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

# Wednesday 13 January 2021

Afternoon (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 (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* ▶

P65795A

©2021 Pearson Education Ltd.

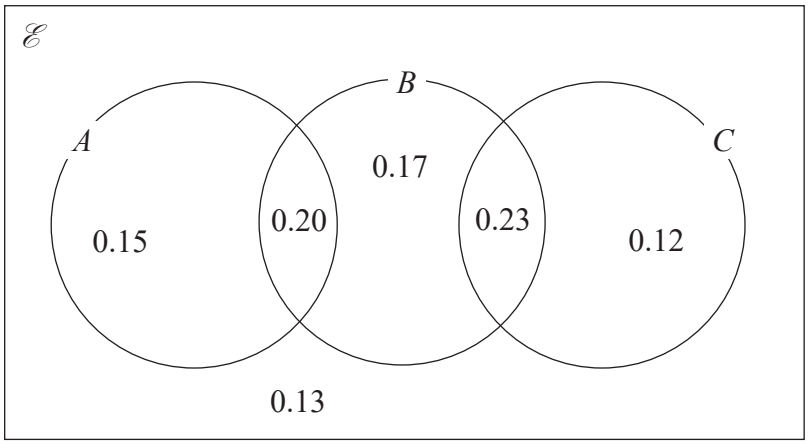
1/1/1/1/

P 6 5 7 9 5 A 0 1 2 0

Pearson

Leave blank

**1.** The Venn diagram shows the events $A$, $B$ and $C$ and their associated probabilities.




Find

(a) $P(B')$

**(1)**

(b) $P(A \cup C)$

**(2)**

(c) $P(A \mid B')$

**(2)**

(a) $P(B') = 0.15 + 0.13 + 0.12 = 0.4$

(b) $P(A \cup C) = 0.15 + 0.20 + 0.23 + 0.12$

$= 0.7$

(c) $P(A \mid B') = \frac{P(A \cap B')}{P(B')} = \frac{0.15}{0.4}$

$= \frac{3}{8}$

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**

**Q1**

**(Total 5 marks)**

Leave blank

**2.** The stem and leaf diagram below shows the ages (in years) of the residents in a care home.

| Age | | | | | | | | | | | | Key: 4\|3 is an age of 43 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 3 | | | | | | | | | | | (1) |
| 5 | 4 | | | | | | | | | | | (1) |
| 6 | 2 | 3 | 5 | 6 | 8 | 8 | 8 | 9 | 9 | | | (9) |
| 7 | 1 | 1 | 3 | 4 | 4 | 6 | 6 | 6 | 8 | 8 | 9 | (11) |
| 8 | 0 | 0 | 2 | 7 | 8 | 8 | 9 | | | | | (7) |
| 9 | 3 | 7 | | | | | | | | | | (2) |

(a) Find the median age of the residents.

**(1)**

(b) Find the interquartile range (IQR) of the ages of the residents.

**(2)**

An outlier is defined as a value that is either

more than 1.5 × (IQR) below the lower quartile or

more than 1.5 × (IQR) above the upper quartile.

(c) Determine any outliers in these data. Show clearly any calculations that you use.

**(3)**

(d) On the grid on page 5, draw a box plot to summarise these data.

**(3)**

(a) $n = 31$, so the median is the 16th value

Median = 74

(b) $Q_1$ is the 8th value = 68, $Q_3$ is the 24th value = 80

$IQR = 80 - 68 = 12$

(c) Lower limit: $Q_1 - 1.5 \times IQR = 68 - 1.5 \times 12 = 50$

Upper limit: $Q_3 + 1.5 \times IQR = 80 + 1.5 \times 12 = 98$

Outliers are values $< 50$ or $> 98$

$43 < 50$, and the largest value $97 < 98$

so there is just one outlier, at 43

(d) Box plot drawn on the grid below:

lower whisker to 54 (smallest value that is not an outlier),

upper whisker to 97, outlier at 43 marked with a cross.

