* 0020057117401 *

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 2024

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 **22** printed pages and **2** blank pages.

Singapore Examinations and Assessment Board

Cambridge Assessment International Education

DC (DE/CB) 329671/6

Section A: Pure Mathematics [40 marks]

1 **Do not use a calculator in answering this question.**

The complex number $-2+\mathrm{i}$ is denoted by z.

(a) Find the real numbers a and b such that $z = az^* + b$. [2]

$$-2+i = a(-2-i) + b$$
$$-2+i = (-2a+b) - ai$$

Comparing coefficients : $-a = 1$, $\underline{a = -1}$

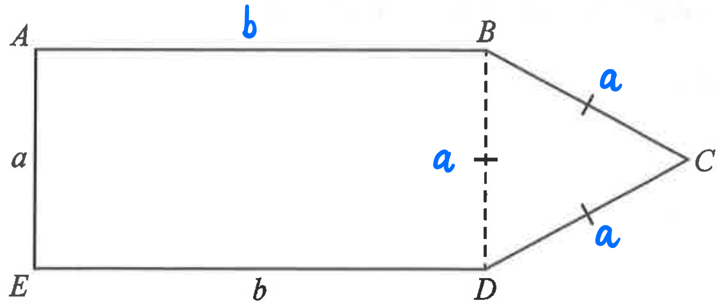
$-2 = -2a + b$
$-2 = 2 + b$
$\therefore \underline{b = -4}$

The complex number $1-3\mathrm{i}$ is denoted by w.

(b) Without using a calculator, evaluate $wz - \dfrac{w}{z}$. Give your answer in the form $c+d\mathrm{i}$, where c and d are real numbers. [4]

$$wz - \frac{w}{z} = (1-3i)(-2+i) - \frac{1-3i}{-2+i}$$
$$= (-2+i+6i+3) - \frac{(1-3i)(-2-i)}{(-2+i)(-2-i)}$$
$$= (1+7i) - \frac{(-2-i+6i-3)}{4+1}$$
$$= (1+7i) - \frac{1}{5}(-5+5i)$$
$$= 1+7i+1-i$$
$$= \underline{2+6i}$$

2



A gardener designs a flower bed *ABCDE* in the shape of a rectangle with an equilateral triangle on one of the shorter sides. Side *AE* is of length a m and side *ED* is of length b m (see diagram). The total perimeter of the flower bed is 20 m.

Find the maximum possible area of the flower bed, showing that it is a maximum value. Give your answer correct to 4 significant figures. [6]

$P = 3a + 2b = 20$

$\therefore \quad 2b = 20 - 3a$

$b = \frac{20-3a}{2}$ —— 1

$A = ab + \frac{1}{2}a^2 \sin 60°$

$= a\left(\frac{20-3a}{2}\right) + \frac{1}{2}a^2 \cdot \frac{\sqrt{3}}{2}$

$= 10a - \frac{3}{2}a^2 + \frac{\sqrt{3}}{4}a^2$

$= 10a + \left(\frac{\sqrt{3}}{4} - \frac{6}{4}\right)a^2$

$= 10a + \frac{\sqrt{3}-6}{4}a^2$

Step 1 : $\frac{dA}{da} = 10 + \frac{\sqrt{3}-6}{2}a$

Step 2 : Let $\frac{dA}{da} = 0$

$10 + \frac{\sqrt{3}-6}{2}a = 0$

$a = 4.68609$

Step 3 : $\frac{d^2A}{da^2} = \frac{\sqrt{3}-6}{2} < 0$

hence , A is max. when $a = 4.68609$ m

Step 4 : $A_{max.} = 10(4.68609) + \frac{\sqrt{3}-6}{4}(4.68609)^2$

$= \underline{23.43 \text{ m}^2}$

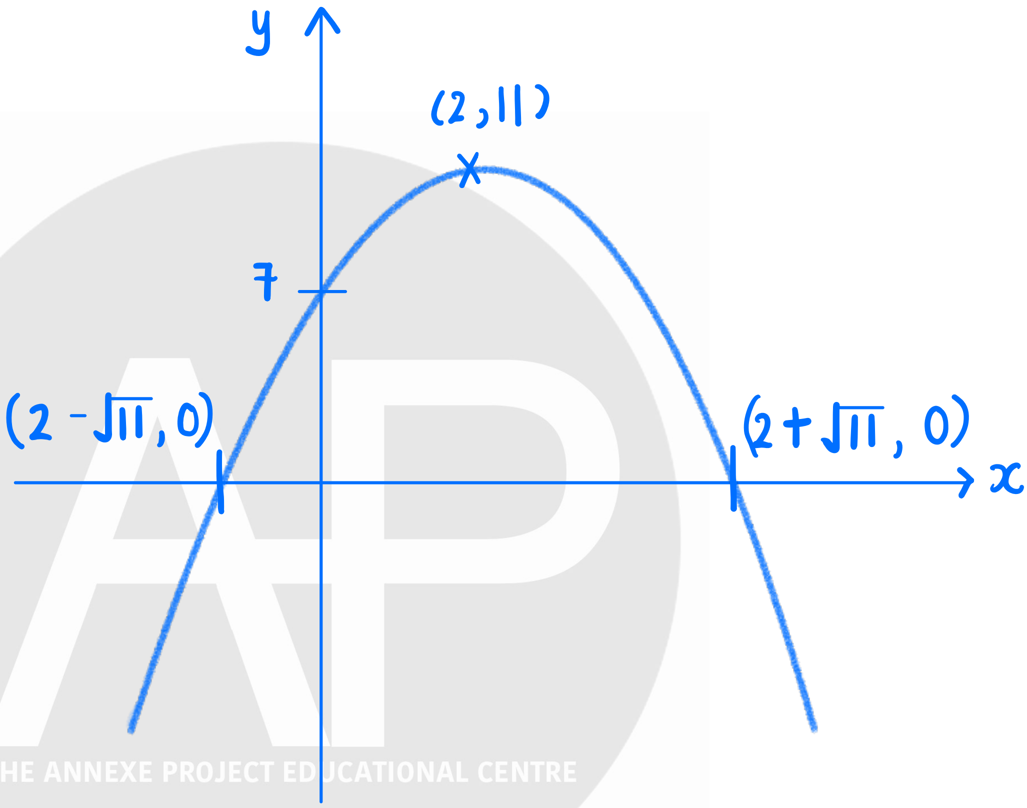
3 The function f is such that $\mathrm{f}(x) = -x^2 + 4x + 7$, for $x \in \mathbb{R}$.

(a) Find the range of f. [2]

$$-(x^2 - 4x - 7) = -[(x-2)^2 - 2^2 - 7]$$
$$= 11 - (x-2)^2$$

$$\underline{R_f = (-\infty, 11]}$$

(b) Sketch the graph of f, giving the exact coordinates of the points where the curve crosses the axes. [2]



Let $11 - (x-2)^2 = 0$

$$x - 2 = \pm\sqrt{11}$$
$$\underline{x = 2 \pm \sqrt{11}}$$

The function g is such that $g(x) = \frac{2}{x} + 1$, for $1 \leqslant x \leqslant 10$.

(c) Explain how you know g^{-1} exists. [1]

$g(x)$ is a 1-1 function for $1 \leq x \leq 10$,
hence, $g^{-1}(x)$ exists.

(d) Find the value of $fg^{-1}(1.5)$. [3]

Let $y = \frac{2}{x} + 1$

$\frac{2}{x} = y - 1$

$x = \frac{2}{y-1}$

$\therefore g^{-1}(x) = \frac{2}{x-1}, \quad 1.2 \leq x \leq 3$

$g^{-1}(1.5) = \frac{2}{1.5-1} = 4$

Hence, $fg^{-1}(1.5) = f(4) = -(4)^2 + 4(4) + 7$
$= \underline{7}$

4 The sum of the first n terms of the series T is given by $2n^3 - 8n^2 - 4n$.

(a) Find an expression for t_n, the nth term of series T. [2]

$$t_n = S_n - S_{n-1}$$
$$= (2n^3 - 8n^2 - 4n) - [2(n-1)^3 - 8(n-1)^2 - 4(n-1)]$$
$$= 2n^3 - 8n^2 - 4n - [2(n^3 - 3n^2 + 3n - 1) - 8(n^2 - 2n + 1) - 4n + 4]$$
$$= \cancel{2n^3} - 8n^2 - 4n - (\cancel{2n^3} - 14n^2 + 18n - 6)$$
$$= \underline{6n^2 - 22n + 6}$$

The nth term of the series U is given by $u_n = 50n - 204$.

(b) Find the values of n for which $u_n = t_n$. [2]

$$50n - 204 = 6n^2 - 22n + 6$$
$$6n^2 - 72n + 210 = 0$$
$$n^2 - 12n + 35 = 0$$
$$(n - 7)(n - 5) = 0$$
$$\underline{n = 5 \text{ or } 7}$$

The nth term of the series V is given by $v_n = 3n + 16$.

(c) Find the smallest number greater than 100 that is in both series U and series V. [2]

By GC : 196

Step 1: Change mode to $\boxed{\text{SEQ}}$

Step 2: key $\boxed{\text{y=}}$

Type: $u(n) = 50n - 204$

$v(n) = 3n + 16$

Step 3: $\boxed{\text{2nd}}$ $\boxed{\text{graph}}$

The nth term of the series W is given by $w_n = 3n^2 - 5n + 7$.

(d) **(i)** Explain why all the terms in series W are odd numbers. [2]

If n is odd : $3n^2$ is odd , $5n$ is odd ,

(odd − odd) + odd

= (even) + odd = odd

If n is even : $3n^2$ is even , $5n$ is even ,

(even − even) + odd

= (even) + odd = odd .

(ii) Hence explain why series U and series W do not have any terms in common. [1]

Every term in series U is an even number, hence series U and W do not have any common terms.

The suggested solutions are prepared by Mr Alvin Yeo. Mr Yeo will hold no liability for any errors.

5 **(a)** The graph of $y = \mathrm{f}(x)$ intersects the x-axis at the point $(a, 0)$ and the y-axis at the point $(0, b)$.

(i) The graph of $y = \mathrm{f}(x)$ is shown in Fig. 1. The scales on the x- and y-axes are the same.

Sketch the graph of $y = \mathrm{f}^{-1}(x)$ on Fig. 1, labelling the intersections with the axes. [1]

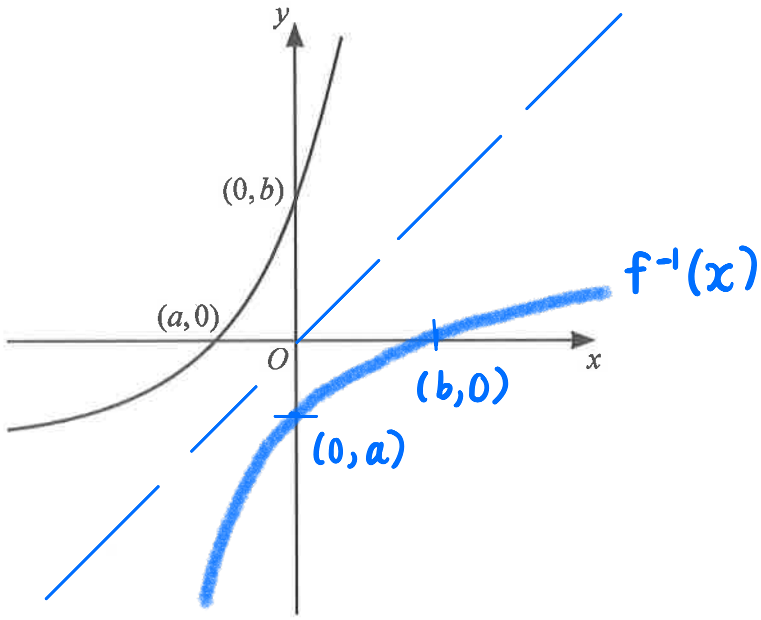


Fig. 1

(ii) The graph of $y = \mathrm{f}(x)$ is also shown in Fig. 2. The scales on the x- and y-axes are the same.

