Please check the examination details below before entering your candidate information

Candidate surname StemUp Model Solution | Other names

Centre Number | Candidate Number

Pearson Edexcel International Advanced Level

Thursday 31 October 2024

Afternoon (Time: 1 hour 30 minutes) | Paper reference WST02/01

Mathematics

International Advanced Subsidiary/Advanced Level
Statistics S2

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 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 ▶

P78915A

©2024 Pearson Education Ltd.
V:1/1/1/

P 7 8 9 1 5 A 0 1 2 8

Pearson

1. During an annual beach-clean, the people doing the clean are asked to conduct a litter survey.
At a particular beach-clean, litter was found at a rate of 4 items per square metre.

(a) Find the probability that, in a randomly selected area of 2 square metres on this beach, exactly 5 items of litter were found.

(2)

Of the litter found on the beach, 30% of the items were face masks.

(b) Find the probability that, in a randomly selected area of 5 square metres on this beach, more than 4 face masks were found.

(2)

(c) Using a suitable approximation, find the probability that, in a randomly selected area of 20 square metres on this beach, less than 60 items of litter were found that were **not** face masks.

(4)

(a) X = number of items of litter in 2 m^2, $X \sim Po(8)$

$$P(X = 5) = \frac{e^{-8} \times 8^5}{5!} = 0.09160... = 0.0916$$

(b) Y = number of face masks in 5 m^2: $\lambda = 5 \times 4 \times 0.3 = 6$, $Y \sim Po(6)$

$P(Y > 4) = 1 - P(Y \leq 4) = 1 - 0.2851$

$= 0.7149$

(c) W = number of items that are not face masks in 20 m^2

$\lambda = 20 \times 4 \times 0.7 = 56$, $W \sim Po(56)$

λ large, so use the normal approximation $W \approx N(56, 56)$

$$P(W < 60) = P(W \leq 59) \approx P\left(Z < \frac{59.5 - 56}{\sqrt{56}}\right)$$

$= P(Z < 0.4677...) = P(Z < 0.47)$

$= 0.6808$

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

Question 1 continued

(Total for Question 1 is 8 marks)

2. A multiple-choice test consists of 25 questions, each having 5 responses, only one of which is correct.

Each correct answer gains 4 marks but each incorrect answer loses 1 mark.

Sam answers all 25 questions by choosing at random one response for each question.

Let X be the number of correct answers that Sam achieves.

(a) State the distribution of X

(1)

Let M be the number of marks that Sam achieves.

(b) (i) State the distribution of M in terms of X

(ii) Hence, show clearly that the number of marks that Sam is expected to achieve is zero.

(4)

In order to pass the test at least 30 marks are required.

(c) Find the probability that Sam will pass the test.

(3)

Past records show that when the test is done properly, the probability that a student answers the first question correctly is 0.5

A random sample of 50 students that did the test properly was taken.

Given that the probability that more than n but at most 30 students answered the first question correctly was 0.9328 to 4 decimal places,

(d) find the value of n

(4)

(a) $X \sim B\left(25, \frac{1}{5}\right)$ (binomial, $n = 25$, $p = 0.2$)

(b) (i) $M = 4X - 1 \times (25 - X)$

$M = 5X - 25$

(ii) $E(X) = np = 25 \times \frac{1}{5} = 5$

$E(M) = E(5X - 25) = 5E(X) - 25$

$E(M) = 5 \times 5 - 25 = 0$

So the expected number of marks is zero, as required.

(c) Pass: $M \geqslant 30 \Rightarrow 5X - 25 \geqslant 30 \Rightarrow 5X \geqslant 55 \Rightarrow X \geqslant 11$

$P(X \geqslant 11) = 1 - P(X \leqslant 10) = 1 - 0.9944$

$= 0.0056$

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

Question 2 continued

(d) Y = number of the 50 students answering the first question correctly

$Y \sim B(50, 0.5)$

$P(n < Y \leqslant 30) = P(Y \leqslant 30) - P(Y \leqslant n) = 0.9328$

$0.9405 - P(Y \leqslant n) = 0.9328$

$P(Y \leqslant n) = 0.9405 - 0.9328 = 0.0077$

From tables, $P(Y \leqslant 16) = 0.0077$

$n = 16$

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

Question 2 continued

(Total for Question 2 is 12 marks)

3. During Monday afternoons, customers are known to enter a certain shop at a mean rate of 7 customers every 10 minutes.

(a) Suggest a suitable distribution to model the number of customers that enter this shop in a 10-minute interval on Monday afternoons.

(1)

(b) State **two** assumptions necessary for this distribution to be a suitable model of this situation.

(2)

A new shop manager wants to find out if the rate of customers has changed since they took over.

(c) Write down suitable null and alternative hypotheses that the shop manager should use.

(1)

The shop manager decides to monitor the number of customers entering the shop in a random 10-minute interval next Monday afternoon.

(d) Using a 3% level of significance, find the critical region to test whether the rate of customers has changed.

(3)

(e) Find the actual significance level of this test based on your critical region from part (d)

(2)

During the random 10-minute interval that Monday afternoon, 12 customers entered the shop.

(f) Comment on this finding, using the critical region in part (d)

(2)

(a) X = number of customers entering in 10 minutes, $X \sim Po(7)$

(b) Customers enter the shop singly (one at a time).
Customers enter the shop independently of each other, at a constant average rate.

(c) $H_0: \lambda = 7$ $H_1: \lambda \neq 7$

(d) Under H_0, $X \sim Po(7)$; two-tailed, 1.5% in each tail
Lower tail: $P(X \leqslant 1) = 0.0073 < 0.015$, $P(X \leqslant 2) = 0.0296 > 0.015$
Upper tail: $P(X \geqslant 13) = 1 - 0.9730 = 0.0270 > 0.015$
$P(X \geqslant 14) = 1 - 0.9872 = 0.0128 < 0.015$
Critical region: $X \leqslant 1$ or $X \geqslant 14$

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

Question 3 continued

(e) Actual significance level = $P(X \leq 1) + P(X \geq 14)$
$= 0.0073 + 0.0128 = 0.0201$ (2.01%)

(f) 12 is not in the critical region, so do not reject H_0
There is insufficient evidence that the rate of customers entering the shop has changed.

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

Question 3 continued

(Total for Question 3 is 11 marks)

4. (i) The continuous random variable X is uniformly distributed over the interval $[a, b]$

Given that

- $P(X > 27) = \frac{3}{4}$
- $\text{Var}(X) = 300$

(a) find the value of a and the value of b

(3)

Given also that

$$4 \times P(X < k - 10) = P(X > k + 20)$$

(b) find the value of k

(2)

(ii) A piece of wire of length 42 cm is cut into 2 pieces at a random point.

Each of the two pieces of the wire is bent to form the outline of a square.

Find the probability that the side length of the larger square minus the side length of the smaller square will be greater than 2 cm.

(4)

(i) (a) $P(X > 27) = \frac{b - 27}{b - a} = \frac{3}{4}$ $\quad \text{Var}(X) = \frac{(b-a)^2}{12} = 300$

$(b - a)^2 = 3600 \Rightarrow b - a = 60$

$b - 27 = \frac{3}{4} \times 60 = 45 \Rightarrow b = 72, \quad a = 72 - 60 = 12$

$a = 12, \quad b = 72$

(b) $X \sim U(12, 72)$: $4 \times P(X < k - 10) = P(X > k + 20)$

$4 \times \frac{(k - 10) - 12}{60} = \frac{72 - (k + 20)}{60}$

$4(k - 22) = 52 - k \Rightarrow 4k - 88 = 52 - k \Rightarrow 5k = 140$

$k = 28$

(ii) L = length of one piece (cm), $L \sim U(0, 42)$

Side lengths of the squares are $\frac{L}{4}$ and $\frac{42 - L}{4}$

Require $\frac{L}{4} - \frac{42 - L}{4} > 2$ or $\frac{42 - L}{4} - \frac{L}{4} > 2$

$2L - 42 > 8 \Rightarrow L > 25$ or $42 - 2L > 8 \Rightarrow L < 17$

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

Question 4 continued

$$P(L < 17) + P(L > 25) = \frac{17 - 0}{42} + \frac{42 - 25}{42} = \frac{34}{42}$$

$$= \frac{17}{21} \quad (= 0.810)$$

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

Question 4 continued

(Total for Question 4 is 9 marks)

5. The continuous random variable X has a probability density function given by

$$f(x) = \begin{cases} \frac{1}{4}(3-x) & 1 \leqslant x \leqslant 2 \\ \frac{1}{4} & 2 < x \leqslant 3 \\ \frac{1}{4}(x-2) & 3 < x \leqslant 4 \\ 0 & \text{otherwise} \end{cases}$$

The cumulative distribution function of X is $F(x)$

(a) Show that $F(x) = \frac{1}{4}\left(3x - \frac{x^2}{2}\right) - \frac{5}{8}$ for $1 \leqslant x \leqslant 2$

(2)

(b) Find $F(x)$ for all values of x

(5)

(c) Find $P(1.2 < X < 3.1)$

(2)

(a) $F(x) = \int_1^x \frac{1}{4}(3-t)\,dt = \frac{1}{4}\left[3t - \frac{t^2}{2}\right]_1^x$

$= \frac{1}{4}\left[\left(3x - \frac{x^2}{2}\right) - \left(3 - \frac{1}{2}\right)\right] = \frac{1}{4}\left(3x - \frac{x^2}{2}\right) - \frac{5}{8}$

So $F(x) = \frac{1}{4}\left(3x - \frac{x^2}{2}\right) - \frac{5}{8}$ for $1 \leqslant x \leqslant 2$, as required

(b) $F(2) = \frac{1}{4}(6-2) - \frac{5}{8} = \frac{3}{8}$

$2 < x \leqslant 3$: $F(x) = \frac{3}{8} + \int_2^x \frac{1}{4}\,dt = \frac{3}{8} + \frac{x-2}{4} = \frac{1}{8}(2x-1)$

$F(3) = \frac{1}{8}(6-1) = \frac{5}{8}$

$3 < x \leqslant 4$: $F(x) = \frac{5}{8} + \int_3^x \frac{1}{4}(t-2)\,dt = \frac{5}{8} + \frac{1}{4}\left[\frac{t^2}{2} - 2t\right]_3^x$

$= \frac{5}{8} + \frac{1}{4}\left(\frac{x^2}{2} - 2x\right) + \frac{3}{8} = \frac{1}{4}\left(\frac{x^2}{2} - 2x\right) + 1$

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

Question 5 continued

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{4}\left(3x - \frac{x^2}{2}\right) - \frac{5}{8} & 1 \leqslant x \leqslant 2 \\ \frac{1}{8}(2x - 1) & 2 < x \leqslant 3 \\ \frac{1}{4}\left(\frac{x^2}{2} - 2x\right) + 1 & 3 < x \leqslant 4 \\ 1 & x > 4 \end{cases}$$

(c) $P(1.2 < X < 3.1) = F(3.1) - F(1.2)$

$$= \left[\frac{1}{4}\left(\frac{3.1^2}{2} - 2 \times 3.1\right) + 1\right] - \left[\frac{1}{4}\left(3 \times 1.2 - \frac{1.2^2}{2}\right) - \frac{5}{8}\right]$$

$$= 0.65125 - 0.095$$

$$= 0.55625 = \frac{89}{160} \quad (= 0.556)$$

Question 5 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

Question 5 continued

(Total for Question 5 is 9 marks)

6. Two boxes, A and B, each contain a large number of coins.

In box A

- there are only 1p coins and 2p coins
- the ratio of 1p coins to 2p coins is 1 : 3

In box B

- there are only 2p coins and 5p coins
- the ratio of 2p coins to 5p coins is 1 : 4

One coin is randomly selected from box A and two coins are randomly selected from box B

The random variable T represents the total of the values of the three coins selected.

(a) Find the sampling distribution of T

(7)

The random variable M represents the median of the values of the three coins selected.

(b) Find the sampling distribution of M

(4)

(a) Box A: $P(1p) = \frac{1}{4}$, $P(2p) = \frac{3}{4}$ Box B: $P(2p) = \frac{1}{5}$, $P(5p) = \frac{4}{5}$

Possible selections (A, B, B): (1, 2, 2), (2, 2, 2), (1, 2, 5), (1, 5, 2), (2, 2, 5), (2, 5, 2), (1, 5, 5), (2, 5, 5)

Totals: $T = 5, 6, 8, 9, 11, 12$

$P(T = 5) = P(1, 2, 2) = \frac{1}{4} \times \frac{1}{5} \times \frac{1}{5} = \frac{1}{100}$

$P(T = 6) = P(2, 2, 2) = \frac{3}{4} \times \frac{1}{5} \times \frac{1}{5} = \frac{3}{100}$

$P(T = 8) = P(1, 2, 5) \times 2 = 2 \times \frac{1}{4} \times \frac{1}{5} \times \frac{4}{5} = \frac{8}{100}$

$P(T = 9) = P(2, 2, 5) \times 2 = 2 \times \frac{3}{4} \times \frac{1}{5} \times \frac{4}{5} = \frac{24}{100}$

$P(T = 11) = P(1, 5, 5) = \frac{1}{4} \times \frac{4}{5} \times \frac{4}{5} = \frac{16}{100}$

$P(T = 12) = P(2, 5, 5) = \frac{3}{4} \times \frac{4}{5} \times \frac{4}{5} = \frac{48}{100}$

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

Question 6 continued

t	5	6	8	9	11	12
$P(T = t)$	$\frac{1}{100}$	$\frac{3}{100}$	$\frac{2}{25}$	$\frac{6}{25}$	$\frac{4}{25}$	$\frac{12}{25}$

