

MINISTRY OF EDUCATION, SINGAPORE
in collaboration with
CAMBRIDGE ASSESSMENT INTERNATIONAL EDUCATION
General Certificate of Education Advanced Level
Higher 2



CANDIDATE
NAME

CENTRE
NUMBER

S				
---	--	--	--	--

INDEX
NUMBER

--	--	--	--

MATHEMATICS

9758/02

Paper 2

October/November 2020

3 hours

Candidates answer on the Question Paper.

Additional Materials: List of Formulae (MF26)

READ THESE INSTRUCTIONS FIRST

Write your Centre number, index number and name on the work you hand in.

Write in dark blue or black pen.

You may use an HB pencil for any diagrams or graphs.

Do not use staples, paper clips, glue or correction fluid.

DO NOT WRITE ON ANY BARCODES.

Answer **all** the questions.

Write your answers in the spaces provided in the Question Paper.

Give non-exact numerical answers correct to 3 significant figures, or 1 decimal place in the case of angles in degrees, unless a different level of accuracy is specified in the question.

You are expected to use an approved graphing calculator.

Unsupported answers from a graphing calculator are allowed unless a question specifically states otherwise.

Where unsupported answers from a graphing calculator are not allowed in a question, you are required to present the mathematical steps using mathematical notations and not calculator commands.

You are reminded of the need for clear presentation in your answers.

The number of marks is given in brackets [] at the end of each question or part question.

This document consists of **21** printed pages and **3** blank pages.



Singapore Examinations and Assessment Board



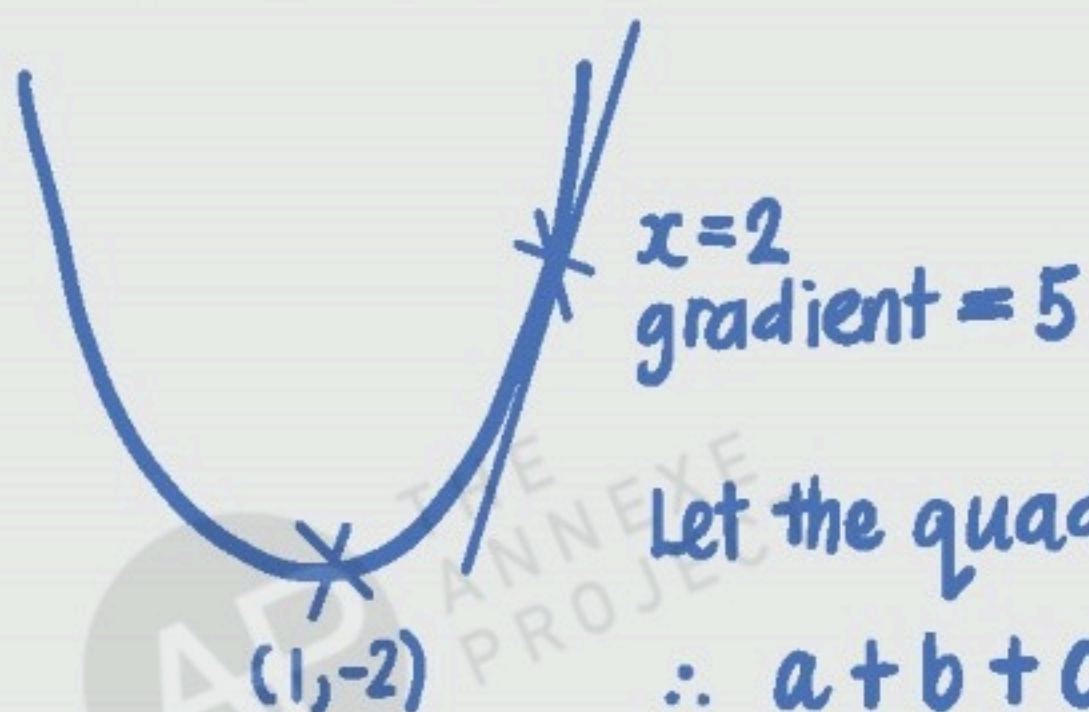
Cambridge Assessment
International Education

JC20 11_9758_02/3R
© UCLES & MOE 2020

[Turn over]

Section A: Pure Mathematics [40 marks]

- 1 A quadratic curve has its minimum point at $(1, -2)$ and has gradient 5 at the point where $x = 2$. Find the equation of the curve. [5]



Let the quadratic equation be $y = ax^2 + bx + c$.

$$\therefore a + b + c = -2 \quad \text{--- (1)}$$

$$\text{When } x=2, \frac{dy}{dx} = 5 : \frac{dy}{dx} = 2ax + b$$

$$5 = 2a(2) + b$$

$$4a + b = 5 \quad \text{--- (2)}$$

$$\text{When } x=1, \frac{dy}{dx} = 0 : 2a + b = 0 \quad \text{--- (3)}$$

$$\text{By GC: } a = \frac{5}{2}, b = -5, c = \frac{1}{2}$$

$$\therefore \underline{y = \frac{5}{2}x^2 - 5x + \frac{1}{2}}$$



- 2 (a) A sequence is such that $u_1 = p$, where p is a constant, and $u_{n+1} = 2u_n - 5$, for $n > 0$.

(i) Describe how the sequence behaves when

(A) $p = 7$,

$$u_1 = 7$$

$$u_2 = 2(7) - 5 = 9$$

$$u_3 = 2(9) - 5 = 13$$

$$u_4 = 2(13) - 5 = 21$$

The sequence increases to infinity.

[1]

(B) $p = 5$.

$$u_1 = 5$$

$$u_2 = 2(5) - 5 = 5$$

$$u_3 = 2(5) - 5 = 5$$

$$u_4 = 2(5) - 5 = 5$$

The sequence is a constant sequence of 5 for each term.

[1]

- (ii) Find the value of p for which $u_5 = 101$.

$$u_1 = p$$

$$u_2 = 2p - 5$$

$$u_3 = 2(2p - 5) - 5 = 4p - 15$$

$$u_4 = 2(4p - 15) - 5 = 8p - 35$$

$$u_5 = 2(8p - 35) - 5 = 16p - 75$$

$$\text{Let } 16p - 75 = 101$$

$$16p = 176$$

$$\therefore p = 11$$

[2]

- (b) Another sequence is defined by $v_1 = a$, $v_2 = b$, where a and b are constants, and

$$v_{n+2} = v_n + 2v_{n+1} - 7, \quad \text{for } n > 0.$$

For this sequence, $v_4 = 2v_3$.

(i) Find the value of b .

[3]

$$\text{when } n=1: \quad v_3 = v_1 + 2v_2 - 7 \\ = a + 2b - 7$$

$$\text{when } n=2: \quad v_4 = v_2 + 2v_3 - 7 \\ = b + 2(a + 2b - 7) - 7 \\ = 2a + 5b - 21$$

$$\text{Since } v_4 = 2v_3$$

$$\text{then } 2a + 5b - 21 = 2(a + 2b - 7)$$

$$2a + 5b - 21 = 2a + 4b - 14$$

$$\underline{b = 7}$$





(ii) Find an expression in terms of a for v_5 .

