	70.8	29.6

(a) Calculate the value of α and the value of β

(3)

(b) Test, at the 5% level of significance, whether there is evidence to support Amala's belief. State your hypotheses clearly.

(6)

(c) Explain the significance of the central limit theorem to the test in part (b).

(1)

(d) State two assumptions you have made in carrying out the test in part (b).

(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 12 marks)

Leave blank

7. A company produces bricks.

The weight of a brick, B kg, is such that $B \sim \mathrm{N}\left(1.96, \sqrt{0.003}^2\right)$

Two bricks are chosen at random.

(a) Find the probability that the difference in weight of the 2 bricks is greater than 0.1 kg

(5)

A random sample of n bricks is to be taken.

(b) Find the minimum sample size such that the probability of the sample mean being greater than 2 is less than 1%

(5)

The bricks are randomly selected and stacked on pallets.

The weight of an empty pallet, E kg, is such that $E \sim \mathrm{N}\left(21.8, \sqrt{0.6}^2\right)$

The random variable M represents the total weight of a pallet stacked with 500 bricks.

The random variable T represents the total weight of a container of cement.

Given that T is independent of M and that $T \sim \mathrm{N}\left(774, \sqrt{1.8}^2\right)$

(c) calculate $\mathrm{P}(4T > 100 + 3M)$

(7)

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

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 17 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