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 31 October 2018

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

# Statistics S2

## Advanced/Advanced Subsidiary

**You must have: Scientific calculator**
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*





P 5 4 9 4 5 A 0 1 2 8

S1 Oct 2018.

Leave blank

1. Each day a restaurant opens between 11 am and 11 pm. During its opening hours, the restaurant receives calls for reservations at an average rate of 6 per hour.

(a) Find the probability that the restaurant receives exactly 1 call for a reservation between 6 pm and 7 pm.
**(2)**

The restaurant distributes leaflets to local residents to try and increase the number of calls for reservations. After distributing the leaflets, it records the number of calls for reservations it receives over a 90 minute period.

Given that it receives 14 calls for reservations during the 90 minute period,

(b) test, at the 5% level of significance, whether the rate of calls for reservations has increased from 6 per hour. State your hypotheses clearly.
**(5)**

(a) 6 → 1 hour.
$X \sim Po(6)$.

$P(X=1)$

$\frac{e^{-6} \times 6^1}{1!} = \underline{0.0149}$

(b) 6 → 60min
x → 90min

$X \sim Po(9)$

$H_0: \lambda = 9$ $\boxed{0.05}$
$H_1: \lambda > 9$

$P(X \geq 14) = 1 - P(X \leq 13) = 1 - 0.9261$
$= 0.0739$ ✗

$P(X \geq 15) = 1 - P(X \leq 14) = 1 - 0.9585$
$= 0.0415$ ✓

$P(X \geq 16) = 1 - P(X \leq 15) = 1 - 0.9780$
$= 0.022$ ✗

Critical region: $x \geq 15$.

14 does not fall in the critical region ∴ we accept the null hypothesis. There is ~~no~~ sufficient evidence that there is an increase in the number of calls for reservations.

Leave blank

**2.** At a cafe, customers ordering hot drinks order either tea or coffee.
Of all customers ordering hot drinks, 80% order tea and 20% order coffee.
Of those who order tea, 35% take sugar and of those who order coffee 60% take sugar.

(a) A random sample of 12 customers ordering hot drinks is selected.

Find the probability that fewer than 3 of these customers order coffee.
**(3)**

(b) (i) A randomly selected customer who orders a hot drink is chosen. Show that the probability that the customer takes sugar is 0.4
**(1)**

(ii) Write down the distribution for the number of customers who take sugar from a random sample of $n$ customers ordering hot drinks.
**(1)**

(c) A random sample of 10 customers ordering hot drinks is selected.

(i) Find the probability that exactly 4 of these 10 customers take sugar.
**(1)**

(ii) Given that at least 3 of these 10 customers take sugar, find the probability that no more than 6 of these 10 customers take sugar.
**(3)**

(d) In a random sample of 150 customers ordering hot drinks, find, using a suitable approximation, the probability that at least half of them take sugar.
**(4)**

(a) $X \sim B(12, 0.2)$
$P(X<3) = P(X \le 2) = \underline{0.558}$

bi) (tea × sugar) + (coffee × sugar)
$(0.8 \times 0.35) + (0.2 \times 0.6)$

$= 0.28 + 0.12 = \underline{0.4}$

ii) $X \sim B(n, 0.4)$

c) i) $X \sim B(10, 0.4)$

$P(X=4) = \binom{10}{4}(0.4)^4(0.6)^6$

$= \underline{0.251}$

(ii) $P(X \le 6 \mid P(X \ge 3))$

$= \dfrac{P(3 \le X \le 6)}{P(X \ge 3)}$

$P(3 \le X \le 6)$
$= P(X \le 6) - P(X \le 2)$
$= 0.9452 - 0.1673 \quad = \underline{0.7779}$

$P(X \ge 3)$
$= 1 - P(X \le 2) = 1 - 0.1673$
$= \underline{0.8327}$

$\dfrac{0.7779}{0.8327} = 0.934.$

Leave blank

**Question ~~1~~ 2 continued**

(a) $X \sim B(150, 0.4)$

$np = 60.$

$np(1-p) = 60(0.6) = 36.$

$X \sim N(60, 36)$

$p(X \geq 75) = p(X > 74.5)$

$1 - P\left(Z \leq \frac{74.5 - 60}{\sqrt{36}}\right)$

$= 1 - P(Z \leq 2.42)$

$= 1 - 0.9922 = \underline{\underline{0.0078}}$

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

**3.** The function f($x$) is defined as

$$f(x) = \begin{cases} \dfrac{1}{9}(x+5)(3-x) & 1 \leqslant x \leqslant 4 \\ 0 & \text{otherwise} \end{cases}$$

Albert believes that f($x$) is a valid probability density function.