DO NOT WRITE IN THIS AREA

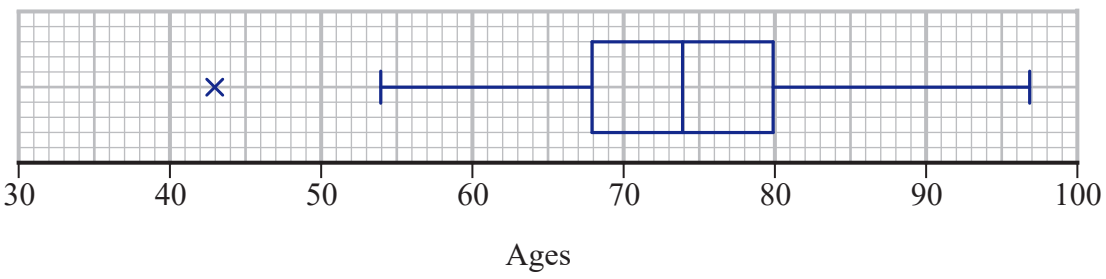
DO NOT WRITE IN THIS AREA

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**




**Q2**

**(Total 9 marks)**

Leave blank

**3.** The weights of packages that arrive at a factory are normally distributed with a mean of 18 kg and a standard deviation of 5.4 kg

(a) Find the probability that a randomly selected package weighs less than 10 kg

**(3)**

The heaviest 15% of packages are moved around the factory by Jemima using a forklift truck.

(b) Find the weight, in kg, of the lightest of these packages that Jemima will move.

**(3)**

One of the packages **not** moved by Jemima is selected at random.

(c) Find the probability that it weighs more than 18 kg

**(4)**

A delivery of 4 packages is made to the factory.
The weights of the packages are independent.

(d) Find the probability that exactly 2 of them will be moved by Jemima.

**(3)**

(a) $W \sim N(18, 5.4^2)$

$P(W < 10) = P\left(Z < \frac{10 - 18}{5.4}\right) = P(Z < -1.4814...)$

$= 1 - 0.9306 = 0.0694$

(b) $P(W > j) = 0.15 \Rightarrow P(W < j) = 0.85$

$\frac{j - 18}{5.4} = 1.0364$

$j = 18 + 5.4 \times 1.0364 = 23.59656 = 23.6$ kg

(c) $P(W > 18 \mid W < 23.6) = \frac{P(18 < W < 23.6)}{P(W < 23.6)}$

$= \frac{0.85 - 0.5}{0.85} = \frac{0.35}{0.85}$

$= \frac{7}{17}$ $(= 0.41176...)$

(d) P(moved by Jemima) = 0.15

Number moved $\sim B(4, 0.15)$

$P(\text{exactly } 2) = {}^4C_2 \times 0.15^2 \times 0.85^2$

$= 6 \times 0.0225 \times 0.7225$

$= 0.0975375 = 0.0975$

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**

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 13 marks)**

**Q3**

Leave blank

**4.** A spinner can land on the numbers 10, 12, 14 and 16 only and the probability of the spinner landing on each number is the same.
The random variable $X$ represents the number that the spinner lands on when it is spun once.

(a) State the name of the probability distribution of $X$.

**(1)**

(b) (i) Write down the value of E($X$)

**(1)**

(ii) Find Var($X$)

**(2)**

A second spinner can land on the numbers 1, 2, 3, 4 and 5 only.
The random variable $Y$ represents the number that this spinner lands on when it is spun once. The probability distribution of $Y$ is given in the table below

| $y$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| P($Y = y$) | $\frac{4}{30}$ | $\frac{9}{30}$ | $\frac{6}{30}$ | $\frac{5}{30}$ | $\frac{6}{30}$ |

(c) Find (i) E($Y$)

**(2)**

(ii) Var($Y$)

**(3)**

The random variable $W = aX + b$, where $a$ and $b$ are constants and $a > 0$
Given that E($W$) = E($Y$) and Var($W$) = Var($Y$)

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

**(5)**

Each of the two spinners is spun once.