Sketch the graph of $y = 2\mathrm{f}^{-1}(x-1)$ on Fig. 2, labelling the intersection with the x-axis. [2]

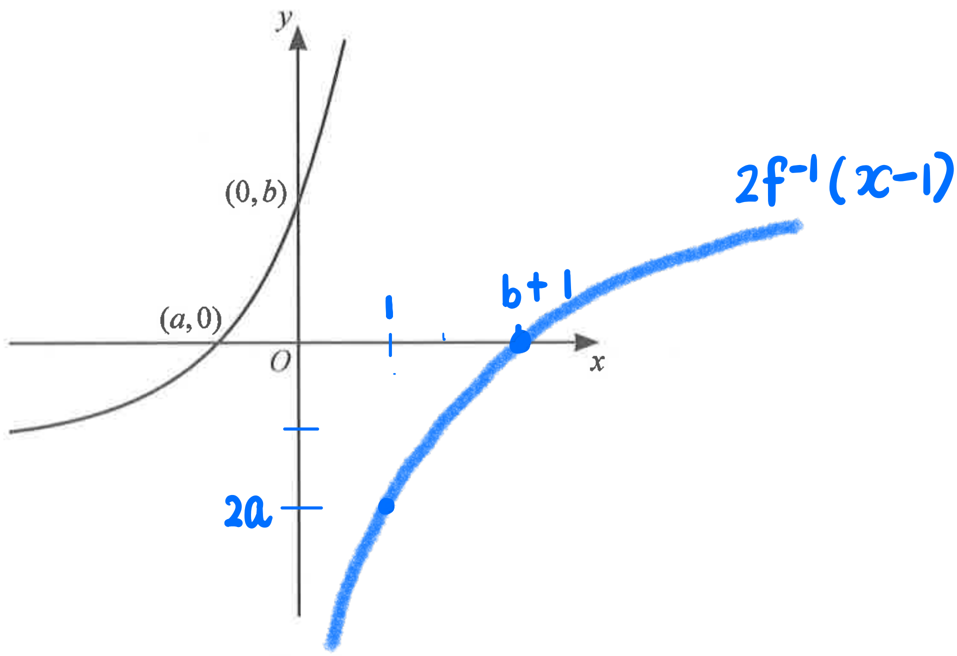


Fig. 2

$\mathrm{f}^{-1}(x) \xrightarrow[\text{y with } \frac{1}{2}y]{\text{replace}} 2\mathrm{f}^{-1}(x) \xrightarrow[x \text{ with } x-1]{\text{replace}} 2\mathrm{f}^{-1}(x-1)$

$(0, a) \longrightarrow (0, 2a) \longrightarrow (1, 2a)$

$(b, 0) \longrightarrow (b, 0) \longrightarrow (b+1, 0)$

(b) The graph of $y = \mathrm{g}(x)$ intersects the x-axis at the points $(s, 0)$, $(t, 0)$ and $(u, 0)$ and the y-axis at the point $(0, v)$.

(i) The graph of $y = \mathrm{g}(x)$ is shown in Fig. 3.1.

Sketch the graph of $y = |\mathrm{g}(x)|$ on Fig. 3.2, labelling the points where the graph intersects or touches the axes. [2]

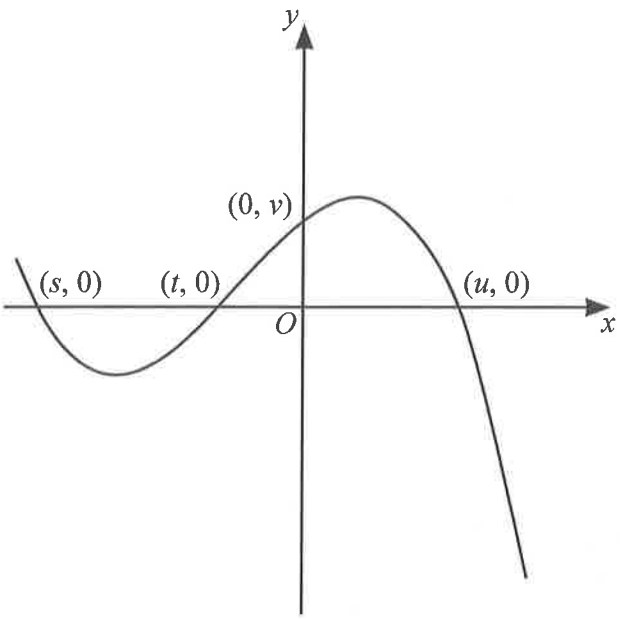


Fig. 3.1

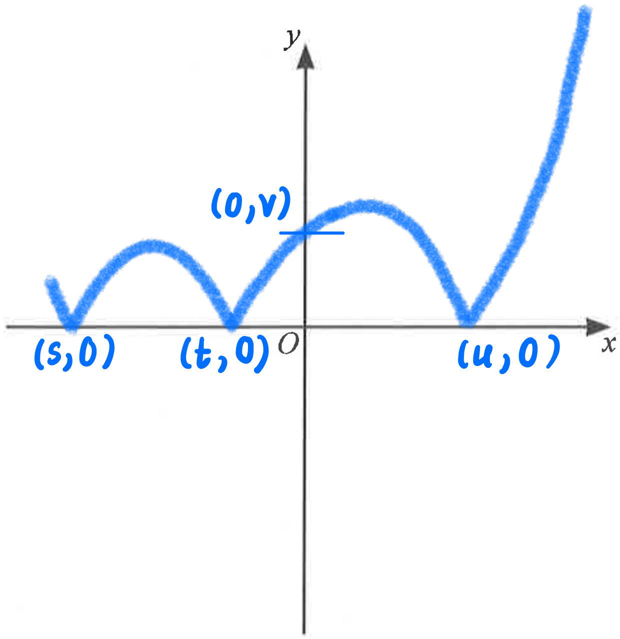


Fig. 3.2

(ii) The graph of $y = \mathrm{g}(x)$ is also shown in Fig. 4.1.

Sketch the graph of $y = \mathrm{g}(|x|)$ on Fig. 4.2, labelling the points where the graph intersects or touches the axes. [2]