(a) Sketch f($x$) and comment on Albert's belief. **(3)**

The continuous random variable $Y$ has probability density function given by

$$g(y) = \begin{cases} ky(12-y^2) & 1 \leqslant y \leqslant 3 \\ 0 & \text{otherwise} \end{cases}$$

where $k$ is a positive constant.

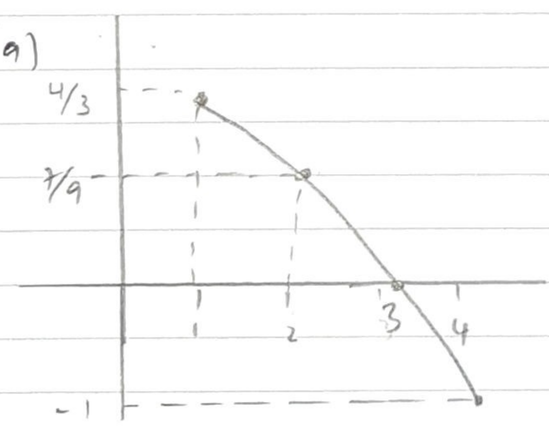
(b) Use calculus to find the mode of $Y$ **(3)**

(c) Use algebraic integration to find the value of $k$ **(3)**

(d) Find the median of $Y$ giving your answer to 3 significant figures. **(3)**

(e) Describe the skewness of the distribution of $Y$ giving a reason for your answer. **(2)**

(a)




This is not a pdf because $f(x) < 0$ when $x > 3$.

b) $g(y) = \begin{cases} ky(12-y^2) & 1 \leq y \leq 3 \\ 0 & \text{otherwise} \end{cases}$

differentiate and equate to zero

$k[12y - y^3] = g(y)$

$g'(y) = k[12 - 3y^2] = 0.$

$12 = 3y^2$

$y^2 = 4$

$\underline{y = 2}$

Leave blank

**Question 3 continued**

(c) $k\int_1^3 12y - y^3\,dy = 1$

$k\left[6y^2 - \frac{y^4}{4}\right]_1^3 = 1$

$k\left[135/4 - 23/4\right] = 1$

$28k = 1$

$k = \frac{1}{28}$.

(d) Integrate.

$\frac{1}{28}\int_1^y 12y - y^3\,dy$

$= \frac{1}{28}\left[\frac{6y^2}{1} - \frac{y^4}{4}\right]_1^y$

$= \frac{1}{28}\left[6y^2 - \frac{y^4}{4} - \frac{23}{4}\right]$.

$= \frac{3y^2}{14} - \frac{1}{112}y^4 - \frac{23}{112} = F(m)$

$= 0.5$

$\frac{3}{14}m^2 - \frac{1}{112}m^4 - \frac{23}{112} = 0.5$.

$\frac{1}{112}m^4 - \frac{3}{14}m^2 + \frac{79}{112} = 0$.

$m^4 - 24m^2 + 79 = 0$

$m = 4.479$ ✗

or

$m = 1.984$ ✓

$m = \underline{\underline{1.984}}$

(e) mean median mode

median $\approx$ mode ∴ no skew.



Turn over

Leave blank

**4.** A bag contains a large number of marbles, each of which is blue or red.

A random sample of 3 marbles is taken from the bag.

The random variable $D$ represents the number of blue marbles taken minus the number of red marbles taken.

Given that 20% of the marbles in the bag are blue,

(a) show that $P(D = -1) = 0.384$ **(2)**

(b) find the sampling distribution of $D$ **(3)**

(c) write down the mode of $D$ **(1)**

Takashi claims that the true proportion of blue marbles is greater than 20% and tests his claim by selecting a random sample of 12 marbles from the bag.

(d) Find the critical region for this test at the 10% level of significance. **(2)**

(e) State the actual significance level of this test. **(1)**

(a)

| $x$ | Red | Blue |
|---|---|---|
| $P(X=x)$ | 0·8 | 0·2 |

$-1$ = 1 blue marble − 2 red marbles.

0·2, 0·8, 0·8
0·8, 0·2, 0·8
0·8, 0·8, 0·2.

$3(0{\cdot}2 \times 0{\cdot}8 \times 0{\cdot}8) = \frac{48}{125}$

$= 0{\cdot}384$

(b)

| $d$ | $-3$ | $-1$ | $1$ | $3$ |
|---|---|---|---|---|
| $P(D=d)$ | $\frac{64}{125}$ | $\frac{48}{125}$ | $\frac{12}{125}$ | $\frac{1}{125}$ |

From:

B B B (3) $\frac{1}{125}$
B B R (1)
B R B (1)
R B B (1) $\frac{12}{125}$

R R R (−3) $\frac{64}{125}$
R R B (−1)
R B R (−1)
B R R (−1) $\frac{48}{125}$

Leave blank

**Question 4 continued**

(c) Mode = −3.

d) $X \sim B(12, 0.2)$

$H_0 : p = 0.2$
$H_1 : p > 0.2$.

$P(X \geq 4) = 1 - P(X \leq 3)$

$= 1 - 0.7946 = 0.2054$ ✗
$(> 0.10)$

$P(X \geq 5) = 1 - P(X \leq 4)$

$1 - 0.9274 = 0.0726$ ✓
$(< 0.10)$

∴ critical region is

$X \geq 15$

(e) $0.0726 \times 100$

$= 7.26\%$

Leave blank

**5.** The random variable $X$ has cumulative distribution function given by

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{100}(ax^3 + bx^2 + 15x) & 0 \leqslant x \leqslant 5 \\ 1 & x > 5 \end{cases}$$

Given that $E(X^2) = 6.25$

(a) show that $6a + b = 0$

**(5)**

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

**(4)**

(c) find $P(3 \leqslant X \leqslant 7)$

**(2)**

(a) differentiate $\frac{1}{100}(ax^3+bx^2+15x)$

$= \frac{1}{100}(3ax^2 + 2bx + 15)$.

$f(x) = \begin{cases} \frac{1}{100}(3ax^2+2bx+15) & 0 \leq x \leq 5 \\ 0 & \text{otherwise} \end{cases}$

$E(X^2) = \int x^2 f(x)\, dx$

$\frac{1}{100}\int_0^5 3ax^4 + 2bx^3 + 15x^2 dx$

$\frac{1}{100}\left[\frac{3a}{5}x^5 + \frac{2bx^4}{4} + 5x^3\right]_0^5$

$\frac{1}{100}\left[1875a + \frac{625b}{2} + 625\right]$.

$\frac{75a}{4} + \frac{25b}{8} + 6.25 = 6.25$.

$\frac{75a}{4} + \frac{25b}{8} = 0 \quad \div 25/8$

$\div \frac{25}{8}$

$6a + b = 0$ as req.

(b) $F(5) = 1$

$\frac{1}{100}(125a + 25b + 75) = 1$.

$1.25a + 0.25b + 0.75 = 1$

$1.25a + 0.25b = 0.25$.

$(6a + b = 0)\ 0.25$

$1.5a + 0.25b = 0$.

$-\ 1.25a + 0.25b = 0.25$

$0.25a = -0.25$

$a = -1$

$b = 0 - 6(-1)$

$b = 6$

(c) $F(5) - F(3)$

$= 1 - F(3) = 1 - \frac{1}{100}(27a + 9b + 45)$

$= 1 - 18/25$

$= 0.28$

Leave blank

**6.** One side of a square is measured to the nearest centimetre and this measurement is multiplied by 4 to estimate the perimeter of the square. The random variable, $W$ cm, represents the estimated perimeter of the square minus the true perimeter of the square.

$W$ is uniformly distributed over the interval $[a, b]$

(a) Explain why $a = -2$ and $b = 2$

**(1)**

The standard deviation of $W$ is $\sigma$

(b) (i) Find $\sigma$

(ii) Find the probability that the estimated perimeter of the square is within $\sigma$ of the true perimeter of the square.

**(4)**

One side of each of 100 squares are now measured. Using a suitable approximation,

(c) find the probability that $W$ is greater than 1.9 for at least 5 of these squares.