(e) Find P($W = Y$)

**(2)**

(a) Discrete uniform distribution

(b)(i) $E(X) = 13$

(ii) $E(X^2) = \frac{10^2 + 12^2 + 14^2 + 16^2}{4} = \frac{696}{4} = 174$

$Var(X) = 174 - 13^2 = 174 - 169 = 5$

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**

(c)(i) $E(Y) = \dfrac{1\times4 + 2\times9 + 3\times6 + 4\times5 + 5\times6}{30} = \dfrac{90}{30} = 3$

(ii) $E(Y^2) = \dfrac{1\times4 + 4\times9 + 9\times6 + 16\times5 + 25\times6}{30} = \dfrac{324}{30} = 10.8$

$\text{Var}(Y) = 10.8 - 3^2 = 1.8$

(d) $E(W) = aE(X) + b = 13a + b$

$E(W) = E(Y) \Rightarrow 13a + b = 3$

$\text{Var}(W) = a^2\,\text{Var}(X) = 5a^2$

$\text{Var}(W) = \text{Var}(Y) \Rightarrow 5a^2 = 1.8 \Rightarrow a^2 = 0.36$

$a > 0$, so $a = 0.6$

$b = 3 - 13 \times 0.6 = 3 - 7.8 = -4.8$

$a = 0.6, \quad b = -4.8$

(e) $W = 0.6X - 4.8$, so the possible values of W are

$0.6(10) - 4.8 = 1.2, \quad 0.6(12) - 4.8 = 2.4,$

$0.6(14) - 4.8 = 3.6, \quad 0.6(16) - 4.8 = 4.8$

None of these are integers, but Y only takes integer values,

so W can never equal Y.

$P(W = Y) = 0$

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 16 marks)**

**Q4**

Leave blank

**5.** A company director wants to introduce a performance-related pay structure for her managers. A random sample of 15 managers is taken and the annual salary, $y$ in £1000, was recorded for each manager. The director then calculated a performance score, $x$, for each of these managers.
The results are shown on the scatter diagram in Figure 1 on the next page.

(a) Describe the correlation between performance score and annual salary.

**(1)**

The results are also summarised in the following statistics.

$$\sum x = 465 \quad \sum y = 562 \quad S_{xx} = 2492 \quad \sum y^2 = 23140 \quad \sum xy = 19428$$

(b) (i) Show that $S_{xy} = 2006$

**(1)**

(ii) Find $S_{yy}$

**(2)**

(c) Find the product moment correlation coefficient between performance score and annual salary.

**(2)**

The director believes that there is a linear relationship between performance score and annual salary.

(d) State, giving a reason, whether or not these data are consistent with the director's belief.

**(1)**

(e) Calculate the equation of the regression line of $y$ on $x$, in the form $y = a + bx$
Give the value of $a$ and the value of $b$ to 3 significant figures.

**(4)**

(f) Give an interpretation of the value of $b$.

**(1)**

(g) Plot your regression line on the scatter diagram in Figure 1

**(2)**

The director hears that one of the managers in the sample seems to be underperforming.

(h) On the scatter diagram, circle the point that best identifies this manager.

**(1)**

The director decides to use this regression line for the new performance related pay structure.

(i) Estimate, to 3 significant figures, the new salary of a manager with a performance score of 30