[1]

$$\begin{aligned}
 \text{When } n=3: \quad v_5 &= v_3 + 2v_4 - 7 \\
 &= (a + 14 - 7) + 2(2a + 35 - 21) - 7 \\
 &= \underline{5a + 28}
 \end{aligned}$$

(c) The sum of the first n terms of a series is $n^3 - 11n^2 + 4n$, where n is a positive integer.

(i) Find an expression for the n th term of this series, giving your answer in its simplest form.

[2]

$$\begin{aligned}
 T_n &= S_n - S_{n-1} \\
 &= (n^3 - 11n^2 + 4n) - [(n-1)^3 - 11(n-1)^2 + 4(n-1)] \\
 &= \cancel{n^3} - 11\cancel{n^2} + 4n - (\cancel{n^3} - 3n^2 + 3n - 1) + 11(\cancel{n^2} - 2n + 1) - 4n + 4 \\
 &= \underline{3n^2 - 25n + 16}
 \end{aligned}$$

(ii) The sum of the first m terms of this series, where $m > 3$, is equal to the sum of the first three terms of this series. Find the value of m .

[2]

$$\begin{aligned}
 \text{Given } S_m &= S_3 \\
 m^3 - 11m^2 + 4m &= 27 - 99 + 12 \\
 m^3 - 11m^2 + 4m + 60 &= 0
 \end{aligned}$$

$$\text{By GC: } \underline{m=10}, 3 \text{ (rej.) or } -2 \text{ (rej.)}$$





3 The curve C is defined by the parametric equations

$$x = 3t^2 + 2, \quad y = 6t - 1 \quad \text{where } t \geq \frac{1}{6}.$$

The line N is the normal to C at the point $(14, 11)$.

- (i) Find the cartesian equation of N . Give your answer in the form $ax + by = c$, where a , b and c are integers to be determined. [5]

$$\frac{dx}{dt} = 6t, \quad \frac{dy}{dt} = 6 \quad \therefore \frac{dy}{dx} = \frac{1}{t}$$

$$\text{When } y = 11, \quad 6t - 1 = 11 \quad \therefore t = 2$$

When $t = 2$; gradient of tangent = $\frac{1}{2}$
hence, gradient of normal = -2

Equation of normal:

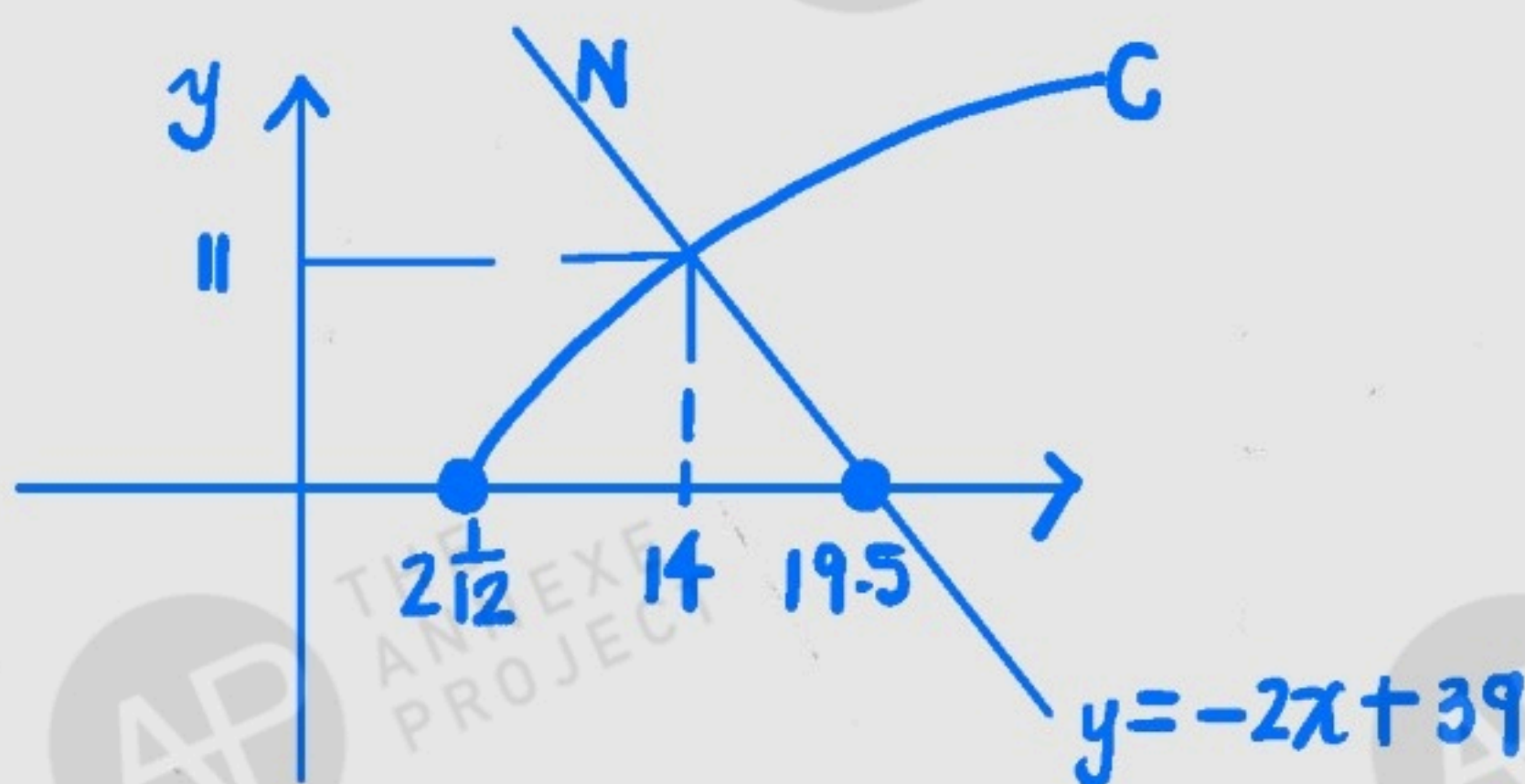
$$y - 11 = -2(x - 14)$$

$$y = -2x + 39$$

$$\underline{2x + y = 39}, \text{ where } a = 2, b = 1, c = 39.$$



(ii) Find the area enclosed by C , N and the x -axis.



$$\begin{aligned} \text{Area} &= \int_{2\frac{1}{2}}^{14} y \, dx + \frac{1}{2}(19.5-14)(11) \\ &= \int_{t=\frac{1}{6}}^{t=2} (6t-1)6t \, dt + 30.25 = 114.28 \\ &= \underline{114 \text{ sq. units}} \end{aligned}$$

(iii) The curve C and the line N are both transformed by a 2-way stretch, scale factor 2 in the x -direction and scale factor 3 in the y -direction, to form the curve D and the line M .

(a) Find the area enclosed by D , M and the x -axis.

[1]

$$\begin{aligned} 114.28 \times 6 &= 685\frac{2}{3} \\ &= \underline{686 \text{ sq. units}} \end{aligned}$$

(b) Find the cartesian equation of D .

[2]

Cartesian eqn. of C :

$$x = 3\left(\frac{y+1}{6}\right)^2 + 2$$

$$12x = (y+1)^2 + 24$$

↓ scale factor 2 stretch in the x -direction

$$12\left(\frac{x}{2}\right) = (y+1)^2 + 24$$

$$6x = (y+1)^2 + 24$$

↓ scale factor 3 stretch in the y -direction

$$6x = \left(\frac{y}{3}+1\right)^2 + 24$$

$$6x = \frac{(y+3)^2}{9} + 24$$

$$\underline{x = \frac{(y+3)^2}{54} + 4}$$

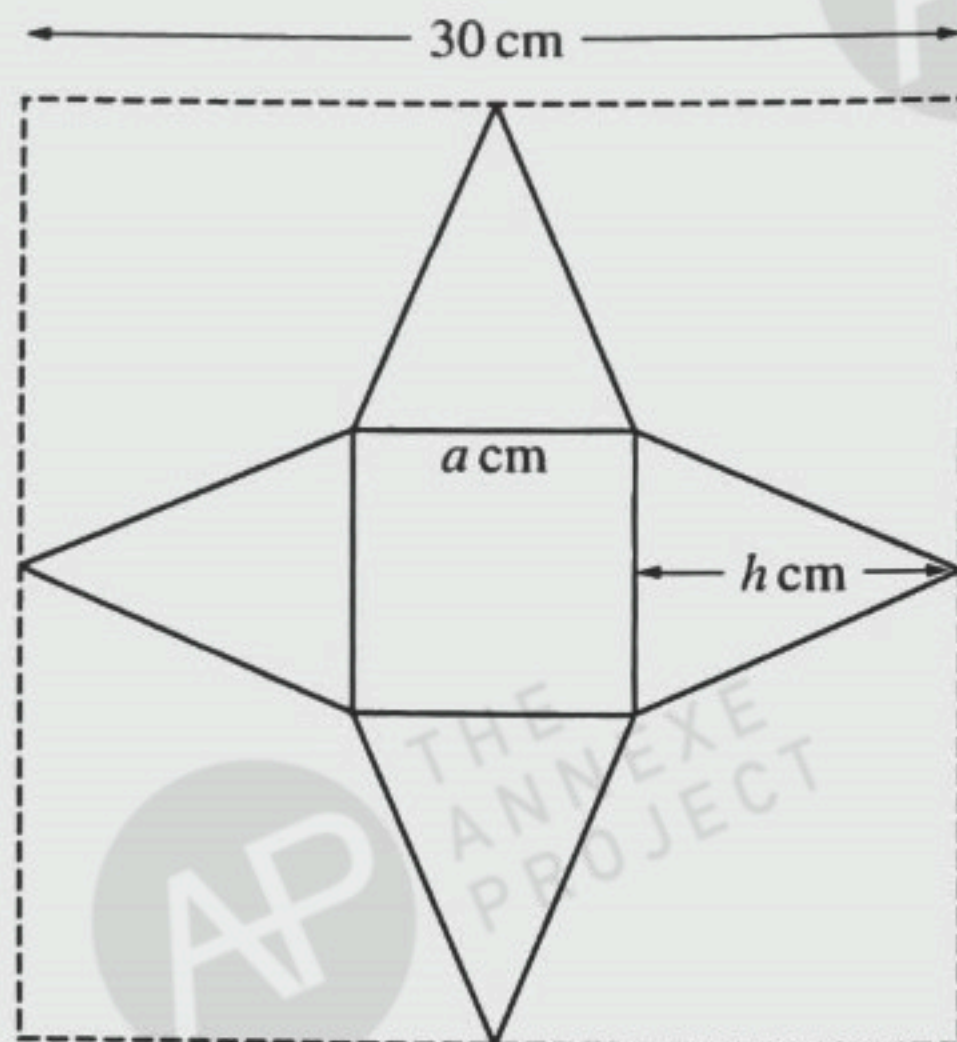


Fig. 1

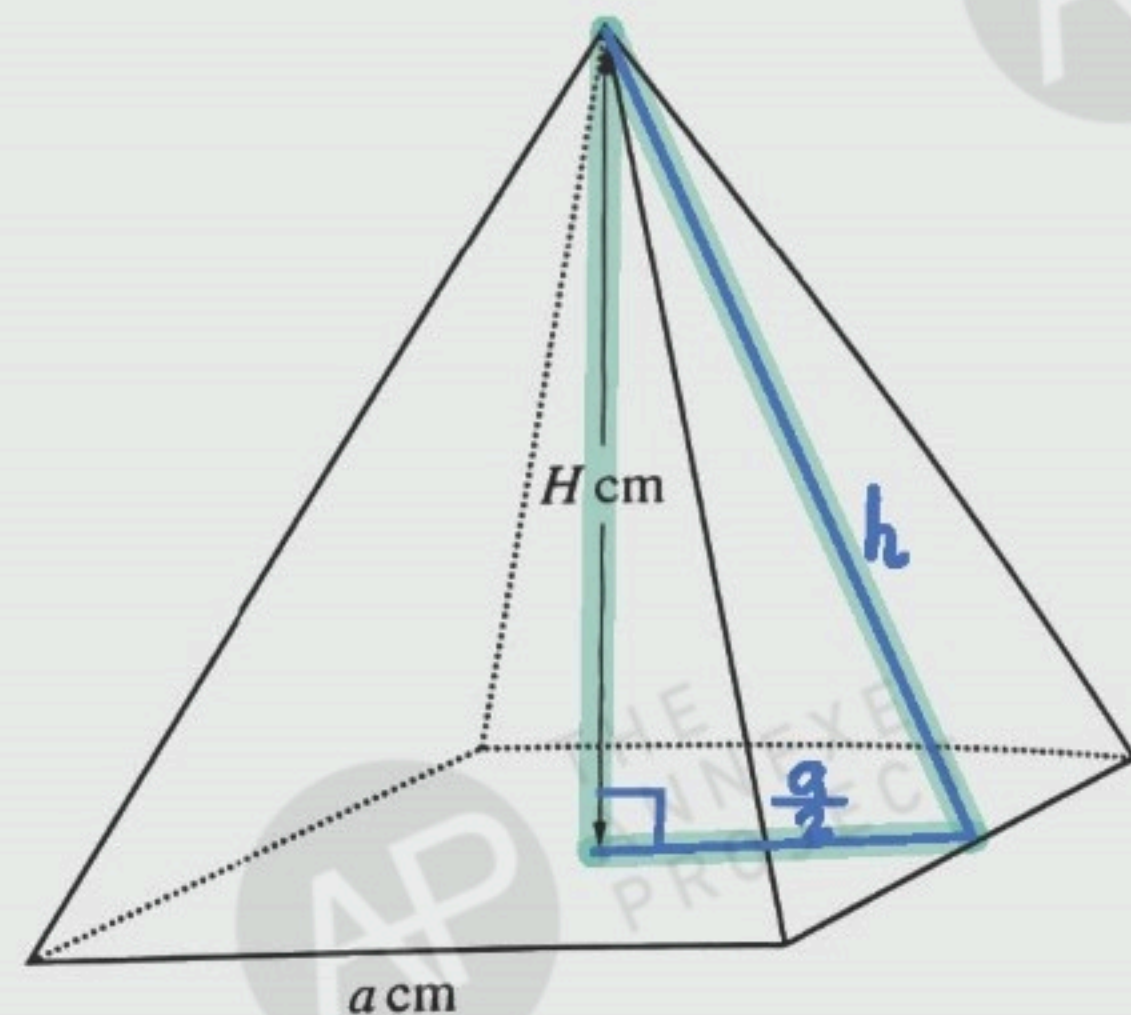


Fig. 2

Fig. 1 shows the net of a square-based pyramid cut from a square of cardboard of side length 30 cm. The net consists of a square of side length a cm and four isosceles triangles, each with base a cm and perpendicular height h cm. The net is folded to form a pyramid which has a square base of side length a cm and vertical height H cm, as shown in Fig. 2.

(i) Show that $H^2 = 225 - 15a$.

[2]

$$H^2 = h^2 - \left(\frac{a}{2}\right)^2 \quad \text{--- (1)}$$

$$2h + a = 30$$

$$h = \frac{30 - a}{2} \quad \text{--- (2)}$$

Sub (2) into (1):

$$\begin{aligned} H^2 &= \left(\frac{30-a}{2}\right)^2 - \left(\frac{a}{2}\right)^2 \\ &= \frac{900 - 60a + a^2 - a^2}{4} \end{aligned}$$

$$= 225 - 15a \quad (\text{shown}).$$





- (ii) Find the maximum possible volume of the pyramid. You do not need to show that this value is a maximum. [5]

[The volume of a square-based pyramid is $\frac{1}{3} \times \text{base area} \times \text{height}$.]

$$V = \frac{1}{3} \times a^2 \times \sqrt{225 - 15a}$$

$$\begin{aligned} \frac{dV}{da} &= \frac{a^2}{3} \cdot \frac{1}{2} (225 - 15a)^{-\frac{1}{2}} (-15) + (225 - 15a)^{\frac{1}{2}} \cdot \left(\frac{2}{3}a\right) \\ &= (225 - 15a)^{-\frac{1}{2}} \left[-\frac{5}{2}a^2 + \frac{2}{3}a(225 - 15a)\right] \\ &= \frac{-\frac{25}{2}a^2 + 150a}{\sqrt{225 - 15a}} \end{aligned}$$

$$\text{Let } \frac{dV}{da} = 0 : -\frac{25}{2}a^2 + 150a = 0$$

$$a(150 - \frac{25}{2}a) = 0$$

$$a = 0 \text{ (rej.) or } \underline{a = 12 \text{ cm}}$$

$$V_{\text{max.}} = \frac{1}{3} \times 144 \times \sqrt{225 - 180} = 321.99 = \underline{322 \text{ cm}^3}$$

- (iii) (a) Find the value of a for which the total surface area of the four triangular faces of the pyramid is a maximum. You do not need to show that this value is a maximum. [3]

$$\begin{aligned} A &= 4 \left[\frac{1}{2} \times a \times \frac{30-a}{2} \right] \\ &= 30a - a^2 \end{aligned}$$

$$\frac{dA}{da} = 30 - 2a$$

$$\text{Let } 30 - 2a = 0 \quad \therefore \underline{a = 15 \text{ cm}}$$

- (b) Describe the shape formed from the net in this case. [1]

$$\text{When } a = 15, \quad H^2 = 225 - 15^2 = 0$$

The height of the pyramid would be 0 cm.





Section B: Probability and Statistics [60 marks]

- 5 Shania and Tina are playing a game. Shania has a bag containing one green disc, r red discs and $2r$ blue discs, where $r > 1$. A red disc is worth 5 points, a blue disc is worth 2 points and the green disc is worth 0 points. Tina takes two discs from the bag at random. Tina's score is found by multiplying together the number of points for each of the two discs she takes.

(i) State Tina's possible scores.

[1]

$$\begin{aligned} \{R, G\} &= 0 & \{G, B\} &= 0 \\ \{R, R\} &= 25 & \{B, B\} &= 4 \\ \{R, B\} &= 10 \end{aligned}$$

(ii) Show that the expectation of Tina's score is $\frac{27r-11}{3r+1}$, and find an expression for the variance.

[7]

Let X be the discrete r.v. denoting Tina's score:

r	0	4	10	25
$P(X=r)$	$\frac{2}{3(3r+1)}$	$\frac{2(2r-1)}{3(3r+1)}$	$\frac{4r}{3(3r+1)}$	$\frac{r-1}{3(3r+1)}$

$$\begin{aligned} P(X=0) &= \{R, G\} + \{G, B\} \\ &= \left[\frac{1}{3r+1} \times \frac{1}{3r} \times 2! \right] + \left[\frac{1}{3r+1} \times \frac{2r}{3r} \times 2! \right] \\ &= \frac{2}{3(3r+1)} + \frac{4}{3(3r+1)} = \frac{2}{3r+1} \end{aligned}$$

$$\begin{aligned} P(X=4) &= \{B, B\} \\ &= \frac{2r}{3r+1} \times \frac{2r-1}{3r} = \frac{2(2r-1)}{3(3r+1)} \end{aligned}$$

$$\begin{aligned} P(X=10) &= \{R, B\} = \frac{r}{3r+1} \times \frac{2r}{3r} \times 2! \\ &= \frac{4r}{3(3r+1)} \end{aligned}$$

$$P(X=25) = \{R, R\} = \frac{1}{3r+1} \times \frac{r-1}{3r} = \frac{r-1}{3(3r+1)}$$



$$E(X) = 4\left(\frac{4r-2}{3(3r+1)}\right) + 10\left(\frac{4r}{3(3r+1)}\right) + 25\left(\frac{r-1}{3(3r+1)}\right)$$

$$= \frac{16r-8+40r+25r-25}{3(3r+1)}$$

$$= \frac{81r-33}{3(3r+1)} = \frac{27r-11}{3r+1} \quad (\text{shown}).$$

$$E(X^2) = 16\left(\frac{4r-2}{3(3r+1)}\right) + 100\left(\frac{4r}{3(3r+1)}\right) + 625\left(\frac{r-1}{3(3r+1)}\right)$$

$$= \frac{64r-32+400r+625r-625}{3(3r+1)}$$

$$= \frac{363r-219}{3r+1}$$

$$\therefore \text{Var}(X) = \frac{363r-219}{3r+1} - \left(\frac{27r-11}{3r+1}\right)^2$$

$$= \frac{(363r-219)(3r+1) - (27r-11)^2}{(3r+1)^2}$$

$$= \frac{360r^2+300r-340}{(3r+1)^2}$$



(iii) Given that the variance of Tina's score is 38, find the value of r .

[2]

$$\frac{360r^2 + 300r - 340}{(3r+1)^2} = 38$$

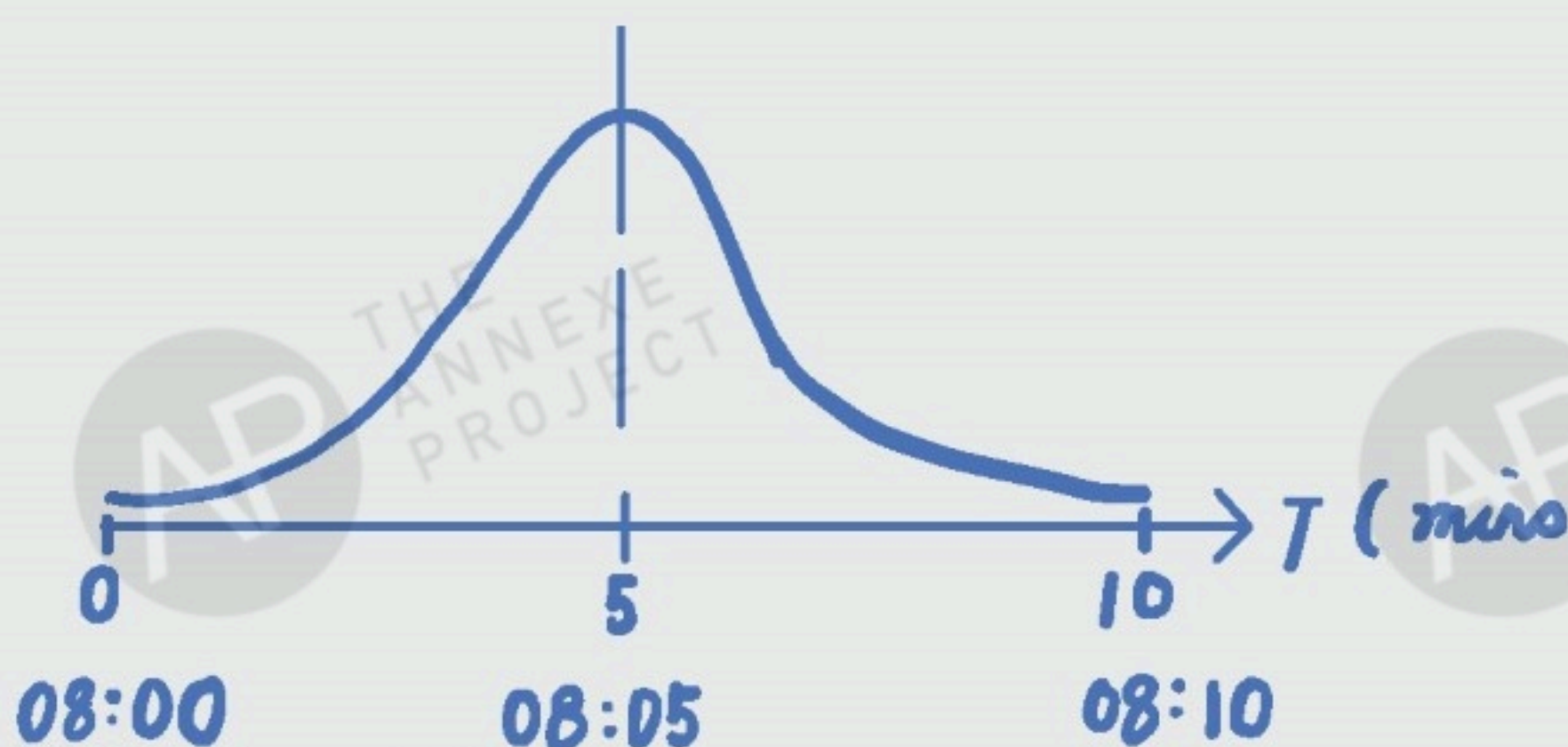
By GC: $r = 3$

- 6 In this question you should assume that T , W and D follow independent normal distributions.

James leaves home to go to work at T minutes past 8 am each day, where T follows the distribution $N(5, 1.2^2)$.

- (i) Sketch this distribution for the period from 8 am to 8.10 am.

[2]



- (ii) Find the probability that, on a randomly chosen day, James leaves for work later than 8.06 am.

[1]

$$T \sim N(5, 1.2^2)$$

$$P(T > 6) = 0.20233$$

$$= \underline{0.202}$$

When the weather is fine, James walks to work. The time, W minutes, he takes to walk to work follows the distribution $N(21, 3^2)$. James is supposed to start work at 8.30 am.

- (iii) Find the probability that, on a randomly chosen day when James walks, he is late for work. [2]

$$\text{Var}(T+W) = 1.2^2 + 3^2 = 10.44$$

$$T+W \sim N(26, 10.44)$$

$$P(T+W > 30) = 0.10786$$

$$= \underline{0.108}$$

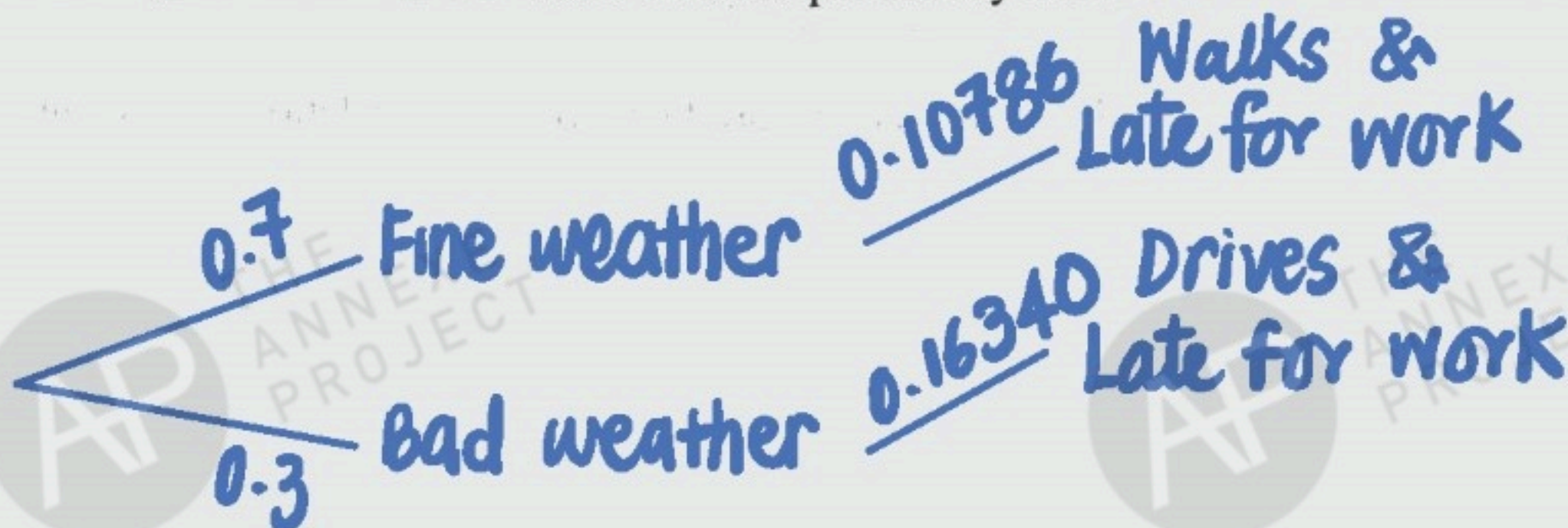


When the weather is not fine, James drives to work. He still leaves at T minutes past 8 am each day; the time, D minutes, he takes to drive to work follows the distribution $N(19, 6^2)$.

On average, the weather is fine on 70% of mornings.

(iv) One day, James is late for work. Find the probability that the weather is fine that day.

[5]



$$\text{Var}(T+D) = 1.2^2 + 6^2 = 37.44$$

$$T+D \sim N(24, 37.44)$$

$$P(T+D > 30) = 0.16340 \\ = 0.163$$

$$P(\text{Fine weather} \mid \text{James is late for work})$$

$$= \frac{0.7 \times 0.10786}{(0.7 \times 0.10786) + (0.3 \times 0.16340)}$$

$$= 0.60633$$

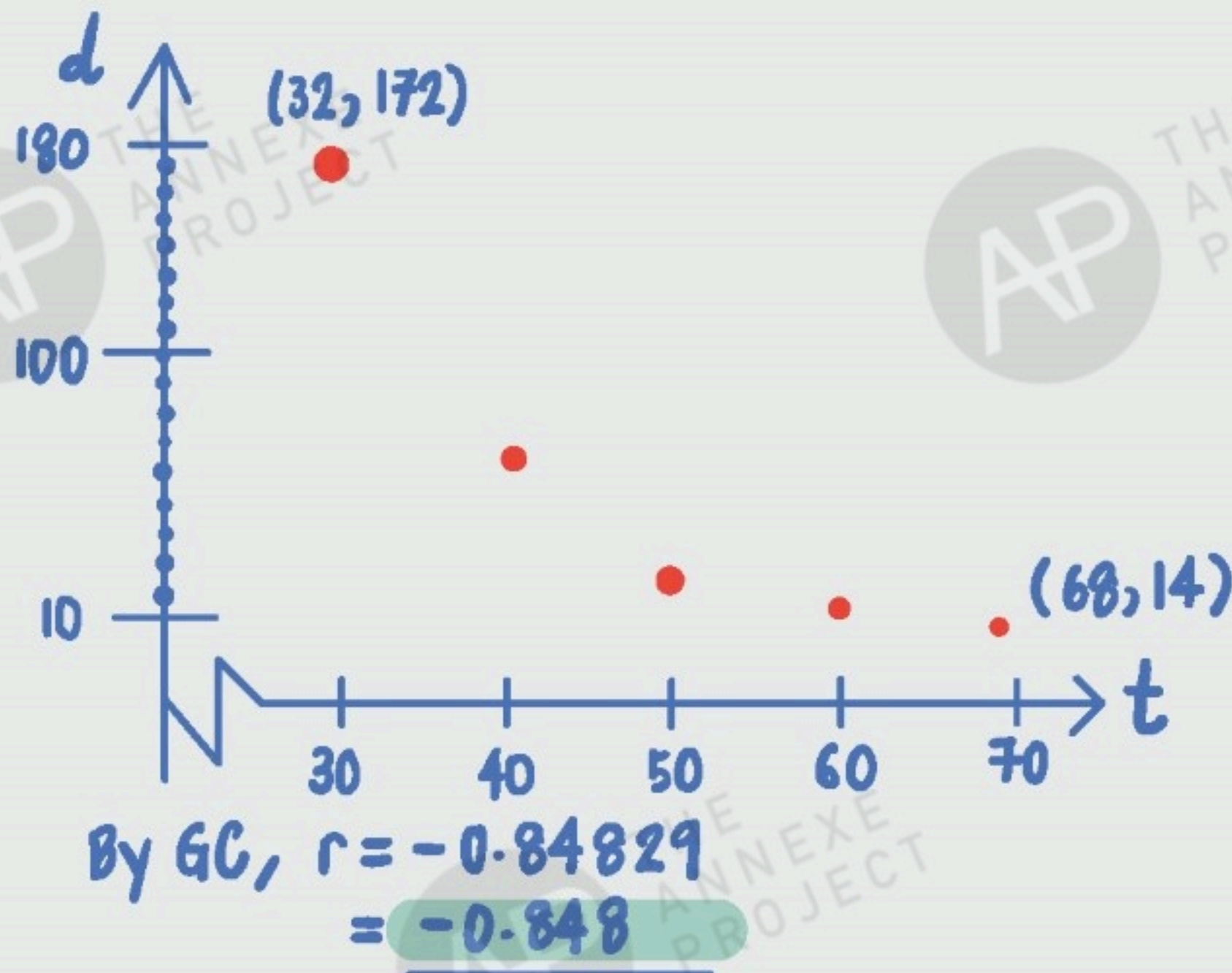
$$= \underline{0.606}$$



- 7 A study into the germination of parsnip seeds in the 1980s produced the following data for the average number of days, d , taken for a seed to germinate at different soil temperatures, t , measured in degrees Fahrenheit.

t	32	41	50	59	68
d	172	57	27	19	14

- (i) Sketch a scatter diagram of the data. State the product moment correlation coefficient between d and t . [2]



- (ii) Lim thinks the data can be modelled by the regression equation $d = -a + bu$, where $u = \frac{1}{t}$. Find the values of a and b for Lim's model, giving the values correct to 3 significant figures. State the product moment correlation coefficient between d and u . [3]

By GC, $d = -145 + 9460u$
 where $a = 145$, $b = 9460$

$r = 0.94123$
 $= 0.941$





The study also found that, at a soil temperature of 86 degrees Fahrenheit, parsnip seeds took an average of 32 days to germinate.

- (iii) Determine whether Lim's model fits this additional data.

[1]

Since the value of $t = 86$ lies outside of the range of data collected, this additional data does not fit Lim's model.

Because, extrapolation does not give a reliable estimation.

A temperature of F degrees Fahrenheit is equivalent to a temperature of C degrees Celsius, where $C = \frac{5}{9}(F - 32)$.

- (iv) Write Lim's equation from part (ii) in terms of d and T , where T is the temperature in degrees Celsius. [2]

$$d = -145 + \frac{9460}{t}$$

$$\text{since } T = \frac{5}{9}(t - 32)$$

$$\frac{9T}{5} = t - 32$$

$$t = 32 + \frac{9T}{5}$$

$$\text{hence, } d = -145 + \frac{9460}{32 + \frac{9T}{5}}$$

$$= -145 + \frac{47300}{160 + 9T}$$





- 8 In a game, a computer randomly chooses 12 shapes from 11 circles and 17 rectangles. The number of rectangles chosen is denoted by R .

(i) Show that $P(R = 1) < P(R = 2)$.

[2]

$$P(R=1) = \frac{{}^{17}C_1 \times {}^{11}C_{11}}{{}^{28}C_{12}} = \frac{5.5881 \times 10^{-7}}{}$$

$$P(R=2) = \frac{{}^{17}C_2 \times {}^{11}C_{10}}{{}^{28}C_{12}} = \frac{4.9175 \times 10^{-5}}{}$$

$$\text{Hence, } P(R=1) < P(R=2)$$

The number of rectangles available is now increased by r . The computer randomly chooses 12 shapes from the 11 circles and $(17 + r)$ rectangles. The probability that 4 rectangles are chosen is now 15 times the probability that 3 rectangles are chosen.

(ii) Find the value of r .

[5]

$$P(R=4) = \frac{{}^{17+r}C_4 \times {}^{11}C_8}{{}^{28+r}C_{12}}$$

$$P(R=3) = \frac{{}^{17+r}C_3 \times {}^{11}C_9}{{}^{28+r}C_{12}}$$

Note: Formula in MF26

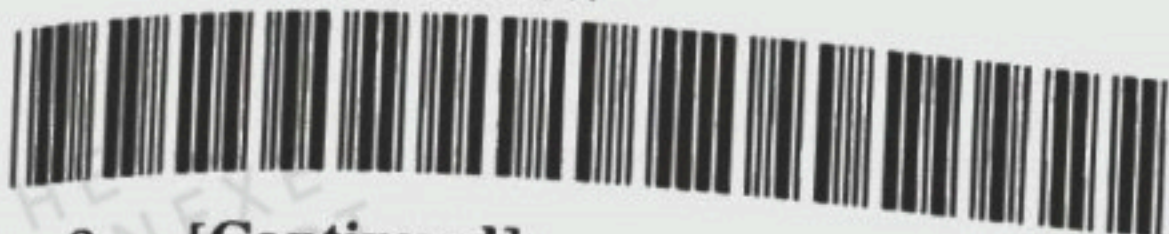
$${}_nC_r = \frac{n!}{(n-r)! r!}$$

$$\text{Given } P(R=4) = 15 \times P(R=3)$$

$$\frac{{}^{17+r}C_4 \times {}^{11}C_8}{{}^{28+r}C_{12}} = 15 \times \frac{{}^{17+r}C_3 \times {}^{11}C_9}{{}^{28+r}C_{12}}$$

$$\frac{{}^{17+r}C_4}{{}^{17+r}C_3} = \frac{15 \times {}^{11}C_9}{{}^{11}C_8}$$





8 [Continued]

17

$$\frac{(17+r)!}{(13+r)! \times 4!} = 5$$

$$\frac{(17+r)!}{(14+r)! \times 3!}$$

$$\frac{\cancel{(17+r)!}}{(13+r)! \times 4!} = \frac{5 \cancel{(17+r)!}}{(14+r)! \times 3!}$$

$$\therefore (14+r)! \times 6 = 5 \times (13+r)! \times 24$$

$$(14+r) \times \cancel{(13+r)!} \times 6 = 120 \times \cancel{(13+r)!}$$

$$14+r = 20$$

$$\underline{r = 6}$$





- 9 A factory produces ballpoint pens. On average 6% of the pens are faulty. The pens are packed in boxes of 100 for sale to retail outlets. It should be assumed that the number of faulty pens in a box of 100 pens follows a binomial distribution.

For quality control purposes a random sample of 10 pens from each box is tested. If 2 or fewer faulty pens are found in the sample of 10, the box is accepted for sale. Otherwise the box is rejected.

- (i) Explain what is meant by a random sample in this context. [1]

A random sample means that every pen in the box has an equal chance of being chosen.

- (ii) Find the probability that a randomly chosen box of 100 pens is accepted for sale. [1]

Let X be the r.v. denoting number of faulty pens in a sample of 10.

$$X \sim B(10, 0.06)$$

$$P(X \leq 2) = 0.98116 = \underline{0.981}$$

- (iii) One morning 75 boxes are tested in this way. Find the probability that more than 5% of these boxes are rejected. [4]

Let Y be the r.v. denoting number of boxes rejected out of 75 boxes.

$$Y \sim B(75, 1 - 0.98116)$$

$$\begin{aligned} P(Y > 3.75) &= P(Y \geq 4) \\ &= 1 - P(Y \leq 3) \\ &= 0.053471 \\ &= \underline{0.0535} \end{aligned}$$

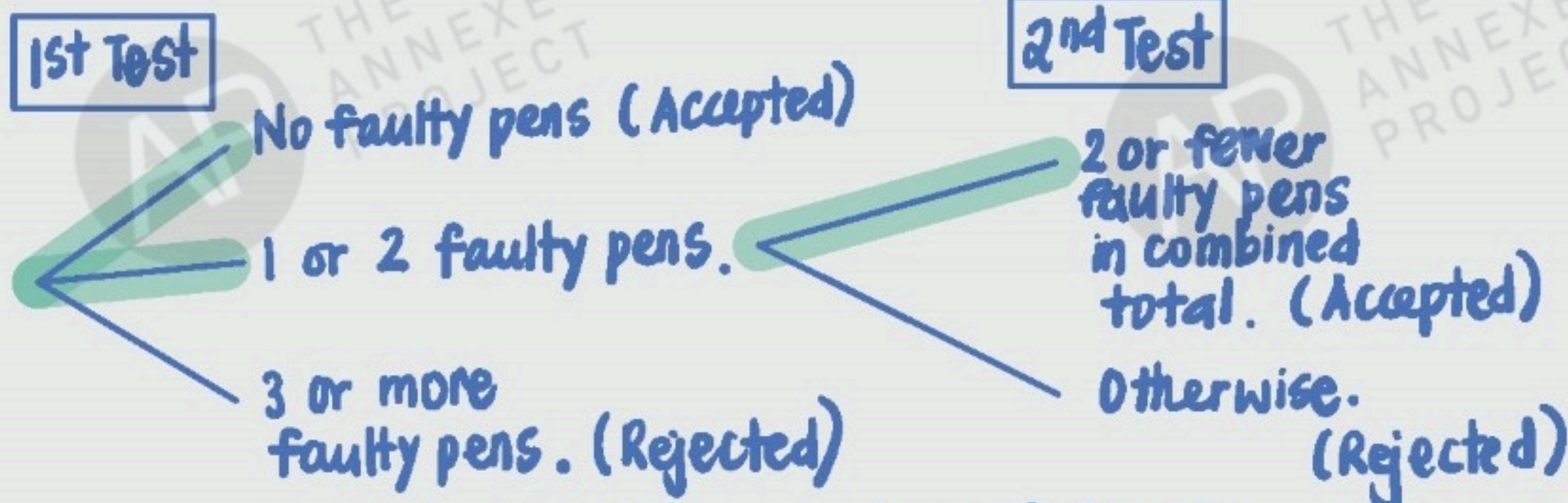


DO NOT WRITE IN THIS MARGIN

An alternative testing procedure is trialled in which a random sample of 5 pens is initially taken from a box and tested.

- If there are no faulty pens in this sample of 5 the box is accepted.
- If there are 3 or more faulty pens in this sample of 5 the box is rejected.
- If there are 1 or 2 faulty pens in this sample a second random sample of 5 pens is taken from the box. When the second sample has been tested, the box is accepted if the **total** number of faulty pens found in the combined sample of 10 is 2 or fewer and rejected otherwise.

(iv) Find the probability that a randomly chosen box of 100 pens is accepted for sale when the alternative testing procedure is used. [5]



Let X be the r.v. denoting number of faulty pens in a sample of 5.

$$X \sim B(5, 0.06)$$

$$\text{Prob. of Acceptance} = P(X=0) + [P(X_1=1) \cdot P(X_2 \leq 1) + P(X_1=2) \cdot P(X_2=0)]$$

$$= 0.73390 + [0.22676 + 0.021944]$$

$$= 0.982604$$

$$= \underline{0.983}$$

(v) Explain why the factory manager might prefer to use the alternative testing procedure. [1]

Less time is required to test random samples of 5 pens compared to 10 pens.

[Turn over]



- 10 Carbon steel is made by adding carbon to iron; this makes the iron stronger, though less flexible. A steel manufacturing company makes carbon steel in the form of round bars. The process is designed to manufacture bars in which the amount of carbon in each bar, by weight, is 1.5%. It is known that the percentage of carbon in the steel bars is distributed normally, and that the standard deviation is 0.09%.

After comments from customers, the production manager wishes to test, at the 5% level of significance, if the percentage of carbon in the steel bars is, in fact, 1.5%. He examines a random sample of 15 bars to determine the percentage of carbon in each bar.

- (i) Find the critical region for this test.

[4]

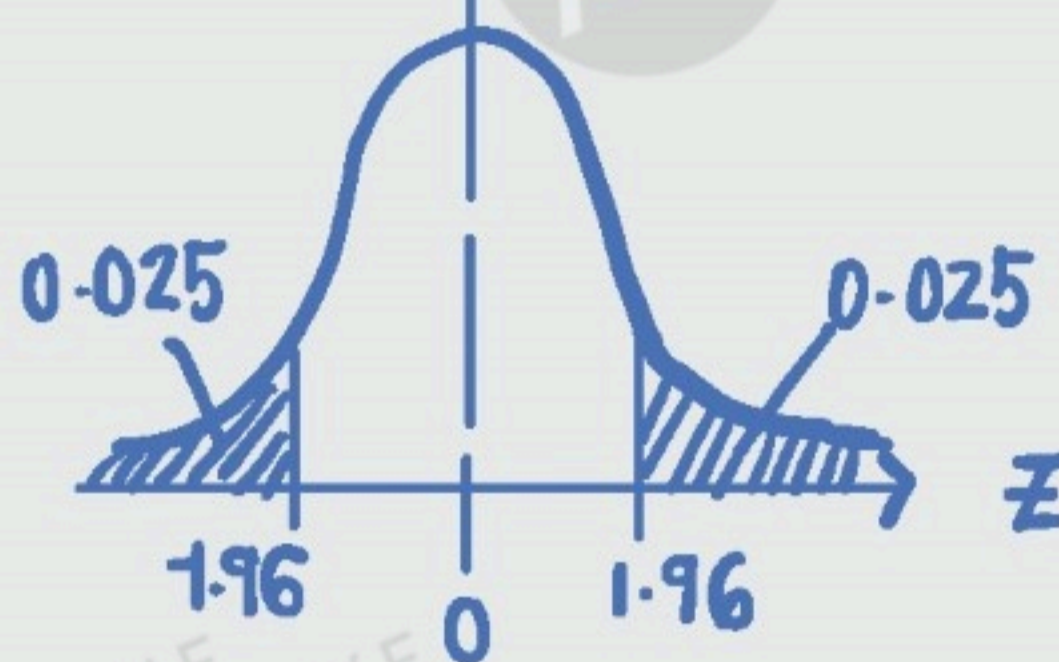
Let X be the percentage of carbon in each bar.
 $X \sim N(\mu, 0.09^2)$

To test $H_0: \mu = 1.5$ against
 $H_1: \mu \neq 1.5$ at 5% level of significance.

$$\bar{X} \sim N\left(\mu, \frac{0.09^2}{15}\right)$$

Using Z -test: $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$

$$= \frac{\bar{X} - 1.5}{0.09/\sqrt{15}} = \frac{\sqrt{15}(\bar{X} - 1.5)}{0.09}$$



H_0 is rejected if z -value is within the shaded region, i.e. critical region.

$$\Rightarrow H_0 \text{ is rejected when } Z \leq -1.96 \text{ or } Z \geq 1.96$$

$$\frac{\sqrt{15}(\bar{X} - 1.5)}{0.09} \leq -1.96 \text{ or } \frac{\sqrt{15}(\bar{X} - 1.5)}{0.09} \geq 1.96$$

$$\bar{X} \leq 1.4545 \text{ or } \bar{X} \geq 1.5456$$

Since \bar{x} is percentage of carbon,

$$\underline{0 \leq \bar{x} \leq 1.45} \text{ or } \underline{1.55 \leq \bar{x} \leq 100}$$





The company recently launched a new line of flat bars made from mild steel. In mild steel the amount of carbon in each bar, by weight, is 0.25%.

Comments from customers suggest that these bars are not sufficiently flexible, which makes the production manager suspect that too much carbon has been added. He decides to perform a hypothesis test on a random sample of 40 of the new flat bars to find out if this is the case.

- (ii) Explain why the production manager takes a sample of 40 flat bars for this test when he only took a sample of 15 round bars in his earlier test. [2]

In the earlier test, the distribution of the carbon in the round bars is normally distributed.

For the flat bars, the distribution of the carbon is unknown. A sample size of $n \geq 30$ is required to perform Central Limit Theorem, so that the distribution can be approximated to a Normal distribution.

The amounts of carbon, $x\%$, in a random sample of 40 flat bars are summarised as follows.

$$n = 40 \quad \Sigma x = 10.16 \quad \Sigma x^2 = 2.586342$$

- (iii) Calculate unbiased estimates of the population mean and variance for the percentage amount of carbon in the flat bars. [2]

$$\bar{x} = \frac{\Sigma x}{n}$$

$$= \frac{10.16}{40}$$

$$= \underline{0.254}$$

$$s^2 = \frac{1}{n-1} \left[\Sigma x^2 - \frac{(\Sigma x)^2}{n} \right]$$

$$= \frac{1}{39} \left[2.586342 - \frac{10.16^2}{40} \right]$$

$$= 1.4621 \times 10^{-4}$$

$$= \underline{1.46 \times 10^{-4}}$$



- (iv) Test, at the $2\frac{1}{2}\%$ level of significance, whether the mean amount of carbon in the flat bars is more than 0.25%. You should state your hypotheses and define any symbols that you use. [5]

To test $H_0 : \mu = 0.25$ against

$H_1 : \mu > 0.25$ at 2.5% level of significance.

(μ represents the population mean of carbon percentage in each flat bar.)

H_0 : null hypothesis

H_1 : alternative hypothesis)

Since n is large, by CLT : $\bar{X} \sim N(0.25, \frac{1.46 \times 10^{-4}}{40})$ approx.

Using Z-test : $Z = \frac{\bar{X} - \mu}{s/\sqrt{n}} \sim N(0,1)$

$$= \frac{0.254 - 0.25}{\frac{\sqrt{1.46 \times 10^{-4}}}{\sqrt{40}}}$$

$$= 2.0937$$

By GC: p-value = 0.0181

Since p-value < 0.025 , we reject H_0 at 2.5% level of significance. We conclude that the mean amount of carbon in each flat bar is more than 0.25%.

