Please check the examination details below before entering your candidate information

Candidate surname	Other names

Centre Number | Candidate Number

Pearson Edexcel International Advanced Level

Thursday 13 June 2024

Morning (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

Total Marks

Candidates may use any calculator permitted by Pearson regulations. Calculators must not have the facility for symbolic algebraic 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 ▶

P75722A

F:1/1/1/1/1/1/1/

P 7 5 7 2 2 A 0 1 2 8

Pearson

1. The names of the 400 employees of a company are listed alphabetically in a book.

The chairperson of the company wishes to select a sample of 8 employees.

The chairperson numbers the employees from 001 to 400

(a) Describe how the list of numbers can be used to select a systematic sample of 8 employees.

(2)

(b) State one disadvantage of systematic sampling in this case.

(1)

(c) Write down the probability that the sample includes both the first name (employee 001) and the last name (employee 400) in the list.

(1)

Question 1 continued

(Total for Question 1 is 4 marks)

2. Aarush is asked to estimate the price of 7 kettles and rank them in order of decreasing price.

Aarush's order of decreasing price is *D A F C B G E*

The actual prices of the 7 kettles are shown in the table below.

Kettle	*A*	*B*	*C*	*D*	*E*	*F*	*G*
Price (£)	99.99	14.99	34.97	49.99	19.97	29.99	8.99

(a) Calculate Spearman's rank correlation coefficient between Aarush's order and the actual order.

Use a rank of 1 for the highest priced kettle.
Show your working clearly.

(4)

(b) Using a 5% level of significance, test whether or not there is evidence to suggest that Aarush is able to rank kettles in order of decreasing price.

You should state your hypotheses and critical value.

(4)

(c) Explain why Aarush did not use the product moment correlation coefficient in this situation.

(1)

Aarush discovered that kettle *A*'s price was recorded incorrectly and should have been £49.99 rather than £99.99

(d) Explain what effect this has on the rankings for the price.

(1)

Question 2 continued

Question 2 continued

Question 2 continued

(Total for Question 2 is 10 marks)

3. The volume of water in a bottle has a normal distribution with **unknown** mean, μ millilitres, and **known** standard deviation, σ millilitres.

A random sample of 150 of the bottles of water gave a 95% confidence interval for μ of

(327.84, 329.76)

(a) Using the confidence interval given, test whether or not $\mu = 328$

State your hypotheses clearly and write down the significance level you have used.

(3)

A second random sample, of 200 of these bottles of water, had a mean volume of 328 millilitres.

(b) Calculate a 98% confidence interval for μ based on this second sample.

You must show all steps in your working.
(*Solutions relying entirely on calculator technology are not acceptable*.)

(6)

Using five different random samples of 200 of these bottles of water, five 98% confidence intervals for μ are to be found.

(c) Calculate the probability that more than 3 of these intervals will contain μ

(3)

Question 3 continued

Question 3 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 3 continued

(Total for Question 3 is 12 marks)

4. The manager of a company making ice cream believes that the proportions of people in the population who prefer vanilla, chocolate, strawberry and other are in the ratio 10 : 5 : 2 : 3

The manager takes a random sample of 400 customers and records their age and favourite ice cream flavour. The results are shown in the table below.

		Ice cream flavour				
		Vanilla	Chocolate	Strawberry	Other	Total
Age	Child	95	25	13	25	158
	Teenager	57	20	17	36	130
	Adult	36	50	10	16	112
	Total	188	95	40	77	400

(a) Use the data in the table to test, at the 5% level of significance, the manager's belief. You should state your hypotheses, test statistic, critical value and conclusion clearly.

(8)

A researcher wants to investigate whether or not there is a relationship between the age of a customer and their favourite ice cream flavour. In order to test whether favourite ice cream flavour and age are related, the researcher plans to carry out a χ^2 test.

(b) Use the table to calculate expected frequencies for the group

(i) teenagers whose favourite ice cream flavour is vanilla,

(ii) adults whose favourite ice cream flavour is chocolate.

(2)

(c) Write down the number of degrees of freedom for this χ^2 test.

(1)

Question 4 continued

Question 4 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 4 continued

(Total for Question 4 is 11 marks)

5. A manager of a large company is investigating the time it takes the company's employees to complete a task.

The manager believes that the mean time for full-time employees to complete the task is more than a minute quicker than the mean time for part-time employees to complete the task.

The manager collects a random sample of 605 full-time employees and 45 part-time employees and records the times, t minutes, it takes each employee to complete the task.

The results are summarised in the table below.

	n	$\bar{t}$	s^2
Full-time employees	605	5.6	9
Part-time employees	45	7.0	4

(a) Test, at the 5% level of significance, the manager's claim.

You should state your hypotheses, test statistic, critical value and conclusion clearly.

(8)

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

(2)

The company increases the size of the sample of part-time employees to 46
The time taken to complete the task by the extra employee is 8 minutes.

(c) Find an unbiased estimate of the variance for the sample of 46 part-time employees.

(4)

Question 5 continued

Question 5 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 5 continued

(Total for Question 5 is 14 marks)

6. The weights of bags of carrots, C kg, are such that $C \sim N(1.2, 0.03^2)$

Three bags of carrots are selected at random.

(a) Calculate the probability that their total weight is more than 3.5 kg.

(4)

The weights of bags of potatoes, R kg, are such that $R \sim N(2.3, 0.03^2)$

Two bags of potatoes are selected at random.

(b) Calculate the probability that the difference in their weights is more than 0.05 kg.

(5)

The weights of trays, T kg, are such that $T \sim N(2.5, \sqrt{0.1}^2)$

The random variable G represents the total weight, in kg, of a single tray packed with 10 bags of potatoes where G and T are independent.

(c) Calculate $P(G < 2T + 20)$

(7)

Question 6 continued

Question 6 continued

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 6 continued

(Total for Question 6 is 16 marks)

7. The continuous random variable D is uniformly distributed over the interval $[x - 1, x + 5]$ where x is a constant.

A random sample of n observations of D is taken, where n is large.

(a) Use the Central Limit Theorem to find an approximate distribution for $\overline{D}$
Give your answer in terms of n and x where appropriate.

(3)

The n observations of D have a sample mean of 24.6

Given that the lower bound of the 99% confidence interval for x is 22.101 to 3 decimal places,

(b) find the value of n
Show your working clearly.

(5)

Question 7 continued

Question 7 continued

Question 7 continued

Question 7 continued

(Total for Question 7 is 8 marks)

TOTAL FOR PAPER IS 75 MARKS

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA