Please check the examination details below before entering your candidate information

Candidate surname	Other names

Pearson Edexcel
International
Advanced Level

Centre Number | Candidate Number

Sample Assessment Materials for first teaching September 2018

(Time: 1 hour 30 minutes) | Paper Reference **WST01/01**

Mathematics

International Advanced Subsidiary/Advanced Level Statistics S1

You must have:
Mathematical Formulae and Statistical Tables, 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.
- 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 ▶

S59767A

1/1/1/1/1/1/

S 5 9 7 6 7 A 0 1 2 8

Pearson

S1 specimen 2018

Leave blank

Answer ALL questions. Write your answers in the spaces provided.

1. The percentage oil content, p, and the weight, w milligrams, of each of 10 randomly selected sunflower seeds were recorded. These data are summarised below.

$$\sum w^2 = 41252 \qquad \sum wp = 27557.8 \qquad \sum w = 640 \qquad \sum p = 431 \qquad S_{pp} = 2.72$$

(a) Find the value of S_{ww} and the value of S_{wp} **(3)**

(b) Calculate the product moment correlation coefficient between p and w **(2)**

(c) Give an interpretation of your product moment correlation coefficient. **(1)**

The equation of the regression line of p on w is given in the form $p = a + bw$

(d) Find the equation of the regression line of p on w **(4)**

(e) Hence estimate the percentage oil content of a sunflower seed which weighs 60 milligrams. **(2)**

(a) $S_{ww} = \sum w^2 - \frac{(\sum w)^2}{10}$

$= 41252 - \frac{(640)^2}{10}$

$= \underline{\underline{292}}$

$S_{wp} = \sum wp - \frac{\sum w \sum p}{10}$

$= 27557.8 - \frac{640 \times 431}{10}$

$= \underline{\underline{-26.2}}$ (3sf)

(b) $r = \frac{S_{wp}}{\sqrt{S_{ww} \times S_{pp}}} = \frac{-26.2}{\sqrt{292 \times 2.72}}$

$= \underline{-0.93}$ (2dp)

(c) negative correlation as weight of seeds increases the percentage of oil content decreases.

(d) $b = \frac{S_{wp}}{S_{ww}} = \frac{-26.2}{292} = \underline{\underline{-0.0897}}$ (3sf)

$a = \bar{p} - b\bar{w} = 43.1 - b \times 64 = 48.8$ (3sf)

$p = 48.8 - 0.0897w$

(e) @ $w = 60$mg.

$p = 43.46$

$= \underline{43.5}$ (3sf)

Leave blank

2. The time taken to complete a puzzle, in minutes, is recorded for each person in a club. The times are summarised in a grouped frequency distribution and represented by a histogram.

One of the class intervals has a frequency of 20 and is shown by a bar of width 1.5 cm and height 12 cm on the histogram. The total area under the histogram is 94.5 cm^2

Find the number of people in the club.

(3)

Area	freq.
18	20
94·5	x

$$x = \frac{94 \cdot 5 \times 20}{18}$$

$$\underline{\underline{x = 105}}$$

Leave blank

3. The discrete random variable X has probability distribution

$$P(X = x) = \frac{1}{5} \qquad x = 1, 2, 3, 4, 5$$

(a) Write down the name given to this distribution. **(1)**

Find

(b) $P(X = 4)$ **(1)**

(c) $F(3)$ **(1)**

(d) $P(3X - 3 > X + 4)$ **(2)**

(e) Write down $E(X)$ **(1)**

(f) Find $E(X^2)$ **(2)**

(g) Hence find $\text{Var}(X)$ **(2)**

Given that $E(aX - 3) = 11.4$

(h) find $\text{Var}(aX - 3)$ **(4)**

X	1	2	3	4	5
P(X=x)	1/5	1/5	1/5	1/5	1/5

(a) discrete uniform distribution

(b) $= 1/5$

(c) $3/5$

(d) $P(2x > 7) \Rightarrow P(x > 7/2)$
$= 2/5$

(e) $\left(1\times\frac{1}{5}\right) + \left(2\times\frac{1}{5}\right) + \left(3\left(\frac{1}{5}\right)\right) + \left(4\times\frac{1}{5}\right) + \left(5\times\frac{1}{5}\right)$
$= 3$

(f) $E(x^2) = (1\times 1/5) + (4\times 1/5) + (9\times 1/5) + (16\times 1/5) + (25\times 1/5)$
$= 11$

(g) $\text{Var}(x) = E(x^2) - ((E x))^2 = 11 - 3^2$
$= 2$

(h) $a(3) - 3 = 11.4$
$a = 4.8$
$\therefore \text{Var}(4.8x - 3) = 46.1$

Leave blank

4. A researcher recorded the time, t minutes, spent using a mobile phone during a particular afternoon, for each child in a club.

The researcher coded the data using $v = \dfrac{t-5}{10}$ and the results are summarised in the table below.

Coded Time (v)	Frequency (f)	Coded Time Midpoint (m)	CF
$0 \leqslant v < 5$	20	2.5	20
$5 \leqslant v < 10$	24	a	44
$10 \leqslant v < 15$	16	12.5	60
$15 \leqslant v < 20$	14	17.5	74
$20 \leqslant v < 30$	6	b	80

(You may use $\sum fm = 825$ and $\sum fm^2 = 12\,012.5$)

(a) Write down the value of a and the value of b. **(1)**

(b) Calculate an estimate of the mean of v. **(1)**

(c) Calculate an estimate of the standard deviation of v. **(2)**

(d) Use linear interpolation to estimate the median of v. **(2)**

(e) Hence describe the skewness of the distribution. Give a reason for your answer. **(2)**

(f) Calculate estimates of the mean and the standard deviation of the time spent using a mobile phone during the afternoon by the children in this club. **(4)**

(a) $a = 7.5 \quad b = 25.$

(b) $\bar{v} = \dfrac{\sum fm}{80} = \dfrac{825}{80} = \underline{\underline{10.3}}(3\!/\!4)$

(c) $\sigma = \sqrt{\dfrac{\sum fm^2}{80} - \left(\dfrac{\sum fm}{80}\right)^2}$

$= \sqrt{\dfrac{12012.5}{80} - \left(\dfrac{825}{80}\right)^2}$

$= \underline{\underline{6.62}}$

(d) At 40^{th} frequency.

20   40   44.

5   m   10

$\dfrac{44-40}{44-20} = \dfrac{10-\cancel{10}m}{10-5}$

$\underline{\underline{m = 9.17}}$

Leave blank

Question 4 continued

(e) $\bar{v} = 10.3$

$m = 9.17$

mean > median ∴ +ve skew.

f) $\bar{r} = \dfrac{\bar{t} - 5}{10}$

$\bar{t} = 108.125$

$= 108$ (3sf)

$\sigma_{coded} = \dfrac{\sigma_{old}}{10}$

$\sigma_{old} = 10 \times 6.62$

$= \underline{\underline{66.2}}$ (3sf)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Leave blank

5. A biased tetrahedral die has faces numbered 0, 1, 2 and 3. The die is rolled and the number face down on the die, X, is recorded. The probability distribution of X is

x	0	1	2	3
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$

If $X = 3$ then the final score is 3
If $X \neq 3$ then the die is rolled again and the final score is the sum of the two numbers.

The random variable T is the final score.

(a) Find $P(T = 2)$ **(2)**

(b) Find $P(T = 3)$ **(3)**

(c) Given that the die is rolled twice, find the probability that the final score is 3 **(3)**

(a) $P(T=2) = p(x=0 \text{ and } x=2) + p(x=2 \text{ and } x=0) + p(x=1 \text{ and } x=1)$

$\left(\frac{1}{6}\times\frac{1}{6}\right) + \left(\frac{1}{6}\times\frac{1}{6}\right) + \left(\frac{1}{6}\times\frac{1}{6}\right)$

$= \frac{1}{12}$

(c) $P(\text{die rolled twice}) = 1 - \frac{1}{2} = \frac{1}{2}$

$p(\text{final score} = 3 \mid \text{rolled twice})$

$= \frac{5/36}{1/2} = 5/18$

(b) $P(T=3) = p(x=3) + p(x=0 \text{ and } x=3) + p(x=1 \text{ and } x=2) + p(x=2 \text{ and } x=1)$

$\frac{1}{2} + \left(\frac{1}{6}\times\frac{1}{2}\right) + \left(\frac{1}{6}\times\frac{1}{6}\right) + \left(\frac{1}{6}\times\frac{1}{6}\right)$

$= \frac{23}{36}$.

Leave blank

6. Three events A, B and C are such that

$$P(A) = \frac{2}{5} \qquad P(C) = \frac{1}{2} \qquad P(A \cup B) = \frac{5}{8}$$

Given that A and C are mutually exclusive find

(a) $P(A \cup C)$ **(1)**

Given that A and B are independent

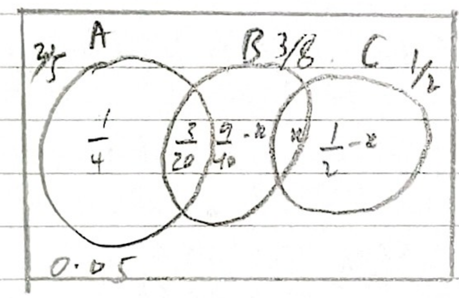
(b) show that $P(B) = \frac{3}{8}$ $\rightarrow p(A \cap B) = p(A) \times p(B)$ **(4)**

(c) Find $P(A \mid B)$ **(1)**

Given that $P(C' \cap B') = 0.3$

(d) draw a Venn diagram to represent the events A, B and C **(5)**

(d)



(a) $p(A \cup C) = p(A) + p(C) - p(A \cap C)$

$= \frac{2}{5} + \frac{1}{2} - 0 \quad = \frac{9}{10}$

(b) $p(A \cup B) = p(A) + p(B) - p(A \cap B)$

$= p(A) + p(B) - p(A) \times p(B)$

$5/8 = 2/5 + p(B) - 2/5\, p(B)$

$9/40 = 3/5\, p(B)$

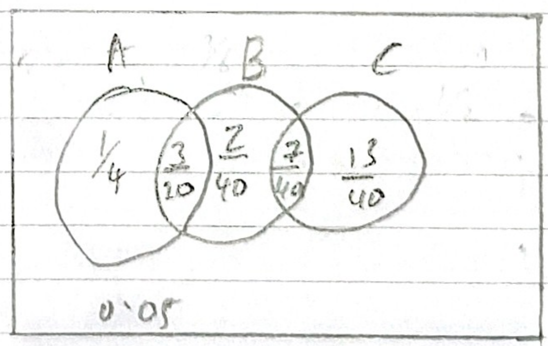
$p(B) = 3/8$

(c) $p(A|B) = \frac{p(A \cap B)}{p(B)} = \frac{2}{5}$ $\rightarrow$ same as $p(A)$

(d) $\frac{1}{4} + \frac{3}{20} + \frac{9}{40} - x + x + \frac{1}{2} - x + 0.05 = 1$

$-x = -7/40$

$x = 7/40$



and $(A \cap C)$ is 0 as they are mutually exclusive

Leave blank

7. A machine fills bottles with water. The volume of water delivered by the machine to a bottle is X ml where $X \sim N(\mu, \sigma^2)$

One of these bottles of water is selected at random.

Given that $\mu = 503$ and $\sigma = 1.6$

(a) find

(i) $P(X > 505)$
(ii) $P(501 < X < 505)$

(5)

(b) Find w such that $P(1006 - w < X < w) = 0.9426$

(3)

Following adjustments to the machine, the volume of water delivered by the machine to a bottle is such that $\mu = 503$ and $\sigma = q$

Given that $P(X < r) = 0.01$ and $P(X > r + 6) = 0.05$

(c) find the value of r and the value of q

(7)

a) $X \sim N(503, 1.6^2)$

i) $P(X > 505) = P\left(Z > \frac{505-503}{1.6}\right)$

$P(Z > 1.25)$

$1 - P(Z < 1.25) = 1 - 0.8944$

$= \underline{0.1056}$

ii) $P(501 < X < 505)$

$= P\left(\frac{501-503}{1.6} < Z < \frac{505-503}{1.6}\right)$

$P(-1.25 < Z < 1.25)$

$2(P(Z < 1.25) - 0.5)$

$2(0.8944 - 0.5)$

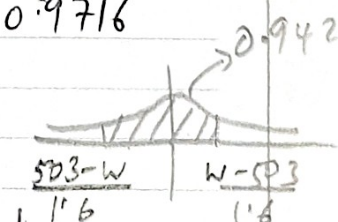
$= \underline{0.7888}$

(b) $P\left(\frac{1006 - W - 503}{1.6} < Z < \frac{W - 503}{1.6}\right)$

$P\left(\frac{503-W}{1.6} < Z < \frac{W-503}{1.6}\right)$

$P\left(Z < \frac{W-503}{1.6}\right) - 0.5 = \frac{0.9426}{2}$

$P\left(Z < \frac{W-503}{1.6}\right) = 0.9716$



$\frac{W-503}{1.6} = 1.9$

$W = 506.04$

$W = 506$

NB

$-\left(\frac{W-503}{1.6}\right) = \frac{503-W}{1.6}$

Leave blank

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Question 7 continued

(c) $X \sim N(503, q^2)$

$P\left(Z < \frac{r-503}{q}\right) = 0.01$

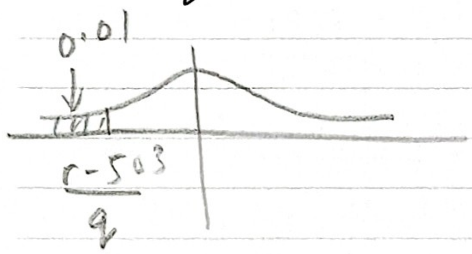
and

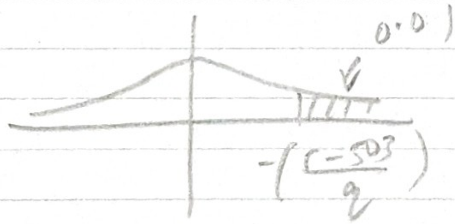
$P\left(Z > \frac{r+6-503}{q}\right) = 0.05$

$P\left(Z < \frac{r-503}{q}\right) = 0.01$

and

$P\left(Z > \frac{r-497}{q}\right) = 0.05.$



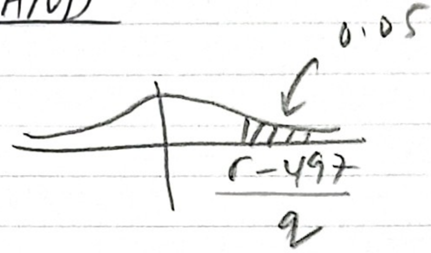


$\frac{r-503}{q} = -2.3263$

$r - 503 = -2.3263q.$

$r + 2.3263q = 503. \quad \text{--} (1)$

AND



$\frac{r-497}{q} = 1.6449.$

$r - 497 = 1.6449q.$

$r - 1.6449q = 497. \quad \text{--} (2)$

By solving simultaneous equations for (1) and (2)

$r = \underline{\underline{499}}$ (3sf)

$q = \underline{\underline{1.51}}$ (3sf)