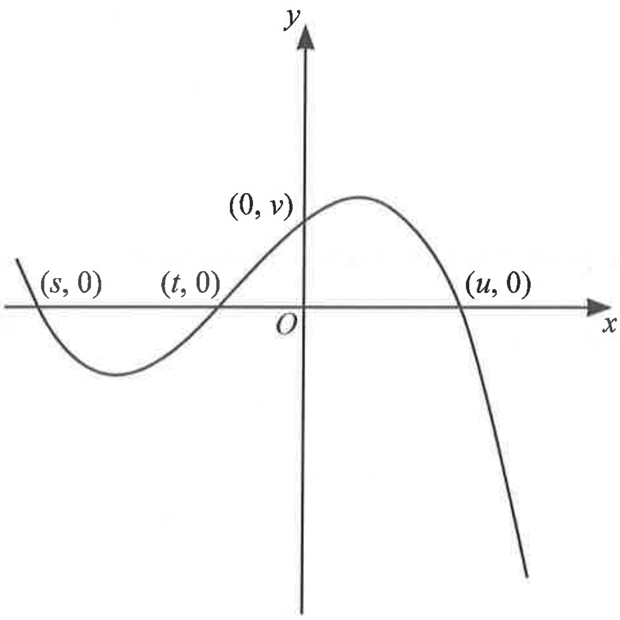


Fig. 4.1

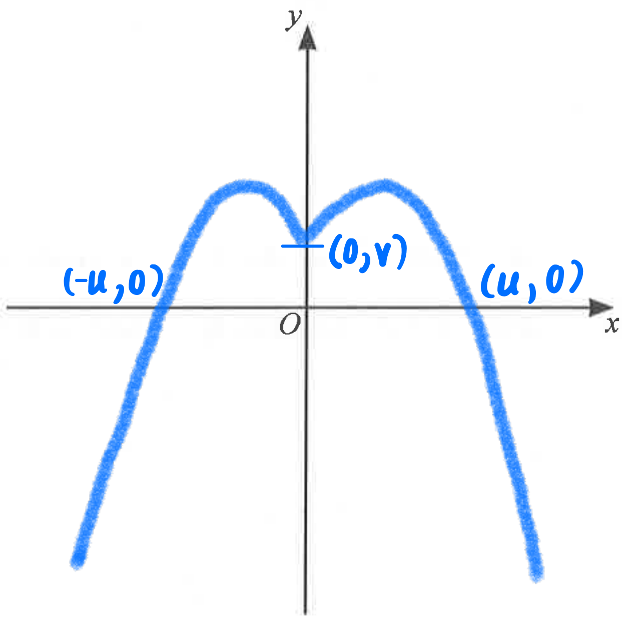


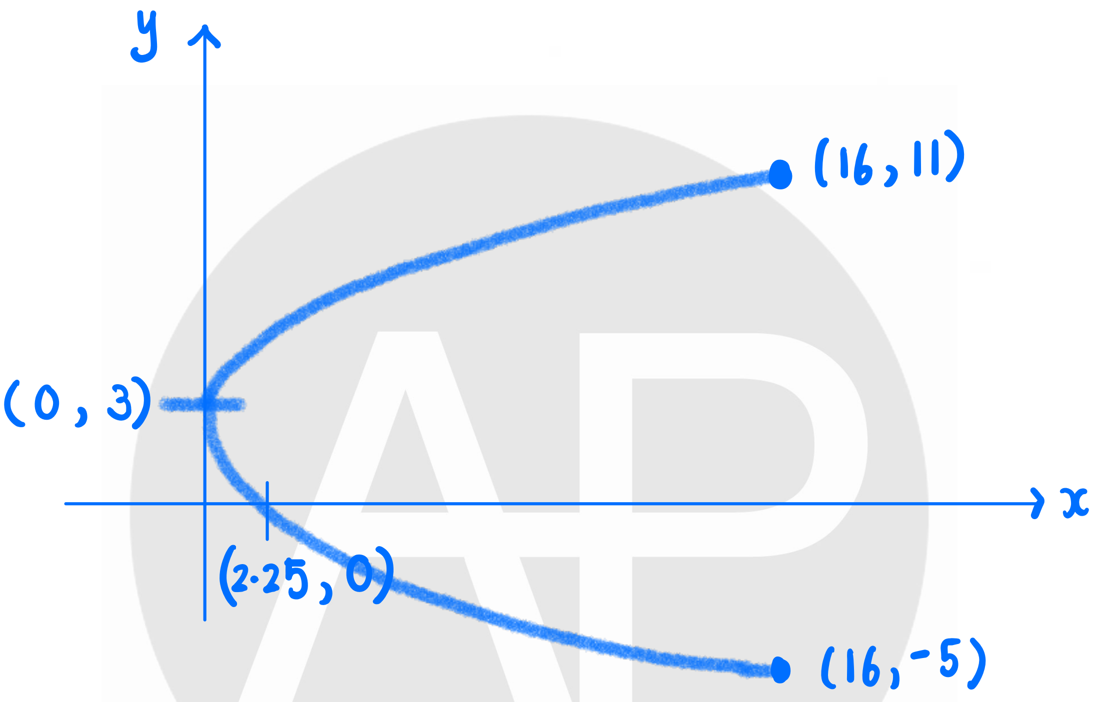
Fig. 4.2

5 [Continued]

(c) The curve C has parametric equations

$$x = t^2, \qquad y = 2t + 3, \qquad \text{for } -4 \leqslant t \leqslant 4.$$

(i) Sketch the graph of C. [1]



Let $y = 0$: $2t + 3 = 0$

$t = -1.5$

$\therefore \ x = (-1.5)^2$

$= 2.25$

Let $t = -4$: $x = 16, \ y = -5$

Let $t = 4$: $x = 16, \ y = 11$

(ii) Mark on your sketch the coordinates of the points where the curve C meets the axes. [2]

(iii) State the equations of any lines of symmetry. [1]

$y = 3$

The suggested solutions are prepared by Mr Alvin Yeo. Mr Yeo will hold no liability for any errors.

BLANK PAGE

Section B: Probability and Statistics [60 marks]

6 A sports club has three categories of membership. There are 27 Youth members, 45 Adult members and 18 Senior members.

The club secretary wishes to find out the opinions of members of the club about the facilities offered. She gives a questionnaire to all the members of the club. She receives replies from 65 of the members.

(a) Explain whether the 65 members comprise a sample or a population. [1]

There are a total of 90 members in the sports club, forming the population.
Hence, the 65 members comprise of a sample.

The secretary decides to form a committee of members to discuss the results of her questionnaire.

(b) Explain an advantage of choosing a random sample of members for her committee. [1]

The discussion will be unbiased.

The secretary decides that the committee will consist of 6 randomly chosen members, but there will be:

- at least 1 member from each category
- more Adult members than Youth members and
- more Adult members than Senior members.

"Given condition"

(c) Find the probability that the committee contains more Youth members than Senior members. [4]

$$\frac{P(Y,Y,S,A,A,A)}{P(Y,S,A,A,A,A) + P(Y,Y,S,A,A,A) + P(Y,S,S,A,A,A)}$$

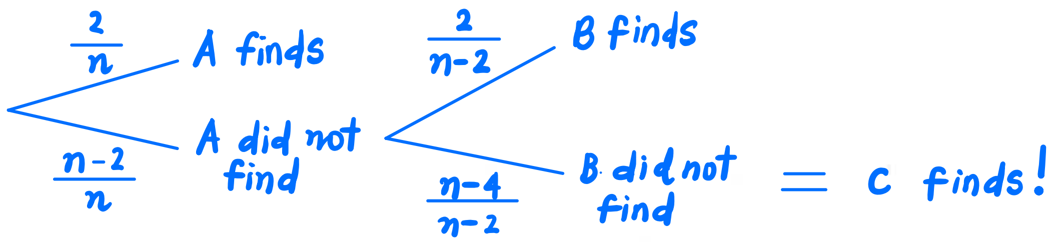
$$= \frac{({}^{27}C_2 \times {}^{18}C_1 \times {}^{45}C_3)}{({}^{27}C_1 \times {}^{18}C_1 \times {}^{45}C_4) + ({}^{27}C_2 \times {}^{18}C_1 \times {}^{45}C_3) + ({}^{27}C_1 \times {}^{18}C_2 \times {}^{45}C_3)}$$

$$= \frac{89\,652\,420}{72\,411\,570 + 89\,652\,420 + 58\,618\,890}$$

$$= \frac{13}{32} \text{ or } 0.40625$$

The suggested solutions are prepared by Mr Alvin Yeo. Mr Yeo will hold no liability for any errors.

Using a Tree Diagram is easier for solving this question:



7 **(a)** A, B and C are taking part in a game in which a prize is hidden in one of n identical closed boxes, where $n > 4$. First, A opens at random two boxes in an attempt to find the prize. If A fails to find the prize, B then opens at random two of the remaining boxes in an attempt to find the prize. If B fails to find the prize, C opens all the remaining boxes and finds the prize.

(i) Find, in terms of n, the probability that A finds the prize. [1]

$\underline{\frac{2}{n}}$

(ii) Show that the probability that B finds the prize is $\frac{2}{n}$. [1]

$$\underbrace{\left(1-\frac{2}{n}\right)}_{\text{Prob. that A did not find the prize}} \times \left(\frac{2}{n-2}\right) = \frac{n-2}{n} \times \frac{2}{n-2}$$

$$= \frac{2}{n}$$

(iii) Find the range of values of n for which the probability that C finds the prize is more than 8 times the probability that either A or B find the prize. [3]

Prob. C finds the prize $= \left(\frac{n-2}{n}\right)\left(\frac{n-4}{n-2}\right) = \frac{n-4}{n}$

Given $\left(\frac{n-4}{n}\right) > 8\left(\frac{2}{n} + \frac{2}{n}\right)$

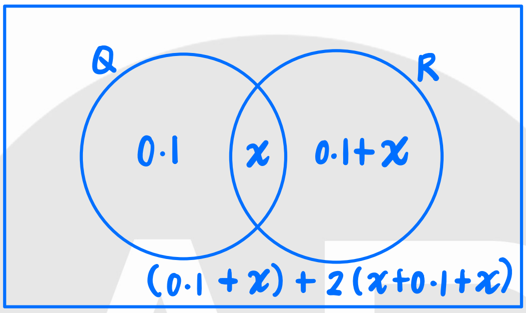
By GC: $n > 36$

i.e. $\underline{n \geq 37,\ n \in \mathbb{Z}^+}$.

The suggested solutions are prepared by Mr Alvin Yeo. Mr Yeo will hold no liability for any errors.

7 **[Continued]**

(b) The events Q and R are such that $P(Q \cap R) = x$, $P(Q \cap R') = 0.1$, $P(Q' \cap R) = P(Q)$ and $P((Q \cup R)') = P(Q) + 2P(R)$. By using a Venn diagram, or otherwise, find the exact value of x. [3]



$$0.1 + x + (0.1+x) + (0.1+x) + 2(0.1+2x) = 1$$

$$7x + 0.5 = 1$$

$$7x = 0.5$$

$$x = \frac{1}{14}$$

8 Lee has a bird feeder on his balcony. Every morning he spends 20 minutes, while he eats his breakfast, counting the number of birds that visit the feeder. Over many months he has found that the mean number of birds visiting the feeder in this time interval is 17.3. When the bird feeder was damaged in a storm, Lee replaced it with a new bird feeder.

He suspects that the mean number of birds visiting the new bird feeder while he eats his breakfast has reduced. He decides to check this with a hypothesis test at the $\alpha\%$ level of significance, where α is an integer. He records the number of birds, x, visiting the feeder in 20 minutes each morning for a random sample of n mornings.

(a) Explain whether Lee should carry out a one-tailed test or a two-tailed test. [1]

One-tailed test, as Lee wants to check if the number of birds visiting his bird feeder has reduced.

(b) State hypotheses for Lee's test, defining any parameters that you use. [2]

Let X be the r.v. denoting the number of birds visiting the bird feeder during 20 mins of Lee's breakfast time.

To test $H_0: \mu = 17.3$ against

$H_1: \mu < 17.3$

where H_0 is the null hypothesis,

H_1 is the alternate hypothesis and

μ is the population mean.

Here is a summary of the data Lee collected.

$$n = 32 \qquad \sum x = 512 \qquad \sum x^2 = 8702$$

Lee carried out his test and concluded that the null hypothesis should be rejected.

(c) Calculate unbiased estimates of the population mean and variance of the number of birds, and determine the minimum possible value of the integer α. [3]

$$\bar{x} = \frac{\sum x}{n} = \frac{512}{32} = \underline{16}$$

$$s^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{(\sum x)^2}{n}\right] = \frac{1}{31}\left(8702 - \frac{512^2}{32}\right)$$

$$= \underline{\frac{510}{31}}$$

H_0 is rejected if p-value $\leqslant \frac{\alpha}{100}$

By GC, $p = 0.034910608 6$

$\therefore \alpha \geqslant 3.49$

$\underline{\alpha_{min} = 4}$

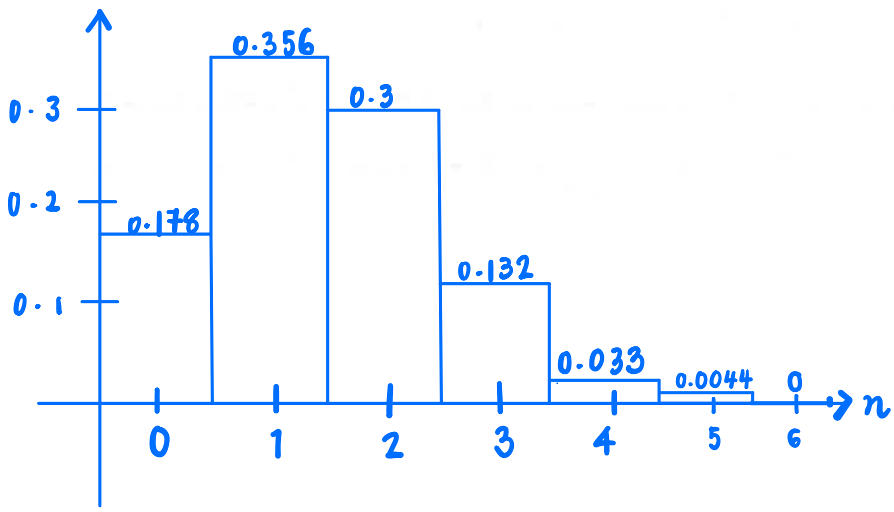
(d) State the conclusion to Lee's test in the context of the question. [1]

There is sufficient evidence at α% to reject H_0, i.e. the number of birds visiting his new bird feeder during his morning breakfast 10-min interval has indeed reduced.

(e) Explain whether using the number of birds for each of 10 mornings would have been sufficient for Lee to carry out his hypothesis test. [1]

It would not be sufficient as Central Limit Theorem cannot be applied to approximate the distribution to be normal.

9 (a) The random variable X has the distribution B(6, 0.25). Sketch the distribution of X. [2]



(b) One of the jobs of a quality control operative in a food-processing factory is to check the quality of a random sample of 50 meat pies from the production line each day. She records the number of pies found to be 'unsatisfactory' each day for 80 days. Her results are shown in the table below.

Number unsatisfactory	0	1	2	3	4	5	6 or more
Frequency	57	9	6	4	3	1	0

Use the information in the table to estimate the probability that a randomly chosen pie is unsatisfactory. [2]

$$\frac{(1\times 9)+(2\times 6)+(3\times 4)+(4\times 3)+(5\times 1)}{50\times 80}$$

$$= \frac{1}{80} \text{ or } 0.0125$$

(c) One of the products made in the food-processing factory is the 'Frozen Cheesee Burger'. A fixed number of the burgers are tested each day and the number found to have insufficient cheese in them is denoted by Y.

(i) State, in context, two assumptions needed for Y to be well modelled by a binomial distribution. [2]

1. The probability of "success" defined as the probability for each burger to have insufficient cheese, remains as a constant.

2. It is assumed that there are 2 mutually exclusive outcomes: with insufficient cheese or with sufficient cheese.

Assume now that Y has the distribution $B(120, 0.03)$.

(ii) Find the probability that, on a randomly chosen day, fewer than 3 burgers are found to have insufficient cheese in them. [1]

$P(Y<3) = P(Y \leq 2) = 0.29844$

$= \underline{0.298}$

(iii) For a randomly chosen period of 28 days, find the expectation of the number of days on which fewer than 3 burgers have insufficient cheese in them. [2]

Let W be the r.v. denoting no. of days out of 28, where there are fewer than 3 burgers with insufficient cheese in them per day.

$W \sim B(28, 0.29844)$

$E(W) = 28 \times 0.29844$

$= 8.3562$

$= \underline{8.36 \text{ days}}$

(iv) Find the probability that, in a randomly chosen period of 28 days, more than 100 burgers are found to have insufficient cheese in them. [2]

Let T be the r.v. denoting no. of burgers out of 3360, with insufficient cheese in them.

$T \sim B(3360, 0.03)$

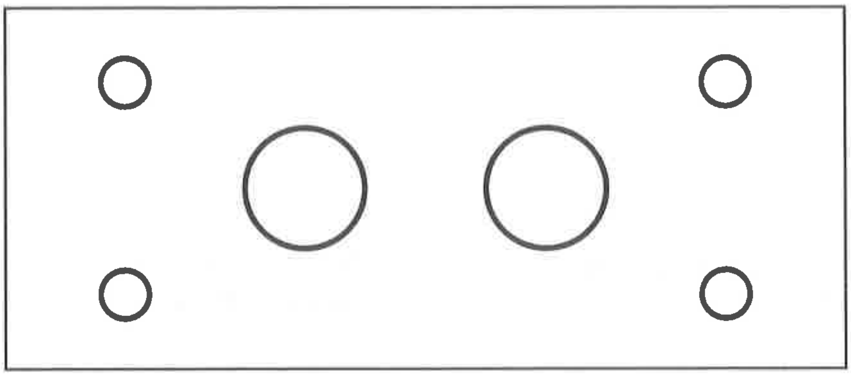
$P(T>100) = 1 - P(T \leq 100)$

$= 0.50578$

$= \underline{0.506}$

10 In this question you should state the parameters of any distributions you use.

A company makes metal plates that can be used to fix fence panels onto posts. The metal plates have four small holes drilled into them for screws and two large holes drilled into them for bolts (see diagram).



Before the holes are drilled, the masses of plates, in grams, follow the distribution $N(200, 1.6^2)$.

(a) Find the probability that the mass of a randomly chosen plate before drilling is more than 197.5 grams. [1]

Let X be the r.v. denoting the mass of a chosen plate (in grams) before drilling.

$X \sim N(200, 1.6^2)$

$P(X > 197.5) = 0.94091 = \underline{0.941}$

Drilling the holes reduces the mass of each plate by 5%. A production worker selects 8 of the drilled plates at random.

(b) Find the probability that at least 5 of these 8 plates have masses between 190 grams and 192 grams. [4]

Let Y be the r.v. denoting the mass of a chosen plate (in grams) after drilling.

$Y = 0.95X$

$E(Y) = 0.95 \times 200 = 190$

$Var(Y) = 0.95^2 \times 1.6^2 = 2.3104$

$\therefore Y \sim N(190, 2.3104)$

$P(190 < Y < 192) = 0.40588$

Let W be the r.v. denoting no. of plates having masses between 190g and 192g out of 8 plates.

$W \sim B(8, 0.40588)$

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

$= 0.18291 = \underline{0.183}$

The drilled plates are sold in packs of 20 randomly chosen plates.

(c) Find the probability that the total mass of a pack of 20 drilled plates is less than 3805 grams. [3]

Let T be the r.v. denoting total mass of a pack of 20 drilled plates (in gram).

$$T = Y_1 + Y_2 + \cdots + Y_{20} \sim N(3800, 46.208)$$

$$P(T < 3805) = 0.768997$$
$$= \underline{0.769}$$

The manufacturer decides to sell 'Value' packs containing all the materials needed to fix fence panels onto posts. Each Value pack consists of 20 drilled plates together with the right number of screws and bolts to fit them; all of these are randomly chosen.

- The masses of screws, in grams, follow the distribution $N(10, 0.3^2)$.
- The masses of bolts, in grams, follow the distribution $N(44, 1.1^2)$.

(d) Find the mass exceeded by just 5% of the Value packs. Give your answer to the nearest gram. You should ignore the mass of any packaging. [4]

Let V be the r.v. denoting the mass of a value pack (in grams).
Let S be the r.v. denoting the mass of a screw (in grams).
Let B be the r.v. denoting the mass of a bolt (in grams).

$$V = T + S_1 + S_2 + \cdots + S_{80} + B_1 + B_2 + \cdots + B_{40}$$
$$E(V) = 3800 + 80(10) + 40(44) = 6360$$
$$Var(V) = 46.208 + 80(0.3^2) + 40(1.1^2) = 101.808$$

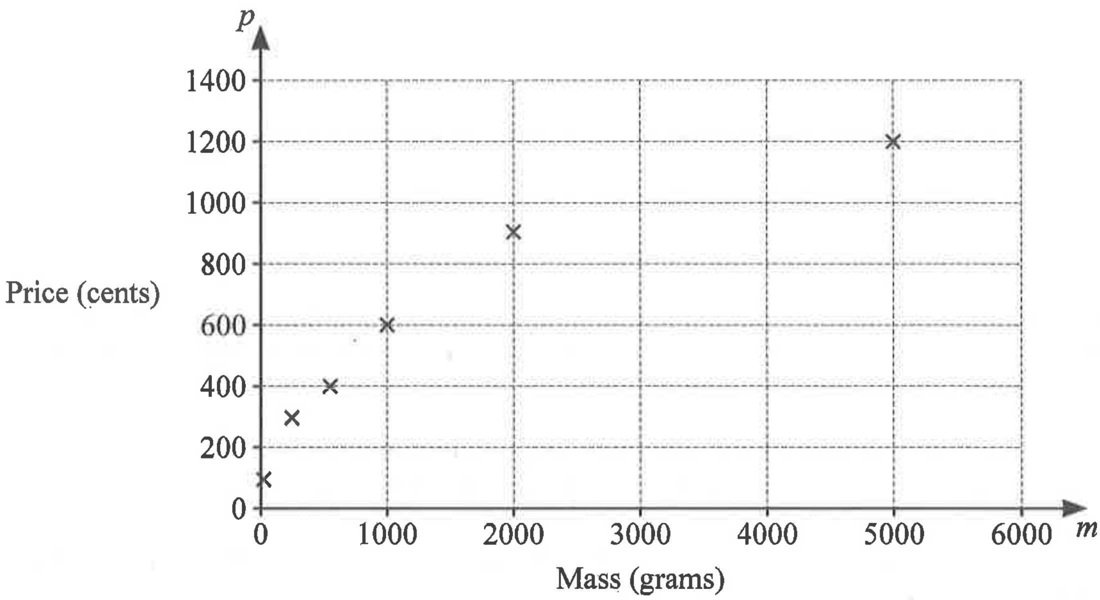
$$\therefore V \sim N(6360, 101.808)$$

$$P(V > m) = 0.05$$
$$1 - P(V \leq m) = 0.05$$
$$P(V \leq m) = 0.95$$
$$m = 6376.596565$$
$$= \underline{6377\text{g}}$$

11 (a) A certain brand of breakfast cereal is sold in a variety of different sized packs. Details of the mass of cereal in each pack, m grams, and the price, p cents, are given in the table below.

m	24	250	550	1000	2000	5000
p	100	300	400	600	900	1200

A scatter diagram for the data is shown below.



(i) Explain what the scatter diagram tells you about the relationship between m and p. [1]

As the mass of the breakfast cereal packs increase, the price increases at a decreasing rate.

THE ANNEXE PROJECT
EDUCATIONAL CENTRE
ESTD 2008

The suggested solutions are prepared by Mr Alvin Yeo. Mr Yeo will hold no liability for any errors.

The following two models are proposed, where a, b, d, and e are constants.

$$p = a + b\ln m \qquad p = d + e\sqrt{m}$$

(ii) Determine which of these models gives the better fit to the data. State the values of the constants and the product moment correlation coefficient in this case. [4]

By GC: for $p = a + b\ln m$, $r = 0.92851$,
$a = -703.84$
$b = 203.05$

i.e. $p = -703.84 + 203.05\ln m$.

for $p = d + e\sqrt{m}$, $r = 0.99144$,
$d = 32.953$
$e = 17.270$

i.e. $p = 32.953 + 17.270\sqrt{m}$.

$p = 32.953 + 17.270\sqrt{m}$ is the better model because r-value is closer to 1.

(iii) A new pack containing 750 grams of cereal is introduced. Use the model you identified in part **(a)(ii)** to estimate the price of this pack, correct to the nearest 10 cents. Explain whether your answer is reliable. [2]

$p = 32.953 + 17.270 \times \sqrt{750}$
$= 505.91$
$= \underline{510\ ¢}$ (nearest 10 cents)

The estimation is <u>reliable</u> as <u>interpolation</u> was practised. ($m = 750$g is within range of data collected)

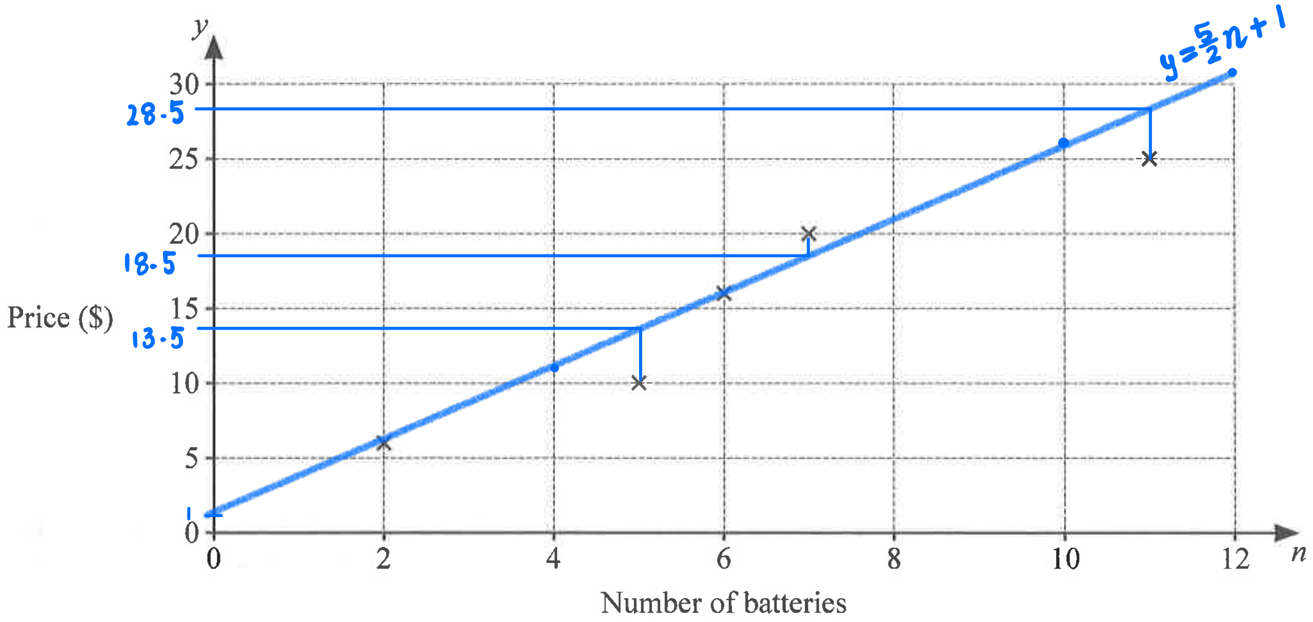
The suggested solutions are prepared by Mr Alvin Yeo. Mr Yeo will hold no liability for any errors.