**(4)**

(a) $a = 4 \times -0.5$
$= \underline{-2}$

$b = 4 \times 0.5 = \underline{2}$

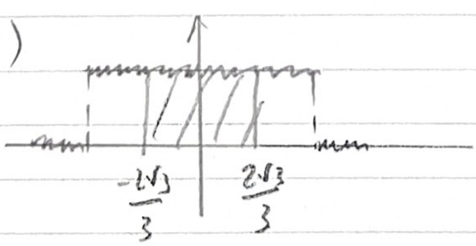
(b) $W \sim U[-2, 2]$.

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

$= \frac{(2--2)^2}{12} = \frac{4}{3}$.

$\sqrt{4/3} = \underline{\underline{1.15}}$

ii)




$P\left(-\frac{2\sqrt{3}}{3} < W < \frac{2\sqrt{3}}{3}\right)$.

$\frac{\frac{2\sqrt{3}}{3} - -\frac{2\sqrt{3}}{3}}{2-(-2)} = \frac{4\sqrt{3}}{3} \times \frac{1}{4}$

$= \frac{\sqrt{3}}{3} = \underline{\underline{0.577}}$

(c) $P(W > 1.9)$

$\frac{2-1.9}{2-(-2)} = 0.1 \times 1/4 = \underline{0.025}$

$X \sim B(100, 0.025)$

$100 \times 0.025 = 2.5$.

$X \sim Po(2.5)$

$P(X \geq 5) = 1 - P(X \leq 4)$
$1 - 0.8912$
$= \underline{\underline{0.109}}$

Leave blank

**7.** Members of a conservation group record the number of sightings of a rare animal. The number of sightings follows a Poisson distribution with a rate of 1 every 2 months.

(a) Find the smallest value of $n$ such that the probability that there are at least $n$ sightings in 2 months is less than 0.05

**(2)**

(b) Find the smallest number of months, $m$, such that the probability of no sightings in $m$ months is less than 0.05

**(2)**

(c) Find the probability that there is at least 1 sighting per month in each of 3 consecutive months.

**(3)**

(d) Find the probability that the number of sightings in an 8 month period is equal to the expected number of sightings for that period.

**(2)**

(e) Given that there were 4 sightings in a 4 month period, find the probability that there were more sightings in the last 2 months than in the first 2 months.

**(3)**

(a) $X \sim Po(1)$

$P(X \geq n) < 0.05$

$1 - P(X \leq n-1) < 0.05$

$P(X \leq n-1) > 0.95$

$n - 1 = 3 \quad \therefore \underline{n = 4}$

(b) $1 \to 2$ months
$x \to m$ months

$2x = m$

$x = \frac{m}{2}$

$X \sim Po(\frac{1}{2}m)$

$P(X = 0) < 0.05$

$\frac{e^{-\frac{1}{2}m} \times (\frac{1}{2}m)^0}{0!}$

$e^{-\frac{1}{2}m} < 0.05$

$\ln e^{-\frac{1}{2}m} < \ln 0.05$

$-\frac{1}{2}m < \ln 0.05$

$m > 5.99$

$\underline{m = 6}$

(c) $1 \to 2$ months
$x \to 1$ month

$\underline{x = 0.5}$

$X \sim Po(0.5) \quad P(X \geq 1) = 1 - P(X \leq 0)$

$1 - 0.6065 = 0.3935$

Leave blank

**Question 7 continued**

$(0.3935)^3 = \underline{\underline{0.0609}}$

(d) $1 \rightarrow 2$ months
$x \rightarrow 8$ months

$\underline{x = 4}$

$X \sim Po(4)$

$P(X=4) = \frac{e^{-4} \times 4^4}{4!} = \underline{\underline{0.195}}$

e) $A \sim Po(1)$
$B \sim Po(2)$

P(sightings in the last 2 months > sightings in first 2 months / 4 sightings in 4 months)

$= \frac{P(A=1) \times P(A=3) + P(A=0) \times P(A=4)}{P(B=4)}$

$= \frac{e^{-1} \times \frac{e^{-1}}{3!} + e^{-1} \times \frac{e^{-1}}{4!}}{\frac{e^{-2} \times 2^4}{4!}}$

$= \frac{(0.368 \times 0.061) + (0.368 \times 0.0158)}{0.092}$

$= \frac{5}{16}$

$= 0.311$



Turn over