Write your name here

Surname	Other names

Pearson Edexcel
International Advanced Level

Centre Number	Candidate Number

Statistics S2

Advanced/Advanced Subsidiary

Wednesday 24 January 2018 – Afternoon
Time: 1 hour 30 minutes

Paper Reference
WST02/01

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 continuous random variable X has cumulative distribution function

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{16}(x-1)^2 & 1 \leqslant x \leqslant 5 \\ 1 & x > 5 \end{cases}$$

(a) Find $P(X > 4)$ **(2)**

(b) Find $P(X > 3 \mid 2 < X < 4)$ **(3)**

(c) Find the exact value of $E(X)$ **(4)**

(a) $P(X>4) = 1 - P(X \leq 4)$

$= 1 - F(4)$

$F(4) = \frac{1}{16}(4-1)^2 = \frac{9}{16}$.

$1 - \frac{9}{16} = \frac{7}{16}$.

(b) $P(X>3 \mid 2<X<4)$

$= \frac{P(3<X<4)}{P(2<X<4)}$

$P(3<X<4) = P(X<4) - P(X<3)$

$= F(4) - F(3)$

$= 9/16 - 1/4 = 5/16$.

$P(2<X<4) = P(X<4) - P(X<2)$

$= F(4) - F(2)$

$= 9/16 - 1/16 = 1/2$.

$= \frac{5/16}{1/2} = \frac{5}{8}$

(c) differentiate $\frac{1}{16}(x-1)^2$

$= \frac{1}{8}(x-1)$

$E(X) = \int_1^5 x \times \frac{1}{8}(x-1)\, dx$

$= \int_1^5 \frac{1}{8}x^2 - \frac{1}{8}x \cdot dx$

$\left[\frac{1}{24}x^3 - \frac{1}{16}x^2\right]_1^5$

$= \left(\frac{5^3}{24} - \frac{5^2}{16}\right) - \left(\frac{1}{24} - \frac{1}{16}\right)$

$= \frac{175}{48} - -\frac{1}{48} = \frac{11}{3}$.

Leave blank

2. A farmer sells boxes of eggs.

The eggs are sold in boxes of 6 eggs and boxes of 12 eggs in the ratio $n : 1$

A random sample of three boxes is taken.

The number of eggs in the first box is denoted by X_1
The number of eggs in the second box is denoted by X_2
The number of eggs in the third box is denoted by X_3
The random variable $T = X_1 + X_2 + X_3$

Given that $P(T = 18) = 0.729$

(a) show that $n = 9$

(2)

(b) find the sampling distribution of T

(4)

The random variable R is the range of X_1, X_2, X_3

(c) Using your answer to part (b), or otherwise, find the sampling distribution of R

(2)

a) 6eggs $\frac{n}{n+1}$ 12eggs $\frac{1}{n+1}$

Sample: 6 6 6 (18) 12 12 12 (36)
6 6 12 (24) 12 12 6 (30)
6 12 6 (24) 12 6 12 (30)
12 6 6 (24) 6 12 12 (30)

$P(T=18) = 0.729$

$\frac{n}{n+1} \times \frac{n}{n+1} \times \frac{n}{n+1} = 0.729$

$\frac{n^3}{(n+1)^3} = 0.729$

$n^3 = 0.729(n+1)^3$

$n^3 = (n^3+3n^2+3n+1) \cdot 0.729$

$0.271n^3 - 2.187n^2 - 2.187n - 0.729 = 0$

$\therefore \underline{n=9}$

(b) Samples ← 6eggs $\frac{9}{10}$ 12eggs $\frac{1}{10}$

T	18	24	30	36
$P(T=t)$	0.729	0.243	0.027	0.001

$P(T=24) = 3 \times \left(\frac{9}{10} \times \frac{9}{10} \times \frac{1}{10}\right) = 0.243$

$P(T=30) = 3\left(\frac{1}{10} \times \frac{1}{10} \times \frac{9}{10}\right) = 0.027$

$P(T=36) = \frac{1}{10} \times \frac{1}{10} \times \frac{1}{10} = \underline{\underline{0.001}}$

Question 2 continued

(c) $P(R=0) = P(T=18) + P(T=36)$

$= 0.73$

$P(R=6) = P(T=24) + P(T=30)$

$= 0.27$

Leave blank

3. Albert uses scales in his kitchen to weigh some fruit.

The random variable D represents, in grams, the weight of the fruit given by the scales **minus** the true weight of the fruit. The random variable D is uniformly distributed over the interval $[-2.5, 2.5]$

(a) Specify the probability density function of D
(2)

(b) Find the standard deviation of D
(2)

Albert weighs a banana on the scales.

(c) Write down the probability that the weight given by the scales equals the true weight of the banana.
(1)

(d) Find the probability that the weight given by the scales is within 1 gram of the banana's true weight.
(1)

Albert weighs 10 bananas on the scales, one at a time.

(e) Find the probability that the weight given by the scales is within 1 gram of the true weight for at least 6 of the bananas.
(3)

(a) $\frac{1}{2.5 - -2.5} = \frac{1}{5}$

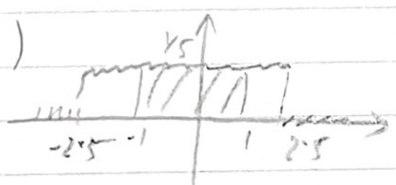
$f(x) \begin{cases} \frac{1}{5} & -2.5 \leq d \leq 2.5 \\ 0 & \text{otherwise} \end{cases}$

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

$s.d. = \sqrt{\frac{(b-a)^2}{12}} = \sqrt{\frac{(-2.5-2.5)^2}{12}} = 1.44 (3sf)$

(c) $p(D=0) = 0$

(d)



$1 - -1 = 2 \times \frac{1}{5} = 2/5$

(e) $X \sim B(10, 2/5)$

$P(X \geq 6) = 1 - P(X \leq 5)$

$1 - 0.8338 = 0.166 (3sf)$

Leave blank

4. A sweet shop produces different coloured sweets and sells them in bags.

The proportion of green sweets produced is p

Each bag is filled with a random sample of n sweets.

The mean number of green sweets in a bag is 4.2 and the variance is 3.57

(a) Find the value of n and the value of p

(4)

The proportion of red sweets produced by the shop is 0.35

(b) Find the probability that, in a random sample of 25 sweets, the number of red sweets exceeds the expected number of red sweets.

(3)

The shop claims that 10% of its customers buy more than two bags of sweets. A random sample of 40 customers is taken and 1 customer buys more than two bags of sweets.

(c) Test, at the 5% level of significance, whether or not there is evidence that the proportion of customers who buy more than two bags of sweets is less than the shop's claim. State your hypotheses clearly.

(4)

(a) $np = 4.2$ $\quad n = \frac{4.2}{p}$

$np(1-p) = 3.57$

$np - np^2 = 3.57$

$4.2 - np^2 = 3.57$

$4.2 - p \times \left(\frac{4.2}{p}\right) = 3.57$

$4.2 - 4.2p = 3.57$

$\underline{\underline{p = 0.15}}$

$np = 4.2$

$n(0.15) = 4.2$

$\underline{\underline{n = 28}}$

(b) $X \sim B(25, 0.35)$

$E(X) = 25 \times 0.35 = 8.75$
$\underline{\underline{= 9}}$

$P(X \geq 9) = 1 - P(X \leq 8)$

$1 - 0.4668 = \underline{\underline{0.5332}}$

$\underline{\underline{0.533}}$ (3sf)

(c) $X \sim B(40, 0.1)$ $\quad \boxed{0.05}$

$H_0: p = 0.1$
$H_1: p < 0.1$

$P(X \leq 1) = 0.0805$

$0.0805 > 0.05$

Don't reject H_0.

Shop's claim is not rejected.

Leave blank

5. A delivery company loses packages randomly at a mean rate of 10 per month.

The probability that the delivery company loses more than 12 packages in a randomly selected month is p

(a) Find the value of p **(1)**

The probability that the delivery company loses more than k packages in a randomly selected month is at least $2p$

(b) Find the largest possible value of k **(2)**

In a randomly selected month,

(c) find the probability that exactly 4 packages were lost in each half of the month, **(3)**

In a randomly selected two-month period, 21 packages were lost.

(d) Find the probability that at least 10 packages were lost in each of these two months. **(4)**

(e) Using a suitable approximation, find the probability that more than 27 packages are lost during a randomly selected 4-month period. **(5)**

(a) $X \sim Po(10)$

$P(X>12) = 1 - P(X \le 12)$

$1 - 0.7916 = \underline{0.208}$ (3sf)

(b) $P(x > k) \geq 2 \times 0.2084$

$P(X > k) \geq 0.4168$

$1 - P(X \le k) \geq 0.4168$

$P(X \le k) \le 0.5832$

$\underline{k = 10}$

(c) $10 \rightarrow 1$ month

$x \rightarrow 0.5$ months $\quad \underline{x = 5}$

$X \sim Po(5)$

$P(X=4)^2$

$= \left(\frac{e^{-5} \times 5^4}{4!}\right)^2 = \underline{0.0308}$ (3sf)

(d) $X \sim Po(10)$

$Y \sim Po(20)$

$P(X_1 \geq 10 \cap X_2 \geq 10 \mid Y = 21)$

$$= \frac{\left(\frac{e^{-10}10^{10}}{10!} \times \frac{e^{-10}10^{11}}{11!}\right) + \left(\frac{e^{-10}10^{11}}{11!} \times \frac{e^{-10}10^{10}}{10!}\right)}{\frac{e^{-20}20^{21}}{21!}}$$

Leave blank

Question 5 continued

$= \dfrac{0.0142 + 0.042}{0.0846}$

$= 0.336$

(e) $10 \to 1$ month.
$x \to 4$ months.

$x = 40$

$X \sim Po(40)$

$X \sim N(40, 40)$

$P(X > 27) = p(X > 27.5)$

$= P\left(Z > \dfrac{27.5 - 40}{\sqrt{40}}\right)$

$= P(Z > -1.98)$

$= P(Z < 1.98)$

$= 0.9761$

Leave blank

6. In a local council, 60% of households recycle at least half of their waste. A random sample of 80 households is taken.

The random variable X represents the number of households in the sample that recycle at least half of their waste.

(a) Using a suitable approximation, find the smallest number of households, n, such that $P(X \geqslant n) < 0.05$

(6)

The number of bags recycled per family per week was known to follow a Poisson distribution with mean 1.5

Following a recycling campaign, the council believes the mean number of bags recycled per family per week has increased. To test this belief, 6 families are selected at random and the total number of bags they recycle the following week is recorded.

The council wishes to test, at the 5% level of significance, whether or not there is evidence that the mean number of bags recycled per family per week has increased.

(b) Find the critical region for the total number of bags recycled by the 6 families.

(2)

(a)

since $X \sim B(80, 0.6)$

$\approx N(48, 19.2)$

$np = 48$

$np(1-p) = 19.2$

$P(X \geqslant n) < 0.05$

$P\left(Z > \frac{(n-0.5)-48}{\sqrt{19.2}}\right) < 0.05$

$\frac{(n-0.5)-48}{\sqrt{19.2}} > 1.6449$

$n > 55.7$

$n = 56$

(b) $X \sim Po(9)$

$H_0: \lambda = 9$

$H_1: \lambda > 9$

$P(X \geqslant 14) = 1 - P(X \leq 13) =$

$1 - 0.9261 = 0.0739$ ✗

$P(X \geqslant 15) = \cancel{P}\,1 - P(X \leq 14)$

$1 - 0.9585 = 0.0415$ ✓

$0.0415 < 0.05$

critical Region $= \cancel{P}(X \geqslant 15)$

Leave blank

7. The continuous random variable X has probability density function $f(x)$ given by

$$f(x) = \begin{cases} \frac{1}{16}x^2 & 1 \leqslant x < 3 \\ k(4-x) & 3 \leqslant x \leqslant 4 \\ 0 & \text{otherwise} \end{cases}$$

(a) Show that $k = \frac{11}{12}$ **(3)**

(b) Sketch $f(x)$ for $1 \leqslant x \leqslant 4$ **(2)**

(c) Write down the mode of X **(1)**

Given that $E(X) = \frac{25}{9}$

(d) use algebraic integration to find Var(X), giving your answer to 3 significant figures. **(4)**

The cumulative distribution function of X is given by

$$F(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{48}(x^3 + c) & 1 \leqslant x < 3 \\ \frac{11}{12}(4x - \frac{1}{2}x^2 + d) & 3 \leqslant x \leqslant 4 \\ 1 & x > 4 \end{cases}$$

(e) (i) Find the exact value of c

(ii) Find the exact value of d **(3)**

(f) Calculate, to 3 significant figures, the upper quartile of X **(2)**

(a) $\int_1^3 \frac{1}{16}x^2\,dx + k\int_3^4 4 - x\,dx = 1$

$\left[\frac{1}{48}x^3\right]_1^3 + k\left[4x - \frac{1}{2}x^2\right]_3^4 = 1$

$\left(\frac{9}{16} - \frac{1}{48}\right) + k(8 - 15/2) = 1$

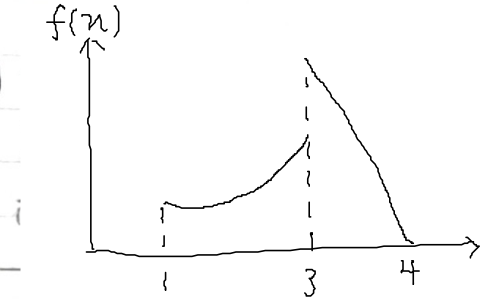
$\frac{13}{24} + \frac{1}{2}k = 1$

$\frac{1}{2}k = \frac{11}{24}$ $\quad k = \frac{11}{12}$ as req.

Leave blank

Question 7 continued

(b)



(c) Mode = 3.

(d) $Var(X) = E(X^2) - [E(X)]^2$

$E(X^2) = \int_1^3 \frac{1}{16}x^2(x^2)\,dx + \int_3^4 \frac{11}{12}(4-x)x^2\,dx$

$= \int_1^3 \frac{x^4}{16}\,dx + \int_3^4 \frac{11}{12}(4x^2 - x^3)\,dx$

$= \left[\frac{x^5}{80}\right]_1^3 + \frac{11}{12}\left[\frac{4}{3}x^3 - \frac{1}{4}x^4\right]_3^4$

$= \frac{3^5}{80} - \frac{1}{80} + \frac{11}{12}\left[\frac{4}{3}(4^3) - \frac{1}{4}(4)^4\right] - \frac{11}{12}\left(\frac{4}{3}(3^3) - \frac{1}{4}(3)^4\right]$

$= \frac{243}{80} - \frac{1}{80} + \frac{11}{12}\left(\frac{256}{3} - 64 - 36 + \frac{81}{4}\right)$

$= \frac{5863}{720}$

$\frac{5863}{720} - \left(\frac{25}{9}\right)^2 = \underline{0.427(3sf)}$

e i) $F(1) = 0$

when $x = 1$, $\frac{1}{48}(x^3 + c) = 0$

$\frac{1}{48}\left[(1)^3 + c\right] = 0$

$c = -1$

ii) $F(4) = 1$

when $x = 4$,

$\frac{11}{12}\left(4x - \frac{1}{2}x^2 + d\right) = 1$

$\frac{11}{12}\left[4(4) - \frac{1}{2}(4)^2 + d\right] = 1$

$\frac{11}{12}(16 - 8 + d) = 1$

$8 + d = \frac{12}{11}$

$d = \frac{-76}{11}$

(f) $F(q) = 0.75$

$\frac{11}{12}\left(4q - \frac{1}{2}q^2 - \frac{76}{11}\right) = 0.75$

$4q - \frac{1}{2}q^2 - \frac{76}{11} = \frac{9}{11}$

$\frac{1}{2}q^2 - 4q + \frac{85}{11} = 0$

$q = 4.74$ or 3.26
n/a

$\underline{3.26 = Q3}$

Q7

(Total for Question 7 is 5 marks)