Please check the examination details below before entering your candidate information

Candidate surname	Other names

Pearson Edexcel
International Advanced Level

Centre Number | Candidate Number

Wednesday 23 January 2019

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

Statistics S2

Advanced/Advanced Subsidiary

You must have:
Mathematical Formulae and Statistical Tables (Blue)

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

Pearson

Leave blank

1. A bus company sells tickets for a journey from London to Oxford every Saturday. Past records show that 5% of people who buy a ticket do not turn up for the journey.

The bus has seats for 48 people.
Each week the bus company sells tickets to exactly 50 people for the journey.

The random variable X represents the number of these people who do not turn up for the journey.

(a) State one assumption required to model X as a binomial distribution.
(1)

For this week's journey find,

(b) (i) the probability that all 50 people turn up for the journey,

(ii) $P(X = 1)$
(3)

The bus company receives £20 for each ticket sold and all 50 tickets are sold. It must pay out £60 to each person who buys a ticket and turns up for the journey but does not have a seat.

(c) Find the bus company's expected total earnings per journey.
(3)

(a) Each person turns up independently from others.

b) $X \sim B(50, 0.05)$ → don't turn up.

$P(X=0) \binom{50}{0}(0.05)^0(0.95)^{50}$

$= \underline{0.0769}$

$P(X=1) \binom{50}{1}(0.05)^1(0.95)^{49}$

$= \underline{0.202}$

(c) Tickets sold → $50 \times £20$
$= £1000$

Probability that all 50 people turn up for the journey $= 0.0769$

$2 \times £60 = £120$

$1000 - 120 = £880.$

$880 \times 0.0769 = £67.67.$

Probability that one person doesn't turn up for the journey $= 0.202$

$1 \times £60 = £60$

$1000 - 60 = £940$

$940 \times 0.202 = 189.88.$

Probability that 2 or more people don't turn up for the journey
$1 - 0.0769 - 0.202$
$= 0.7211.$

Leave blank

Question 1 continued

$1000 \times 0.7211 = 721.$

$67.67 + 189.88 + 721.1$

$= 978.65$

$= \underline{\underline{£979}}$

Q1

(Total 7 marks)

Turn over

Leave blank

2. During morning hours, employees arrive randomly at an office drinks dispenser at a rate of 2 every 10 minutes.

The number of employees arriving at the drinks dispenser is assumed to follow a Poisson distribution.

(a) Find the probability that fewer than 5 employees arrive at the drinks dispenser during a 10-minute period one morning.
(2)

During a 30-minute period one morning, the probability that n employees arrive at the drinks dispenser is the same as the probability that $n + 1$ employees arrive at the drinks dispenser.

(b) Find the value of n
(3)

During a 45-minute period one morning, the probability that between c and 12, inclusive, employees arrive at the drinks dispenser is 0.8546

(c) Find the value of c
(3)

(d) Find the probability that exactly 2 employees arrive at the drinks dispenser in exactly 4 of the 6 non-overlapping 10-minute intervals between 10 am and 11am one morning.
(4)

(a) $X \sim Po(2)$

$P(X<5) = P(X \leq 4)$

$= \underline{0.947}$

(b) 2 → 10 minutes
x → 30 minutes

$\underline{x = 6}$

$X \sim Po(6)$

$P(X=n) = P(X=n+1)$

$$\frac{e^{-6} \times 6^n}{n!} = \frac{e^{-6} \times 6^{n+1}}{(n+1)!}$$

$$\therefore \frac{6^n}{n!} = \frac{6^{n+1}}{(n+1)!}$$

$n=1 \rightarrow 6 \neq 18$ ×
$n=2 \rightarrow 18 \neq 36$ ×
$n=3 \rightarrow 36 \neq 54$ ×
$n=4 \rightarrow 54 \neq 64.8$ ×
$n=5 \rightarrow 64.8 = 64.8$ ✓

$\therefore \underline{\underline{n=5}}$

Leave blank

Question 2 continued

(c) $2 \rightarrow 10$mins
$x \rightarrow 45$mins.

$x = 9$

$X \sim Po(9)$

$P(C \leq X \leq 12) = 0.8546$

$0.8758 - P(X \leq C-1) = 0.8546$

$P(X \leq C-1) = 0.0212$

$C - 1 = 3$
$C = 4$

(d) $P(X = 2)$ $X \sim Po(2)$.

$\frac{e^{-2} \times 2^2}{2!} = 0.27067$

$X \sim B(6, 0.2707)$

$P(X = 4)$ $\binom{6}{4}(0.2707)^4(0.7293)^2$

$= 0.043$

Leave blank

3. Figure 1 shows an accurate graph of the cumulative distribution function, F(x), for the continuous random variable X

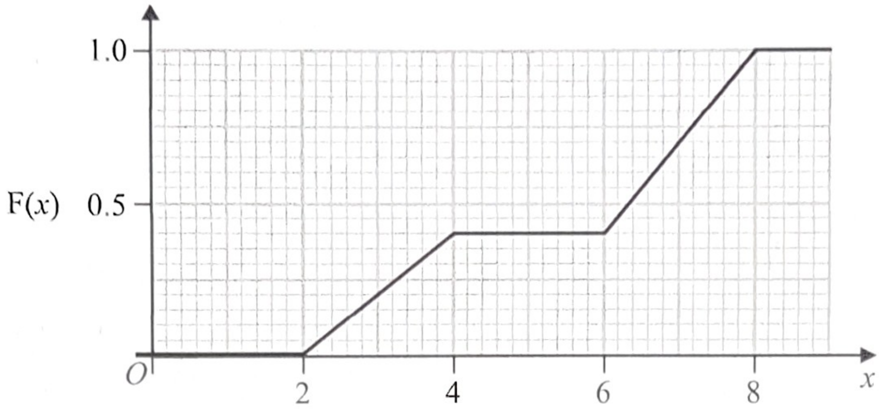


Figure 1

(a) Find $P(3 < X < 7)$

(2)

The probability density function of X is given by

$$f(x) = \begin{cases} a & 2 \leqslant x < 4 \\ b & 4 \leqslant x < 6 \\ c & 6 \leqslant x \leqslant 8 \\ 0 & \text{otherwise} \end{cases}$$

where a, b and c are constants.

(b) Find the value of a, the value of b and the value of c

(3)

(c) Find E(X)

(3)

(a) $F(7) - F(3)$

$0.7 - 0.2 = \underline{0.5}$

(b) $(2, 0)$ $(4, 0.4)$

$\frac{0.4 - 0}{4-2} = \frac{0.4}{2} = \underline{0.2}$

$(4, 0.4)$ $(6, 0.4)$

$\frac{0.4 - 0.4}{6-4} = \underline{0}$

$(6, 0.4)$ $(8, 1)$

$\frac{1 - 0.4}{8-6} = \frac{0.6}{2} = 0.3$

$\therefore$ $\underline{a = 0.2}$ $\underline{b = 0}$ $\underline{c = 0.3}$

(c) $E(X) = \int_2^4 0.2x(x)dx + \int_4^6 0x\,dx$

$+ \int_6^8 0.3(x)\,dx$

$= \left[\frac{0.2x^2}{2}\right]_2^4 + \left[\frac{0.3x^2}{2}\right]_6^8$

$= \frac{8}{5} - \frac{2}{5} + \frac{48}{5} - \frac{27}{5}$

$= 5.4$

Leave blank

4. At a shop, past figures show that 35% of customers pay by credit card. Following the shop's decision to no longer charge a fee for using a credit card, a random sample of 20 customers is taken and 11 are found to have paid by credit card.

Hadi believes that the proportion of customers paying by credit card is now greater than 35%

(a) Test Hadi's belief at the 5% level of significance. State your hypotheses clearly.
(5)

For a random sample of 20 customers,

(b) show that 11 lies less than 2 standard deviations above the mean number of customers paying by credit card.
You may assume that 35% is the true proportion of customers who pay by credit card.
(4)

(a) $X \sim B(20, 0.35)$

$\boxed{0.05}$

$H_0: p = 0.35$
$H_1: p > 0.35$.

$P(X \geq 11) = 1 - P(X \leq 10) = 1 - 0.9468$
$= 0.0532 \times$

$P(X \geq 12) = 1 - P(X \leq 11) = 1 - 0.9804$
$= 0.0196 \checkmark$

$CR = X \geq 12$.
It doesn't fall in CR.
$\therefore$ Hadi's Belief isn't supported

(b) $X \sim B(20, 0.35)$

mean $= np = 7$.
Variance $= np(1-p)$
$= 7(1-0.35)$
$= 4.55$

$S.D = \sqrt{4.55} = 2.13$

$7 + 2(2.13) = 11.26$.