11 [Continued]

(b) Kai is investigating the relationship between the number, n, of a particular type of high-performance batteries in a pack and the price, $\$y$, of the pack she found in different stores. Her results are shown in the table below.

n	2	5	6	7	11
y	6	10	16	20	25

This information is shown in the scatter diagram below.



Kai decides to investigate whether the model $y = \frac{5}{2}n + 1$ is a good fit for this data.

(i) Draw the line $y = \frac{5}{2}n + 1$ on the scatter diagram above.

[1]

(ii) For the model $y = f(n)$, the residual for a point (a, b) is $b - f(a)$.

(A) Mark the residuals for the points on the scatter diagram above. [1]

(B) Explain why Kai should use the sum of the squares of the residuals rather than the sum of the residuals when assessing the fit of the model. [1]

When adding the sum of the residuals, the positive and negative residuals will cancel out each other. Hence squaring the residuals convert all measurements to positive.

(C) Calculate the sum of the squares of the residuals for the line $y = \frac{5}{2}n + 1$. [1]

$(10 - 13.5)^2 + (20 - 18.5)^2 + (25 - 28.5)^2 = \underline{26.75}$

Kai's friend points out that other lines parallel to Kai's line can be drawn which are a better fit for the data.

(iii) Explain how the sum of the squares of the residuals for a line that is a better fit for the data differs from the sum of the squares of the residuals for the line $y=\frac{5}{2}n+1$ found in part **(ii)(C)**. [1]

The sum of the squares of the residuals for a line that is a better fit should be less than 26.75.

(iv) Find the range of values of c for which the line $y=\frac{5}{2}n+c$ is a better fit for the data than the line $y=\frac{5}{2}n+1$. [3]

the sum of the squares of the residuals between f(a) and the new line $y=\frac{5}{2}n+c$.

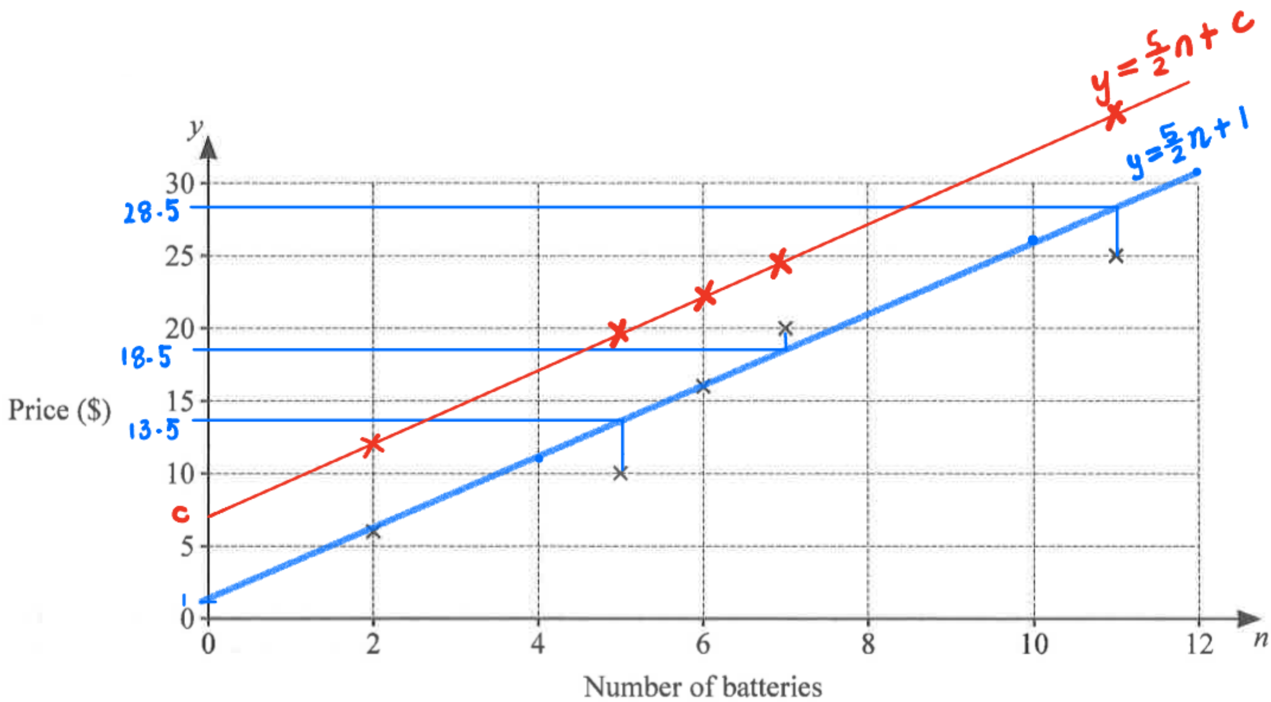
$\boxed{(c-1)^2+(c-1+3.5)^2+(c-1)^2+(c-1-1.5)^2+(c-1+3.5)^2} < 26.75$

$(c^2-2c+1)+(c^2+5c+6.25)+(c^2-2c+1)+(c^2-5c+6.25)+(c^2+5c+6.25)-26.75<0$

$$5c^2+c-6<0$$

$$(5c+6)(c-1)<0$$

$$\therefore -\frac{6}{5}<c<1$$



BLANK PAGE