(b) Medians: (1, 2, 2) → 2, (2, 2, 2) → 2, (1, 2, 5) → 2, (2, 2, 5) → 2, (1, 5, 5) → 5, (2, 5, 5) → 5, so $M = 2$ or 5

$$P(M = 2) = \frac{1}{100} + \frac{3}{100} + \frac{8}{100} + \frac{24}{100} = \frac{36}{100} = \frac{9}{25}$$

$$P(M = 5) = \frac{16}{100} + \frac{48}{100} = \frac{64}{100} = \frac{16}{25}$$

m	2	5
$P(M = m)$	$\frac{9}{25}$	$\frac{16}{25}$

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

Question 6 continued

(Total for Question 6 is 11 marks)

7. The continuous random variable X has probability density function given by

$$f(x) = \begin{cases} ax & 0 \leqslant x \leqslant 4 \\ bx + c & 4 < x \leqslant 8 \\ 0 & \text{otherwise} \end{cases}$$

where a, b and c are constants.

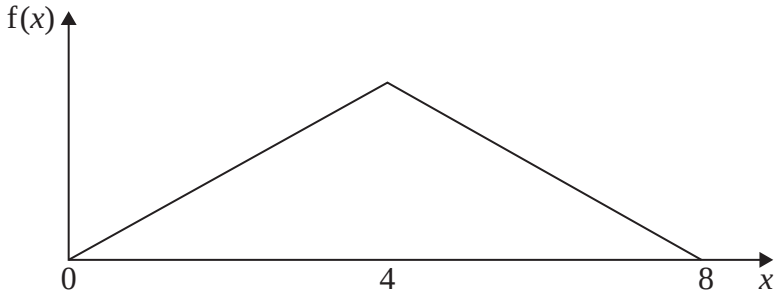


Figure 1

Figure 1 shows the graph of the probability density function $f(x)$

The graph consists of two straight line segments of equal length joined at the point where $x = 4$

(a) Show that $a = \frac{1}{16}$

(1)

(b) Hence find

(i) the value of b

(ii) the value of c

(3)

(c) Using algebraic integration, show that $\text{Var}(X) = \frac{8}{3}$

(6)

(d) Find, to 2 decimal places, the lower quartile and the upper quartile of X

(3)

A statistician claims that

$$P(-\sigma < X - \mu < \sigma) > 0.5$$

where μ and σ are the mean and standard deviation of X

(e) Show that the statistician's claim is correct.

(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

Question 7 continued

(a) $f(4) = 4a$, and the total area under $f(x)$ is 1:

Area of triangle $= \frac{1}{2} \times 8 \times 4a = 16a = 1 \Rightarrow a = \frac{1}{16}$ as required

(b) (i) By symmetry (equal length segments), $b = -\frac{1}{16}$

(ii) At (8, 0): $0 = -\frac{1}{16} \times 8 + c \Rightarrow c = \frac{1}{2}$

(c) By symmetry, $E(X) = 4$

$$E(X^2) = \int_0^4 \frac{x^3}{16}\,dx + \int_4^8 \left(-\frac{x^3}{16} + \frac{x^2}{2}\right)dx$$

$$= \left[\frac{x^4}{64}\right]_0^4 + \left[-\frac{x^4}{64} + \frac{x^3}{6}\right]_4^8$$

$$= 4 + \left[\left(-64 + \frac{256}{3}\right) - \left(-4 + \frac{32}{3}\right)\right] = 4 + \frac{44}{3} = \frac{56}{3}$$

$Var(X) = E(X^2) - [E(X)]^2 = \frac{56}{3} - 4^2 = \frac{8}{3}$ as required

(d) $\int_0^{Q_1} \frac{x}{16}\,dx = 0.25 \Rightarrow \frac{Q_1^2}{32} = 0.25 \Rightarrow Q_1^2 = 8$

$Q_1 = \sqrt{8} = 2.828\ldots = 2.83$

By symmetry, $Q_3 = 8 - \sqrt{8} = 5.171\ldots = 5.17$

(e) $\mu = 4$, $\sigma = \sqrt{\left(\frac{8}{3}\right)} = 1.6329\ldots$, $\mu - \sigma = 2.367\ldots$, $\mu + \sigma = 5.632\ldots$

$P(-\sigma < X - \mu < \sigma) = P(2.37 < X < 5.63)$

$P(Q_1 < X < Q_3) = 0.5$ and $2.37 < 2.83$, $5.63 > 5.17$

so the interval (2.37, 5.63) contains the interval (Q_1, Q_3)

(Indeed $P(2.37 < X < 5.63) = 1 - 2 \times \frac{2.367\ldots^2}{32} = 0.6498\ldots$)

$\therefore$ $P(-\sigma < X - \mu < \sigma) > 0.5$, so the statistician's claim is correct.

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

Question 7 continued

Question 7 continued

(Total for Question 7 is 15 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