**(2)**

(a) Positive correlation

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**

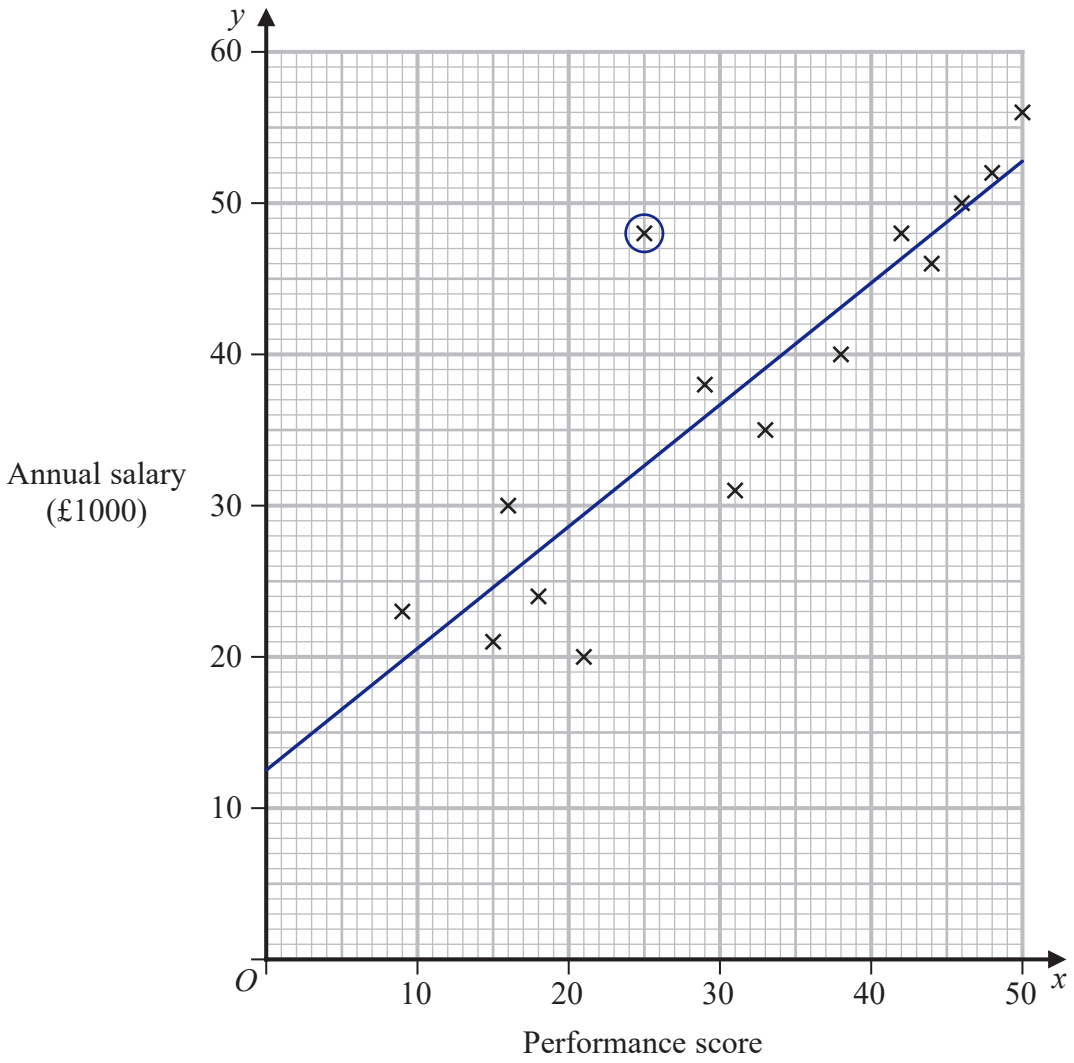



**Figure 1**

(b)(i) $S_{xy} = \Sigma xy - \frac{\Sigma x \Sigma y}{n} = 19428 - \frac{465 \times 562}{15}$

$= 19428 - 17422 = 2006$

(ii) $S_{yy} = \Sigma y^2 - \frac{(\Sigma y)^2}{n} = 23140 - \frac{562^2}{15}$

$= 23140 - 21056.2666... = 2083.7333... = 2080$ (3 s.f.)

(c) $r = \frac{S_{xy}}{\sqrt{(S_{xx} S_{yy})}} = \frac{2006}{\sqrt{(2492 \times 2083.7333...)}}$

$= 0.88031... = 0.880$

**Turn over for a spare copy of the scatter diagram if you need to redraw your line.**

Leave blank

**Question 5 continued**

(d) Yes, consistent: $r = 0.880$ is close to 1 (strong positive correlation), so the points lie close to a straight line.