$11.26 > 11$

Leave blank

5. The continuous random variable X is uniformly distributed over the interval $[a, b]$ where $0 < a < b$

Given that $\mathrm{P}(X < b - 2a) = \frac{1}{3}$

(a) (i) show that $\mathrm{E}(X) = \frac{5a}{2}$ **(3)**

(ii) find $\mathrm{P}(X > b - 4a)$ **(1)**

The continuous random variable Y is uniformly distributed over the interval $[3, c]$ where $c > 3$

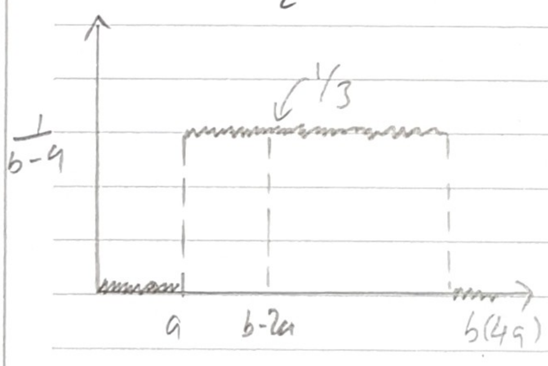
Given that $\mathrm{Var}(Y) = 3c - 9$, find

(b) (i) the value of c **(3)**

(ii) $\mathrm{P}(2Y - 7 < 20 - Y)$ **(3)**

(iii) $\mathrm{E}(Y^2)$ **(3)**

ai) $E(X) = \frac{a+b}{2}$



$b-2a-a \times \frac{1}{b-a} = \frac{1}{3}$

$\frac{b-3a}{b-a} = \frac{1}{3}$. -- (1).

$E(X) = \frac{a+4a}{2} = \frac{5a}{2}$ -- (2).

$3b - 9a = b - a$.

$2b = 8a$

$b = 4a$

ii) $p(X > b-4a)$.

$\cancel{\frac{4a-b-4a}{b-a}} = \cancel{\frac{-b}{b-a}}$.

$p(X > b-4a)$

$= p(X > 0) = 1$

Leave blank

Question 5 continued

bi) $Y \sim U[3, 6]$.

$\text{Var}(x) = \frac{(b-a)^2}{12}$

$= \frac{(c-3)^2}{12} = 3c - 9.$

$(c-3)^2 = 12(3c-9)$

$(c-3)^2 = 36c - 108.$

$c(c-3) = 3(c-3).$

$= c^2 - 6c + 9 = 36c - 108$

$c^2 - 42c + 117 = 0$

$c = 39$ or $c = 3$

n/a

$\underline{\underline{c = 39}}$

ii) $Y \sim U[3, 39]$.

$P(2Y - 7 < 20 - Y)$

$= P(Y < 9) = \frac{9-3}{39-3} = \underline{\underline{\frac{1}{6}}}$

iii) $E(Y^2)$

$\text{Var}(Y) = E(Y^2) - (E(Y))^2$

$3(39) - 9 = E(Y^2) - \left(\frac{39+3}{2}\right)^2$

$E(Y^2) = 549$

Leave blank

6. (i) (a) State the conditions under which the Poisson distribution may be used as an approximation to the binomial distribution.
(1)

A factory produces tyres for bicycles and 0.25% of the tyres produced are defective.

A company orders 3000 tyres from the factory.

(b) Find, using a Poisson approximation, the probability that there are more than 7 defective tyres in the company's order.
(3)

(ii) At the company 40% of employees are known to cycle to work. A random sample of 150 employees is taken. The random variable C represents the number of employees in the sample who cycle to work.

(a) Describe a suitable sampling frame that can be used to take this sample.
(1)

(b) Explain what you understand by the sampling distribution of C
(1)

Louis uses a normal approximation to calculate the probability that at most α employees in the sample cycle to work. He forgets to use a continuity correction and obtains the incorrect probability 0.0668

Find, showing all stages of your working,

(c) the value of α
(4)

(d) the correct probability.
(2)

i a) n is large and p is small.

$X \sim B(3000, 0.0025)$.

3000×0.0025

$= 7.5$

$X \sim Po(7.5)$

$P(X>7) = 1 - P(X \le 7)$

$1 - 0.5246$

$= 0.4754$.

ii a) A list of all the employees.

b) Probability distribution of the no. of employees that cycle to work.

(c) $C \sim B(150, 0.4)$

$P(C \le \alpha) = 0.0668$.

$np = 150 \times 0.4 = 60$

$np(1-p) = 60(1-0.4) = 36$.

