Please check the examination details below before entering your candidate information

Candidate surname	StemUp Model Solution	Other names

Pearson Edexcel
International
Advanced Level

Centre Number	Candidate Number

Tuesday 19 January 2021

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 (Blue), 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 6 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 ▶

P66442A

©2021 Pearson Education Ltd.

1/1/1/1/

P 6 6 4 4 2 A 0 1 2 4

Pearson

Leave blank

1. Jim farms oysters in a particular lake. He knows from past experience that 5% of young oysters do not survive to be harvested.

In a random sample of 30 young oysters, the random variable X represents the number that do not survive to be harvested.

(a) Write down a suitable model for the distribution of X.

(1)

(b) State an assumption that has been made for the model in part (a).

(1)

(c) Find the probability that

(i) exactly 24 young oysters **do survive** to be harvested,

(ii) at least 3 young oysters do not survive to be harvested.

(4)

A second random sample, of 200 young oysters, is taken.
The probability that at least n of these young oysters do not survive to be harvested is more than 0.8

(d) Using a suitable approximation, find the maximum value of n.

(3)

Jim believes that the level of salt in the lake water has changed and it has altered the survival rate of his oysters. He takes a random sample of 25 young oysters and places them in the lake.
When Jim harvests the oysters, he finds that 21 **do survive** to be harvested.

(e) Use a suitable test, at the 5% level of significance, to assess whether or not there is evidence that the proportion of oysters not surviving to be harvested is more than 5%. State your hypotheses clearly.

(5)

(a) $X \sim B(30, 0.05)$

(b) The probability that an oyster does not survive is constant (0.05) and each oyster survives or not independently of the others.

(c)(i) 24 survive $\Rightarrow$ $X = 6$

$P(X = 6) = {}^{30}C_6 (0.05)^6 (0.95)^{24}$

$= 0.0027089... = 0.00271$

(ii) $P(X \geq 3) = 1 - P(X \leq 2)$

$= 1 - 0.8122$

$= 0.1878$

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

(d) $Y \sim B(200, 0.05)$, n large and p small $\Rightarrow$ $Y \approx A \sim Po(10)$

$P(A \geqslant n) > 0.8 \Rightarrow P(A \leqslant n - 1) < 0.2$

$P(A \leqslant 6) = 0.1301 < 0.2$, $P(A \leqslant 7) = 0.2202 > 0.2$

so $n - 1 = 6$

maximum value $n = 7$

(e) $H_0 : p = 0.05$ $H_1 : p > 0.05$

C = number not surviving in 25, $C \sim B(25, 0.05)$ under H_0

21 survive $\Rightarrow$ 4 do not survive

$P(C \geqslant 4) = 1 - P(C \leqslant 3) = 1 - 0.9659$

$= 0.0341$

$0.0341 < 0.05$ so the result is significant: reject H_0

There is evidence that the proportion of oysters not surviving has increased (Jim's belief is supported).

Leave blank

Question 1 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 1 continued

(Total 14 marks)

Q1

Leave blank

2. The distance, in metres, a novice tightrope artist, walking on a wire, walks before falling is modelled by the random variable W with cumulative distribution function

$$F(w) = \begin{cases} 0 & w < 0 \\ \frac{1}{3}\left(w - \frac{w^4}{256}\right) & 0 \leqslant w \leqslant 4 \\ 1 & w > 4 \end{cases}$$

(a) Find the probability that a novice tightrope artist, walking on the wire, walks at least 3.5 metres before falling.

(2)

A random sample of 30 novice tightrope artists is taken.

(b) Find the probability that more than 1 of these novice tightrope artists, walking on the wire, walks at least 3.5 metres before falling.

(3)

Given $E(W) = 1.6$

(c) use algebraic integration to find $\text{Var}(W)$

(5)

(a) $P(W \geqslant 3.5) = 1 - F(3.5) = 1 - \frac{1}{3}\left(3.5 - \frac{3.5^4}{256}\right)$

$= 1 - 0.971272...$

$= 0.028727... = 0.0287$

(b) $Y \sim B(30, 0.028727...)$

$P(Y > 1) = 1 - P(Y \leqslant 1)$

$= 1 - [\,(0.971272...)^{30} + 30(0.028727...)(0.971272...)^{29}\,]$

$= 1 - (0.417096... + 0.370092...)$

$= 1 - 0.787188... = 0.21281... = 0.213$

(c) $f(w) = \frac{dF}{dw} = \frac{1}{3}\left(1 - \frac{w^3}{64}\right)$ for $0 \leqslant w \leqslant 4$

$E(W^2) = \int_0^4 \frac{1}{3}\left(w^2 - \frac{w^5}{64}\right) dw$

$= \frac{1}{3}\left[\frac{w^3}{3} - \frac{w^6}{384}\right]_0^4$

$= \frac{1}{3}\left(\frac{64}{3} - \frac{4096}{384}\right) = \frac{1}{3} \times \frac{32}{3} = \frac{32}{9}$

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Leave blank

Question 2 continued

$\text{Var}(W) = E(W^2) - [E(W)]^2 = \frac{32}{9} - 1.6^2$

$= \frac{224}{225}$ $(= 0.99555\ldots)$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

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. The number of water fleas, in 100 ml of pond water, has a Poisson distribution with mean 7

(a) Find the probability that a sample of 100 ml of the pond water does **not** contain exactly 4 water fleas.

(2)

Aja collects 5 separate samples, each of 100 ml, of the pond water.

(b) Find the probability that exactly 1 of these samples contains exactly 4 water fleas.

(3)

Using a normal approximation, the probability that more than 3 water fleas will be found in a random sample of n ml of the pond water is 0.9394 correct to 4 significant figures.

(c) (i) Show that $n - 1.55\sqrt{\frac{n}{0.07}} - 50 = 0$

(5)

(ii) Hence find the value of n

(2)

After the pond has been cleaned, the number of water fleas in a 100 ml random sample of the pond water is 15

(d) Using a suitable test, at the 1% level of significance, assess whether or not there is evidence that the number of water fleas per 100 ml of the pond water has increased. State your hypotheses clearly.

(5)

(a) $X \sim Po(7)$

$P(X \neq 4) = 1 - P(X = 4) = 1 - \frac{e^{-7}\,7^4}{4!}$

$= 1 - 0.091226...$

$= 0.908773... = 0.909$

(b) Y = number of samples with exactly 4 fleas, $Y \sim B(5, 0.091226...)$

$P(Y = 1) = {}^5C_1\,(0.091226...)(0.908773...)^4$

$= 0.31110... = 0.311$

(c)(i) Y = fleas in n ml, $Y \sim Po(0.07n)$

approximate by $N(0.07n, 0.07n)$

$P(Y > 3) \approx P(N > 3.5) = 0.9394$

$\frac{3.5 - 0.07n}{\sqrt{(0.07n)}} = -1.55$

$0.07n - 3.5 = 1.55\sqrt{(0.07n)}$

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Question 3 continued

$\div 0.07: \quad n - 50 = \frac{1.55\sqrt{(0.07n)}}{0.07} = 1.55\frac{\sqrt{n}}{\sqrt{0.07}}$

$n - 50 = 1.55\sqrt{\frac{n}{0.07}}$

$\therefore \quad n - 1.55\sqrt{\frac{n}{0.07}} - 50 = 0$

(ii) $(\sqrt{n})^2 - \frac{1.55}{\sqrt{0.07}}\sqrt{n} - 50 = 0$

$\sqrt{n} = \frac{5.85844... \pm \sqrt{(5.85844...^2 + 200)}}{2}$

$\sqrt{n} = 10.583...$ (or $-4.7245...$, rejected as $\sqrt{n} > 0$)

$n = 10.583...^2 = 112$

(d) $H_0 : \lambda = 7 \qquad H_1 : \lambda > 7$

Under H_0, $X \sim Po(7)$

$P(X \geqslant 15) = 1 - P(X \leqslant 14) = 1 - 0.9943$

$= 0.0057$

$0.0057 < 0.01$ so the result is significant: reject H_0

There is evidence that the number of water fleas per 100 ml of the pond water has increased.

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

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 17 marks)

Q3

Leave blank

4. A continuous random variable X has probability density function

$$f(x) = \begin{cases} k(a-x)^2 & 0 \leqslant x \leqslant a \\ 0 & \text{otherwise} \end{cases}$$

where k and a are constants.

(a) Show that $ka^3 = 3$

(3)

Given that $E(X) = 1.5$

(b) use algebraic integration to show that $a = 6$

(4)

(c) Verify that the median of X is 1.2 to one decimal place.

(3)

(a) $\int_0^a k(a-x)^2\,dx = 1$

$\left[\frac{-k(a-x)^3}{3}\right]_0^a = 1$

$0 - \left(\frac{-ka^3}{3}\right) = 1 \Rightarrow \frac{ka^3}{3} = 1$

$\therefore ka^3 = 3$

(b) $E(X) = \int_0^a kx(a-x)^2\,dx = \int_0^a k(a^2x - 2ax^2 + x^3)\,dx$

$= \left[k\left(\frac{a^2x^2}{2} - \frac{2ax^3}{3} + \frac{x^4}{4}\right)\right]_0^a$

$= k\left(\frac{a^4}{2} - \frac{2a^4}{3} + \frac{a^4}{4}\right) = \frac{ka^4}{12}$

$\frac{ka^4}{12} = 1.5 \Rightarrow ka^4 = 18$

$\frac{ka^4}{ka^3} = \frac{18}{3} \Rightarrow a = 6$

(c) $k = \frac{3}{6^3} = \frac{1}{72}$

$F(x) = \int_0^x \frac{1}{72}(36 - 12t + t^2)\,dt = \frac{1}{72}\left(36x - 6x^2 + \frac{x^3}{3}\right)$

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

$F(1.15) = 0.47183... < 0.5$

$F(1.25) = 0.50383... > 0.5$

Since $F(1.15) < 0.5 < F(1.25)$, the median lies between 1.15 and 1.25

$\therefore$ the median of X is 1.2 to one decimal place.

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 10 marks)

Q4

Leave blank

5. A piece of wood AB is 3 metres long. The wood is cut at random at a point C and the random variable W represents the length of the piece of wood AC.

(a) Find the probability that the length of the piece of wood AC is more than 1.8 metres.

(2)

The two pieces of wood AC and CB form the two shortest sides of a right-angled triangle. The random variable X represents the length of the longest side of the right-angled triangle.

(b) Show that $X^2 = 2W^2 - 6W + 9$

(2)

[You may assume for random variables S, T and U and for constants a and b that if $S = aT + bU$ then $\mathrm{E}(S) = a\mathrm{E}(T) + b\mathrm{E}(U)$]

(c) Find $\mathrm{E}(X^2)$

(6)

(d) Find $\mathrm{P}(X^2 > 5)$

(4)

(a) $W \sim U[0, 3]$

$P(W > 1.8) = \frac{3 - 1.8}{3} = 0.4$

(b) $AC = W$, $CB = 3 - W$. By Pythagoras:

$X^2 = W^2 + (3 - W)^2$

$X^2 = W^2 + 9 - 6W + W^2$

$X^2 = 2W^2 - 6W + 9$

(c) $E(W) = \frac{0 + 3}{2} = 1.5$

$Var(W) = \frac{(3 - 0)^2}{12} = \frac{9}{12} = \frac{3}{4}$

$E(W^2) = Var(W) + [E(W)]^2 = \frac{3}{4} + 1.5^2 = 3$

$E(X^2) = 2E(W^2) - 6E(W) + 9$

$= 2 \times 3 - 6 \times 1.5 + 9$

$= 6$

(d) $P(X^2 > 5) = P(2W^2 - 6W + 9 > 5) = P(2W^2 - 6W + 4 > 0)$

$= P(2(W - 1)(W - 2) > 0)$

critical values $W = 1$, $W = 2$ $\Rightarrow$ $W < 1$ or $W > 2$

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

$= P(W < 1) + P(W > 2) = \frac{1}{3} + \frac{1}{3}$

$= \frac{2}{3}$

Leave blank

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

Leave blank

Question 5 continued

(Total 14 marks)

Q5

Leave blank

6. The owner of a very large youth club has designed a new method for allocating people to teams. Before introducing the method he decided to find out how the members of the youth club might react.

(a) Explain why the owner decided to take a random sample of the youth club members rather than ask all the youth club members.

(1)

(b) Suggest a suitable sampling frame.

(1)

(c) Identify the sampling units.

(1)

The new method uses a bag containing a large number of balls. Each ball is numbered either 20, 50 or 70
When a ball is selected at random, the random variable X represents the number on the ball where

$$P(X = 20) = p \qquad P(X = 50) = q \qquad P(X = 70) = r$$

A youth club member takes a ball from the bag, records its number and replaces it in the bag.
He then takes a second ball from the bag, records its number and replaces it in the bag.

The random variable M is the mean of the 2 numbers recorded.

Given that

$$P(M = 20) = \frac{25}{64} \qquad P(M = 60) = \frac{1}{16} \qquad \text{and} \quad q > r$$

(d) show that $P(M = 50) = \frac{1}{16}$

(7)

(a) Taking a sample is quicker and cheaper than asking every member.

(b) A list (register) of all the members of the youth club.

(c) The (individual) youth club members.

(d) $P(M = 20) = p^2 = \frac{25}{64} \Rightarrow p = \frac{5}{8}$

$p + q + r = 1 \Rightarrow q + r = \frac{3}{8}$ 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 6 continued

$M = 60$ from $(50, 70)$ or $(70, 50)$: $2qr = \frac{1}{16} \Rightarrow qr = \frac{1}{32}$ ②

Sub $r = \frac{3}{8} - q$ into ②: $q\left(\frac{3}{8} - q\right) = \frac{1}{32}$

$\frac{3}{8}q - q^2 = \frac{1}{32} \Rightarrow 32q^2 - 12q + 1 = 0$

$(8q - 1)(4q - 1) = 0$

$q = \frac{1}{4}$, $r = \frac{1}{8}$ or $q = \frac{1}{8}$, $r = \frac{1}{4}$

$q > r \Rightarrow q = \frac{1}{4}$, $r = \frac{1}{8}$

$M = 50$ only from $(50, 50)$:

$P(M = 50) = q^2 = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$ as required

Leave blank

Question 6 continued

Q6

(Total 10 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