(e) $b = \frac{S_{xy}}{S_{xx}} = \frac{2006}{2492} = 0.804975...$

$\bar{x} = \frac{465}{15} = 31, \qquad \bar{y} = \frac{562}{15} = 37.4666...$

$a = \bar{y} - b\bar{x} = 37.4666... - 0.804975... \times 31 = 12.5124...$

$y = 12.5 + 0.805x$

(f) Each extra performance point increases annual salary by about £805 (about £800).

(g) Line $y = 12.5 + 0.805x$ plotted on Figure 1: through (0, 12.5) and (50, 52.8).

(h) Circled point (25, 48) on Figure 1: a high salary for a low performance score.

(i) $y = 12.5124... + 0.804975... \times 30$

$= 36.6616...$

New salary ≈ £36 700

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

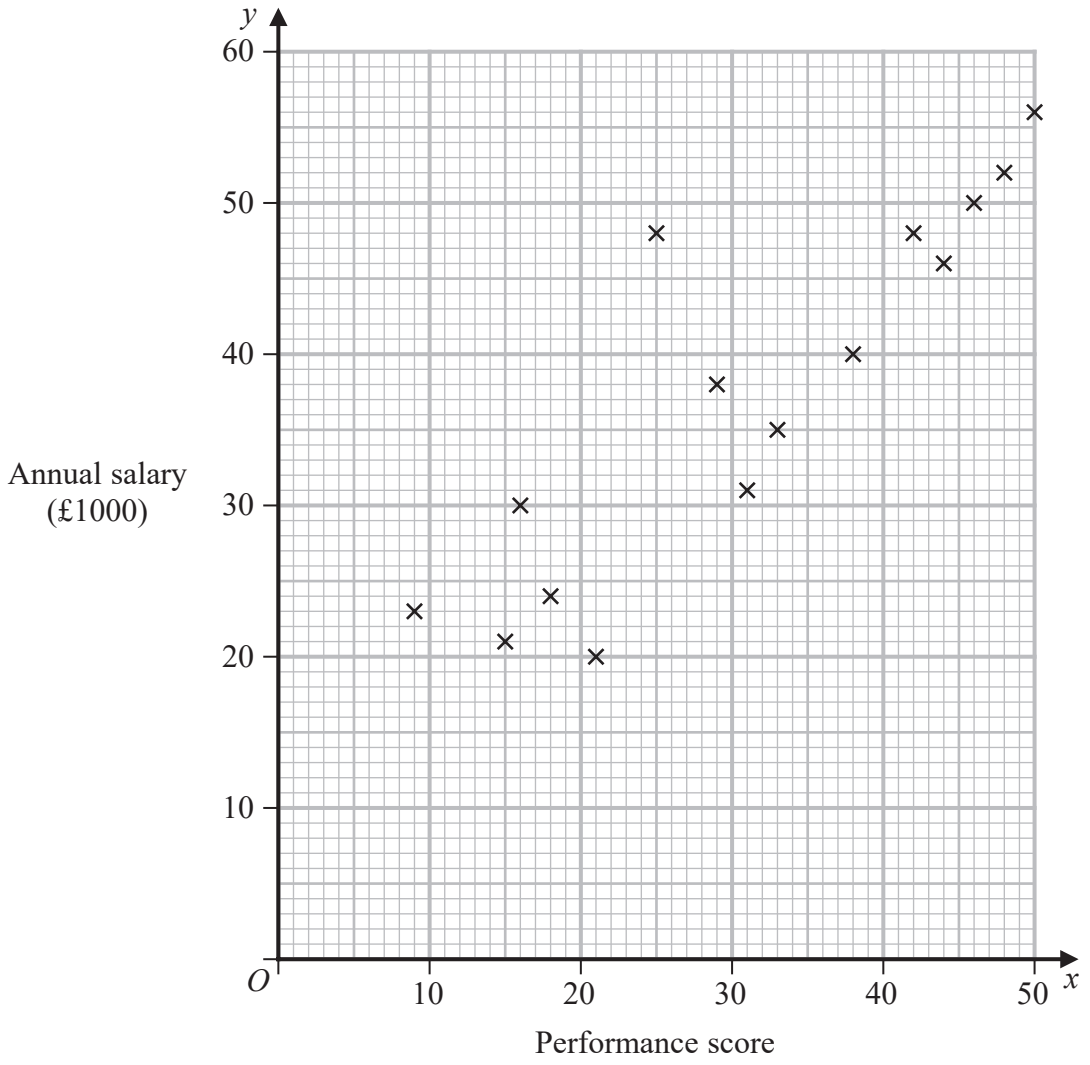
DO NOT WRITE IN THIS AREA

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

Leave blank

**Question 5 continued**

**Only use this scatter diagram if you need to redraw your line.**




**Q5**

**(Total 17 marks)**

Leave blank

**6.** A disc of radius 1 cm is rolled onto a horizontal grid of rectangles so that the disc is equally likely to land anywhere on the grid. Each rectangle is 5 cm long and 3 cm wide. There are no gaps between the rectangles and the grid is sufficiently large so that no discs roll off the grid.

If the disc lands inside a rectangle without covering any part of the edges of the rectangle then a prize is won.

By considering the possible positions for the centre of the disc,

(a) show that the probability of winning a prize on any particular roll is $\frac{1}{5}$

**(3)**

A group of 15 students each roll the disc onto the grid twenty times and record the number of times, $x$, that each student wins a prize. Their results are summarised as follows

$$\sum x = 61 \qquad \sum x^2 = 295$$

(b) Find the standard deviation of the number of prizes won per student.

**(2)**

A second group of 12 students each roll the disc onto the grid twenty times and the mean number of prizes won per student is 3.5 with a standard deviation of 2