Leave blank

Question 6 continued

$C \sim N(60, 36)$

$P\left(Z \le \frac{\alpha - 60}{\sqrt{36}}\right) = 0.0668$

$\frac{\alpha - 60}{6} = -1.5$

$\alpha - 60 = -9$

$\therefore \alpha = 51$

d) $P(C \le 51) = P(X < 51.5)$

$X \sim N(60, 36)$

$P\left(Z \le \frac{51.5 - 60}{\sqrt{36}}\right) = P(Z \le -1.42)$

$= 1 - 0.9222$

$= 0.0778$

$= 0.078$

Leave blank

7. The continuous random variable X has probability density function

$$f(x) = \begin{cases} c(x+3) & -3 \leqslant x < 0 \\ \frac{5}{36}(3-x) & 0 \leqslant x \leqslant 3 \\ 0 & \text{otherwise} \end{cases}$$

where c is a positive constant.

(a) Show that $c = \frac{1}{12}$ **(3)**

(b) (i) Sketch the probability density function.

(ii) Explain why the mode of $X = 0$ **(3)**

(c) Find the cumulative distribution function of X, for all values of x **(4)**

(d) Find, to 3 significant figures, the value of d such that $P(X > d \mid X > 0) = \frac{2}{5}$ **(4)**

$c \int_{-3}^{0} x+3 \cdot dx$

$= c\left[\frac{x^2}{2} + 3x\right]_{-3}^{0}$

$= c\left(0 - -\frac{9}{2}\right) = \frac{9}{2}c$

$\frac{5}{36}\int_{0}^{3} 3-x \cdot dx$

$\frac{5}{36}\left[3x - \frac{x^2}{2}\right]_{0}^{3}$

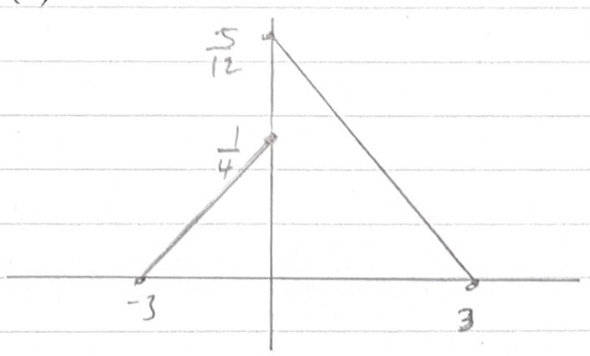
$\frac{5}{36}\left(\frac{9}{2} - 0\right) = \frac{5}{8}$

$\frac{9}{2}c + \frac{5}{8} = 1$

$\frac{9}{2}c = \frac{3}{8}$

$c = \frac{1}{12}$

bi)



ii) The highest point on the graph occurs when $x = 0$ so mode is 0

Leave blank

Question 7 continued

(c) Integrate.

$\int_{-3}^{x} \frac{1}{12}x + \frac{1}{4} \, dx$

$= \left[\frac{1}{24}x^2 + \frac{1}{4}x\right]_{-3}^{x}$

$= \frac{1}{24}x^2 + \frac{1}{4}x + \frac{3}{8}$

$F(0) = \frac{3}{8}$.

$\int_{0}^{x} \frac{5}{12} - \frac{5}{36}x \, dx$

$= \left[\frac{5}{12}x - \frac{5}{72}x^2\right]_0^x + F(0)$

$= \frac{5}{12}x - \frac{5}{72}x^2 + \frac{3}{8}$

$F(x) = \begin{cases} 0 & x < -3 \\ \frac{1}{24}x^2 + \frac{1}{4}x + \frac{3}{8} & -3 \le x < 0 \\ \frac{5}{12}x - \frac{5}{72}x^2 + \frac{3}{8} & 0 \le x \le 3 \\ 1 & x > 3. \end{cases}$

(d) $\frac{P(X > d \cap X > 0)}{P(X > 0)} = \frac{2}{5}$

$\frac{P(X > d)}{P(X > 0)}$

$P(X > 0) \to \frac{1}{2} \times 3 \times \frac{5}{12} = \frac{5}{8}$

$\frac{1 - F(d)}{5/8} = \frac{2}{5}$

$1 - F(d) = 2/5 \times 5/8$.

$1 - F(d) = 1/4$.

$1 - (5/12\, d - 5/72\, d^2 + 3/8) = 1/4$.

$1 - 5/12\, d + 5/72\, d^2 - 3/8 - 1/4 = 0$.

$5/72\, d^2 - 5/12\, d + 3/8 = 0$

$d = 1.10$ ✓

$d = 4.897$ ✗

$\underline{\underline{d = 1.10}}$