(c) Find the mean and standard deviation of the number of prizes won per student for the whole group of 27 students.

**(7)**

The 27 students also recorded the number of times that the disc covered a corner of a rectangle and estimated the probability to be 0.2216 (to 4 decimal places).

(d) Explain how this probability could be used to find an estimate for the value of $\pi$ and state the value of your estimate.

**(3)**

(a) No edge is covered when the centre of the disc is at least 1 cm from every side of the rectangle.

So the centre must lie in a smaller rectangle of length $5 - 2 = 3$ cm and width $3 - 2 = 1$ cm.

$P(\text{prize}) = \frac{3 \times 1}{5 \times 3} = \frac{3}{15} = \frac{1}{5}$ as required

(b) $\sigma_x^2 = \frac{295}{15} - \left(\frac{61}{15}\right)^2 = 3.12888...$

$\sigma_x = \sqrt{3.12888...} = 1.76886... = 1.77$

(c) Second group (y): $\Sigma y = 12 \times 3.5 = 42$

Whole group: $\Sigma = 42 + 61 = 103$

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**

$\sigma_y^2 = \frac{\Sigma y^2}{12} - 3.5^2 \quad \Rightarrow \quad 2^2 = \frac{\Sigma y^2}{12} - 3.5^2$

$\Sigma y^2 = 12(2^2 + 3.5^2) = 12 \times 16.25 = 195$

Whole group: $\Sigma$ squares $= 195 + 295 = 490$

New mean $= \frac{103}{27} = 3.81481\ldots = 3.81$

New variance $= \frac{490}{27} - 3.81481\ldots^2 = 3.59533\ldots$

New standard deviation $= \sqrt{3.59533\ldots} = 1.89613\ldots = 1.90$

mean $= 3.81$, standard deviation $= 1.90$

(d) The disc covers a corner when its centre is within 1 cm of a vertex. Each rectangle has 4 corners, each giving a quarter circle of radius 1, so the area is a full circle $= \pi$.

P(covers a corner) $= \frac{\pi \times 1^2}{5 \times 3} = \frac{\pi}{15}$

Set $\frac{\pi}{15} = 0.2216 \quad \Rightarrow \quad \pi \approx 15 \times 0.2216$

Estimate of $\pi = 3.324$

Leave blank

**Question 6 continued**

**Q6**

**(Total 15 marks)**

**TOTAL FOR PAPER IS 75 MARKS**

**END**

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA