TEN-YEAR SERIES · WORKED SOLUTIONS

2024 A-Level H2 Mathematics Paper 2

11 questions · 100 marks

The question paper is © UCLES/SEAB and is not reproduced here — have your own copy of the paper open as you work through these solutions. The worked solutions and diagrams are original work by The Learning Cave and are for personal study only.

thelearningcave.com.sg

Physics. Math. Chemistry.

Question 1 6 marks

Complex numbers · conjugates · roots

Part (a) 2 marks

Since $z = -2 + i$, we have $z^* = -2 - i$. Substituting into $az^* + b$,

$$\begin{aligned} az^* + b &= a(-2 - i) + b \\ &= (-2a + b) - ai \end{aligned}$$

Equating this to $z = -2 + i$,

Real: $-2a + b = -2$

Imaginary: $-a = 1 \Rightarrow a = -1$

Substituting into the real equation: $-2(-1) + b = -2 \Rightarrow b = -4$.

To check your answer: $(-1)(-2 - i) + (-4) = 2 + i - 4 = -2 + i = z$ ✓

$\therefore a = -1, b = -4.$

Part (b) 4 marks

APPROACH

We compute wz and $\frac{w}{z}$ separately, each by hand without a calculator, then subtract.

$$\begin{aligned} wz &= (1 - 3i)(-2 + i) \\ &= -2 + i + 6i - 3i^2 \\ &= -2 + 7i + 3 \\ &= 1 + 7i \end{aligned}$$

To divide, we multiply numerator and denominator by the conjugate of z:

$$\begin{aligned} \frac{w}{z} &= \frac{1 - 3i}{-2 + i} \times \frac{-2 - i}{-2 - i} \\ &= \frac{(1 - 3i)(-2 - i)}{(-2)^2 - i^2} \\ &= \frac{-2 - i + 6i + 3i^2}{4 + 1} \\ &= \frac{-5 + 5i}{5} \\ &= -1 + i \end{aligned}$$

Subtracting,

$$\begin{aligned} wz - \frac{w}{z} &= (1 + 7i) - (-1 + i) \\ &= 2 + 6i \end{aligned}$$

THE LEARNING CAVE
thelearningcave.com.sg
Physics. Math. Chemistry.

$\therefore wz - \frac{w}{z} = 2 + 6i$, so $c = 2$, $d = 6$.

Question 2 6 marks

Differentiation · maximum area

APPROACH

With two side-lengths a and b related by a fixed perimeter, we express the area as a function of a alone, then use calculus to maximise it.

Let $AE = a$ and $ED = b$. Since ABDE is a rectangle, $BD = AE = a$ and $AB = ED = b$. Since triangle BCD is equilateral, $BC = CD = BD = a$.

Using the perimeter of the pentagon,

$$\begin{aligned} AB + BC + CD + DE + EA &= 20 \\ &= b + a + a + b + a \\ &= 3a + 2b \\ \Rightarrow b &= \frac{20 - 3a}{2} \quad ...(1) \end{aligned}$$

The area is the rectangle ABDE plus the equilateral triangle BCD:

$$A = ab + \frac{\sqrt{3}}{4}a^2$$

Substituting (1),

$$\begin{aligned} A &= a\left(\frac{20 - 3a}{2}\right) + \frac{\sqrt{3}}{4}a^2 \\ &= 10a - \frac{3}{2}a^2 + \frac{\sqrt{3}}{4}a^2 \\ &= 10a + \left(\frac{\sqrt{3} - 6}{4}\right)a^2 \end{aligned}$$

Differentiating with respect to a,

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

Setting $\frac{dA}{da} = 0$,

$$\begin{aligned} 10 &= \left(\frac{6 - \sqrt{3}}{2}\right)a \\ \Rightarrow a &= \frac{20}{6 - \sqrt{3}} \\ &= \frac{20\left(6 + \sqrt{3}\right)}{33} \\ &\approx 4.686 \end{aligned}$$

Since $\frac{d^2A}{da^2} = \frac{\sqrt{3}-6}{2} < 0$ (as $6 > \sqrt{3}$) for all a, A is a maximum at this value of a.

From (1), $b \approx \frac{20-3(4.686)}{2} \approx 2.971 > 0$, so this value of a lies in the valid range $0 < a < \frac{20}{3}$.

The maximum area is

$$\begin{aligned} A_{max} &= ab + \frac{\sqrt{3}}{4}a^2 \\ &= \frac{100}{6-\sqrt{3}} \\ &= \frac{100\left(6+\sqrt{3}\right)}{33} \\ &\approx 23.4305 \end{aligned}$$

To check your answer: substituting $a \approx 4.686$, $b \approx 2.971$ directly into $A = ab + \frac{\sqrt{3}}{4}a^2$ gives $\approx 13.923 + 9.508 = 23.431$, matching the exact-form value above ✓

$\therefore$ the maximum possible area of the flower bed is $\frac{100\left(6+\sqrt{3}\right)}{33} \approx 23.43\ \text{m}^2$ (4 s.f.).

Question 3 8 marks

Functions · range · inverse

Part (a) 2 marks

Completing the square,

$$\begin{aligned} f(x) &= -x^2 + 4x + 7 \\ &= -\left(x^2 - 4x\right) + 7 \\ &= -(x-2)^2 + 4 + 7 \\ &= -(x-2)^2 + 11 \end{aligned}$$

Since $(x-2)^2 \geq 0$ for all real x, $-(x-2)^2 \leq 0$, so $f(x) \leq 11$.

$\therefore R_f = (-\infty,\ 11]$.

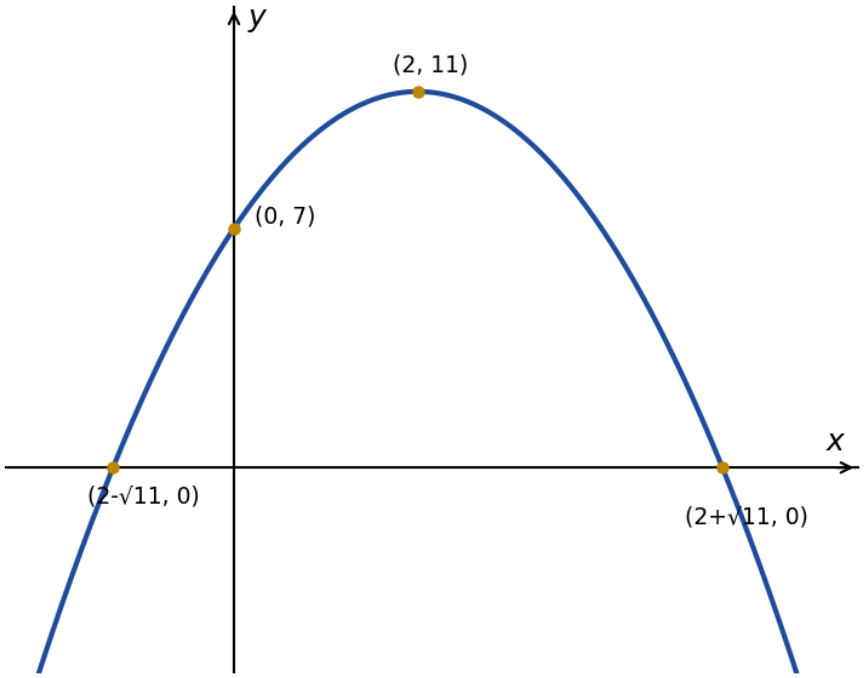
Part (b) 2 marks

y-intercept: $f(0) = 7$, giving $(0, 7)$.

x-intercepts: solving $f(x) = 0$,

$$\begin{aligned} -x^2 + 4x + 7 &= 0 \\ \Rightarrow x^2 - 4x - 7 &= 0 \\ \Rightarrow x &= \frac{4 \pm \sqrt{16+28}}{2} \\ &= 2 \pm \sqrt{11} \end{aligned}$$

giving $\left(2-\sqrt{11}, 0\right)$ and $\left(2+\sqrt{11}, 0\right)$, with maximum turning point $(2, 11)$ from part (a).



Part (c) 1 mark

thelearningcave.com.sg

APPROACH

A function has an inverse exactly when it is one-one on its domain; for a function given by a formula, it is enough to show it is strictly monotonic there.

$$g(x) = \frac{2}{x} + 1, \quad 1 \leq x \leq 10$$

$$\Rightarrow g'(x) = -\frac{2}{x^2}$$

Since $x^2 > 0$ on $[1, 10]$, $g'(x) < 0$ throughout, so g is strictly decreasing on $[1, 10]$.

$\therefore g$ is one-one on $[1, 10]$, so g^{-1} exists.

Part (d) 3 marks

APPROACH

$fg^{-1}(x)$ means $f\left(g^{-1}(x)\right)$: we first find g^{-1} and its domain, then substitute $x = 1.5$ into g^{-1}, then apply f to the result.

Let $y = g(x) = \frac{2}{x} + 1$ and solve for x in terms of y:

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

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

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

Physics. Math. Chemistry.

Replacing y by x,

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

Since g is strictly decreasing on $[1, 10]$ (part (c)), its range runs from $g(10) = 1.2$ up to $g(1) = 3$, so $D_{g^{-1}} = R_g = [1.2,\ 3]$. This contains $x = 1.5$, so $g^{-1}(1.5)$ is defined.

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

Applying f,

$$\begin{aligned} fg^{-1}(1.5) &= f(4) \\ &= -(4)^2 + 4(4) + 7 \\ &= 7 \end{aligned}$$

To check your answer: $g^{-1}(1.5) = 4$ lies in $[1, 10]$, the domain of g (and hence the range of g^{-1}), and $g(4) = \frac{2}{4} + 1 = 1.5$ confirms $g^{-1}(1.5) = 4$ ✓

$\therefore fg^{-1}(1.5) = 7.$

Question 4 9 marks

Sequences and series · nth term · sums

Part (a) 2 marks

Since $S_n = t_1 + t_2 + \cdots + t_n$, we have $t_n = S_n - S_{n-1}$ for $n \geq 2$, with $t_1 = S_1$ checked separately.

$$\begin{aligned} S_{n-1} &= 2(n-1)^3 - 8(n-1)^2 - 4(n-1) \\ &= 2\left(n^3 - 3n^2 + 3n - 1\right) - 8\left(n^2 - 2n + 1\right) - 4(n-1) \\ &= 2n^3 - 14n^2 + 18n - 6 \end{aligned}$$

Subtracting,

$$\begin{aligned} t_n &= S_n - S_{n-1} \\ &= \left(2n^3 - 8n^2 - 4n\right) - \left(2n^3 - 14n^2 + 18n - 6\right) \\ &= 6n^2 - 22n + 6 \end{aligned}$$

Note: this formula also holds at $n = 1$, since $S_1 = 2 - 8 - 4 = -10$ agrees with $6 - 22 + 6 = -10$.

To check your answer: $t_1 = 6(1)^2 - 22(1) + 6 = -10$, matching $S_1 = 2(1)^3 - 8(1)^2 - 4(1) = -10$ ✓

$\therefore t_n = 6n^2 - 22n + 6.$

Part (b) 2 marks

$$\begin{aligned} 50n - 204 &= 6n^2 - 22n + 6 \\ \Rightarrow 6n^2 - 72n + 210 &= 0 \end{aligned}$$

$$\Rightarrow n^2 - 12n + 35 = 0$$
$$\Rightarrow (n-5)(n-7) = 0$$
$$\Rightarrow n = 5 \text{ or } n = 7$$

$\therefore n = 5$ or $n = 7$.

Part (c) 2 marks

APPROACH

U has common difference 50 and V has common difference 3, so a number common to both progressions recurs with period $lcm(50, 3) = 150$. Rather than listing terms of each series until one matches, we find one common term algebraically using modular arithmetic, restricting the search to terms exceeding 100.

Terms of U exceed 100 when

$$50n - 204 > 100$$
$$\Rightarrow n > 6.08$$
$$\Rightarrow n \geq 7,$$

giving $u_7 = 146,\ u_8 = 196,\ u_9 = 246, \ldots$

For a term to lie in both series, we need $50a - 204 = 3b + 16$ for positive integers a, b:

$$50a - 204 = 3b + 16$$
$$\Rightarrow 50a - 220 = 3b$$

Taking both sides mod 3 (using $50 \equiv 2$ and $220 \equiv 1$),

$$2a \equiv 1 \pmod 3$$
$$\Rightarrow a \equiv 2 \pmod 3,$$

since 2 is its own inverse mod 3. So $a = 2, 5, 8, 11, \ldots$. The smallest such a with $a \geq 7$ (so that $u_a > 100$) is $a = 8$:

$$u_8 = 50(8) - 204 = 196$$

To check your answer: $196 = 3b + 16 \Rightarrow b = 60$, an integer, so $196 = v_{60}$ ✓

$\therefore$ the smallest number greater than 100 in both series is 196.

Part (d)(i) 2 marks

APPROACH

since w_n depends on the parity of n, we check n even and n odd separately, writing $n = 2k$ or $n = 2k + 1$.

If $n = 2k$,

$$w_n = 3(2k)^2 - 5(2k) + 7$$
$$= 12k^2 - 10k + 7$$

$12k^2$ and $10k$ are both even, so $w_n = even - even + odd = odd$.

If $n = 2k + 1$,

$$w_n = 3(2k+1)^2 - 5(2k+1) + 7$$
$$= 3\left(4k^2 + 4k + 1\right) - 5(2k+1) + 7$$
$$= 12k^2 + 2k + 5$$

$12k^2$ and $2k$ are both even, so $w_n = even + even + odd = odd$.

∴ in both cases w_n is odd, so every term of series W is odd. (shown)

Part (d)(ii) 1 mark

$$u_n = 50n - 204$$

$50n$ is even for every integer n, and 204 is even, so $u_n = even - even = even$ for every n — every term of U is even.

Since every term of W is odd (part (i)) and every term of U is even, no term of W can equal a term of U.

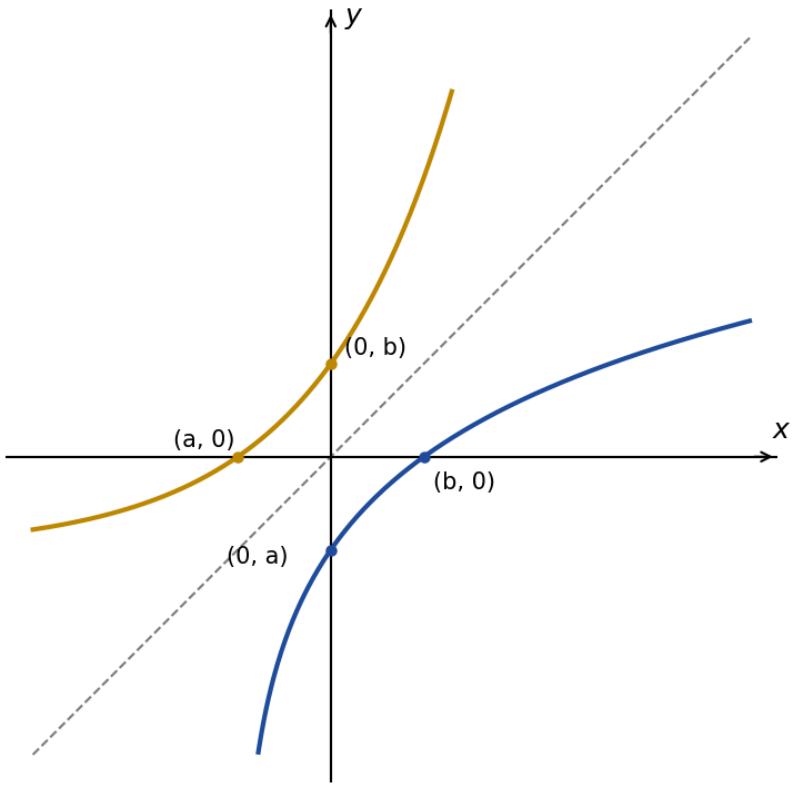
∴ series U and series W have no terms in common.

Question 5 11 marks

Graphs and transformations · sketches of related graphs

Part (a)(i) 1 mark

Since $y = f(x)$ meets the axes at $(a, 0)$ and $(0, b)$, and $y = f^{-1}(x)$ is the reflection of $y = f(x)$ in the line $y = x$, each labelled point swaps its coordinates: $(a, 0)$ maps to $(0, a)$, and $(0, b)$ maps to $(b, 0)$.



$\therefore y = f^{-1}(x)$ is the reflection of $y = f(x)$ in $y = x$, cutting the axes at $(0, a)$ and $(b, 0)$.

Part (a)(ii) 2 marks

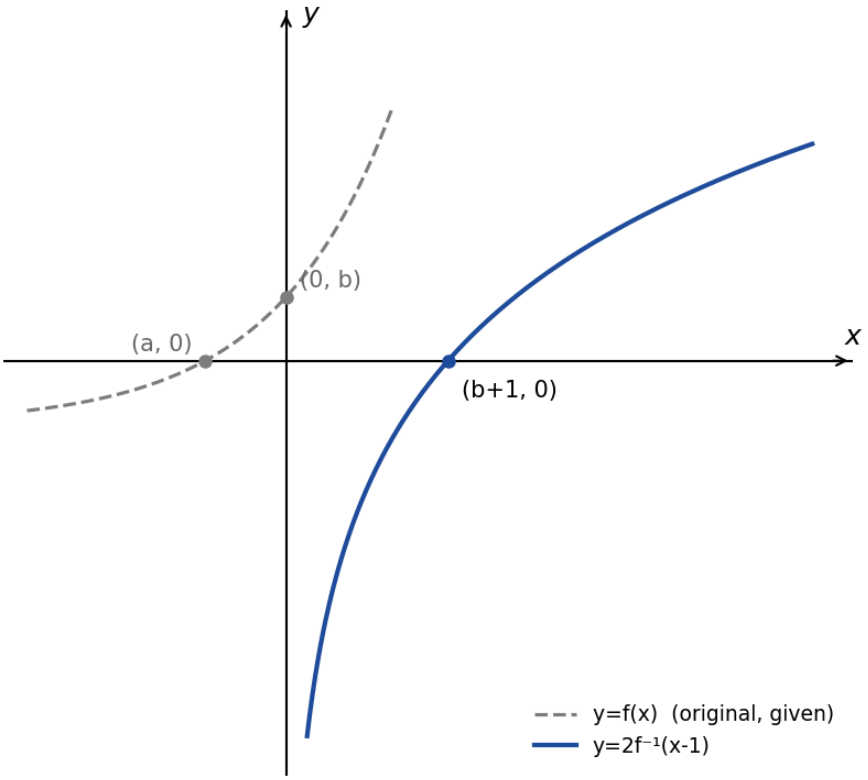
APPROACH

$y = 2f^{-1}(x - 1)$ is built from $y = f^{-1}(x)$ by two transformations: replace x with $x - 1$ (translation $+1$ in the x-direction), then double the resulting y-values (stretch, scale factor 2, parallel to the y-axis). Only the x-intercept is required, and a y-stretch does not move a point already on the x-axis, so we only need to track how the translation moves the x-intercept of $y = f^{-1}(x)$.

From part (i), $y = f^{-1}(x)$ has x-intercept $(b, 0)$. Setting $y = 0$ in the transformed curve,

$$2f^{-1}(x - 1) = 0$$
$$\Rightarrow f^{-1}(x - 1) = 0$$
$$\Rightarrow x - 1 = b$$
$$\Rightarrow x = b + 1$$

Physics. Math. Chemistry.

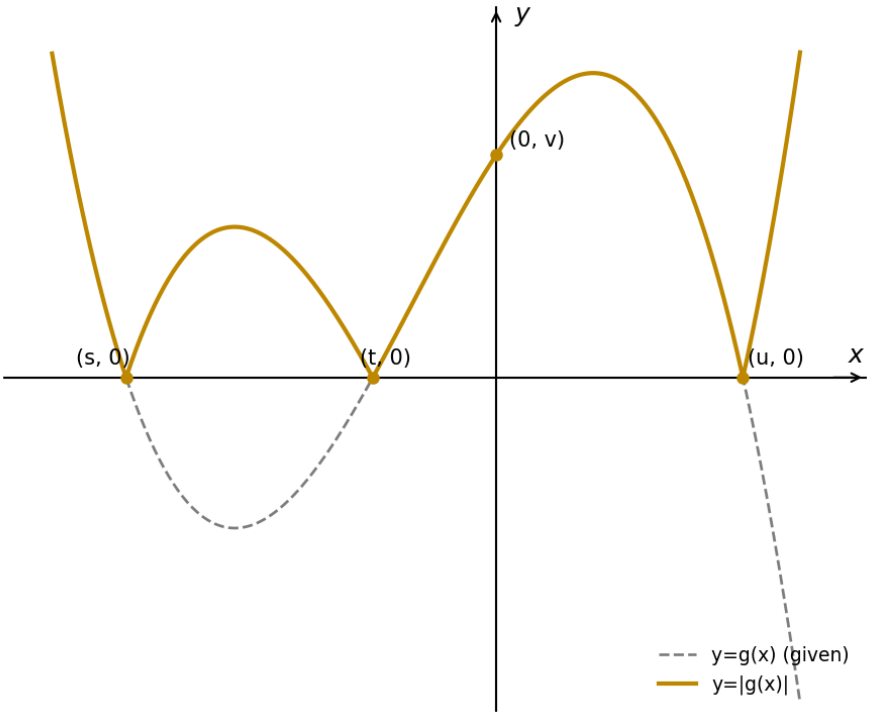


$\therefore y = 2f^{-1}(x-1)$ has x-intercept $(b+1, 0)$.

Part (b)(i) 2 marks

$y = |g(x)|$ agrees with $y = g(x)$ wherever $g(x) \geq 0$, and reflects upward (in the x-axis) the part where $g(x) < 0$. The dip between s and t, where $g(x) < 0$, is reflected up so the curve touches the x-axis at s and t instead of crossing it; the arch from t to u, already non-negative, is unchanged; beyond u, where $g(x)$ falls below the axis, that branch is reflected to rise instead.

Physics. Math. Chemistry.



$\therefore y = |g(x)|$ touches the x-axis at $(s, 0)$ and $(t, 0)$, meets it at $(u, 0)$, and cuts the y-axis at $(0, v)$.

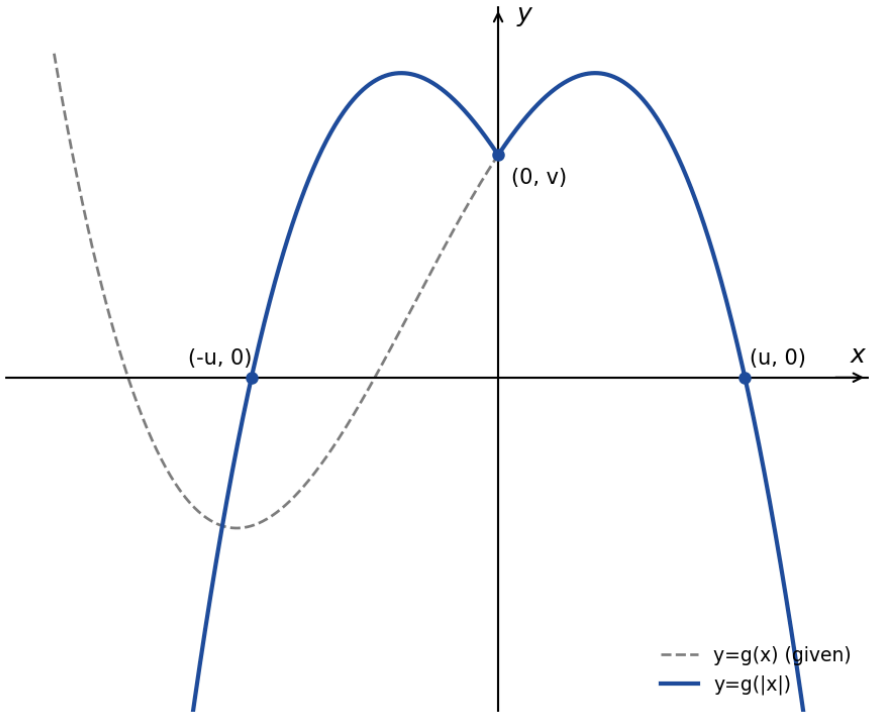
Part (b)(ii) 2 marks

APPROACH

$y = g(|x|)$ is an even function, so the sketch has two parts: for $x \geq 0$ it is identical to $y = g(x)$, and for $x < 0$ it is the mirror image, in the y-axis, of that same $x \geq 0$ branch — the original branch of g for $x < 0$ is discarded entirely.

The right-hand branch of g passes through $(0, v)$, rises to a local maximum, then falls through $(u, 0)$. Reflecting this branch in the y-axis gives a symmetric curve with a local minimum at $(0, v)$, rising to two symmetric maxima, and crossing the x-axis at $(u, 0)$ and $(-u, 0)$.

thelearningcave.com.sg
Physics. Math. Chemistry.



$\therefore y = g(|x|)$ is the even function agreeing with g for $x \geq 0$ and mirrored for $x < 0$, cutting the axes at $(-u, 0)$, $(u, 0)$ and $(0, v)$.

Part (c)(i) and (ii) 3 marks

APPROACH

to sketch a parametric curve we eliminate t to find its Cartesian shape, then use the endpoints of the given range of t and the axis-crossings to fix where the sketch starts, stops, and meets the axes.

Since $y = 2t + 3$,

$$t = \frac{y - 3}{2}$$

Substituting into $x = t^2$,

$$x = \left(\frac{y - 3}{2}\right)^2$$
$$= \frac{(y - 3)^2}{4}$$

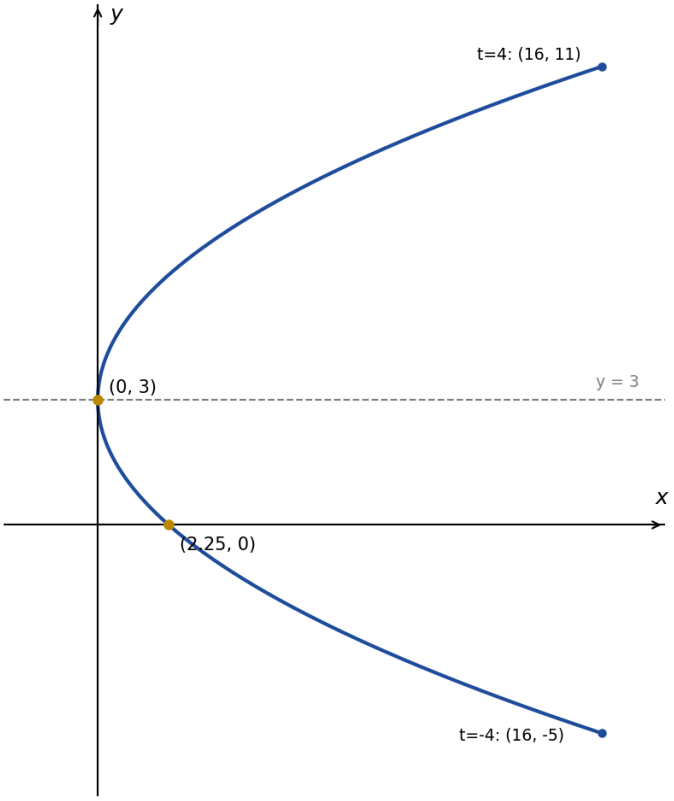
This is a sideways parabola, opening in the positive x-direction, with vertex at $(0, 3)$ (at $t = 0$).

Endpoints: $t = -4$ gives $(16, -5)$; $t = 4$ gives $(16, 11)$.

y-axis: $x = 0 \Rightarrow t^2 = 0 \Rightarrow t = 0 \Rightarrow y = 3$, giving $(0, 3)$ — this is also the vertex, since $t = 0$ is the only value of t with $x = 0$.

x-axis: $y = 0 \Rightarrow 2t + 3 = 0 \Rightarrow t = -1.5$, which lies in $[-4, 4]$, giving $x = (-1.5)^2 = 2.25$, so $(2.25, 0)$.

thelearningcave.com.sg
Physics · Math · Chemistry



$\therefore C$ runs from $(16, -5)$ through $(2.25, 0)$ and the vertex $(0, 3)$ to $(16, 11)$.

Part (c)(iii) 1 mark

The curve is symmetric about the horizontal line through its vertex $(0, 3)$.

$\therefore$ the line of symmetry is $y = 3$.

Question 6 6 marks

Sampling · sampling methods

Part (a) 1 mark

The 65 respondents form a sample, not the population.

Explanation: the population here is all 90 members of the club (27 Youth $+$ 45 Adult $+$ 18 Senior), since the questionnaire was sent to every member. Only 65 of them replied, so these 65 are part — not all — of that group: a self-selected sample of the population, not the population itself.

Part (b) 1 mark

Choosing a random sample of members for the committee gives every member an equal chance of selection, so the committee is not biased by the secretary favouring particular members — it gives a fairer, more representative selection of opinion.

Part (c) 4 marks

APPROACH

let y, a, s be the numbers of Youth, Adult and Senior members on the committee, so $y + a + s = 6$ with $y \geq 1$, $a \geq 1$, $s \geq 1$, $a > y$ and $a > s$. Since a is a strict majority-holder among only 6 members split three ways, we can simply list the few triples satisfying these constraints, rather than set up a general formula.

Listing all triples with $y + a + s = 6$, each at least 1, and $a > y$, $a > s$:

Case 1: $y = 1,\ a = 4,\ s = 1$

Case 2: $y = 1,\ a = 3,\ s = 2$

Case 3: $y = 2,\ a = 3,\ s = 1$

(every other combination fails $a > y$ or $a > s$).

Choosing each committee from the population of 27 Youth, 45 Adult and 18 Senior members,

$$\text{Case 1: } \binom{27}{1}\binom{45}{4}\binom{18}{1} = 27 \times 148\,995 \times 18 = 72\,411\,570$$

$$\text{Case 2: } \binom{27}{1}\binom{45}{3}\binom{18}{2} = 27 \times 14\,190 \times 153 = 58\,618\,890$$

$$\text{Case 3: } \binom{27}{2}\binom{45}{3}\binom{18}{1} = 351 \times 14\,190 \times 18 = 89\,652\,420$$

Adding the three cases, the total number of committees satisfying all the stated constraints is

$$72\,411\,570 + 58\,618\,890 + 89\,652\,420 = 220\,682\,880$$

Of these, only Case 3 has $y > s$ (here $2 > 1$); Case 1 has $y = s$ and Case 2 has $y < s$, so only Case 3 is favourable.

$$P(Y > S) = \frac{89\,652\,420}{220\,682\,880} = \frac{13}{32}$$

To check your answer: $\frac{89\,652\,420}{220\,682\,880}$ simplifies exactly to $\frac{13}{32} = 0.40625$ ✓

$\therefore P(Y > S) = \frac{13}{32} = 0.406$ (3 s.f.).

Question 7 8 marks

Probability · combined events

Part (a)(i) 1 mark

$$P\,(A \text{ finds prize}) = \frac{2}{n}$$

Explanation: A opens 2 of the n boxes at random, and the prize is in exactly one of the n boxes, so the probability that it is among the 2 A opens is $\frac{2}{n}$.

Part (a)(ii) 1 mark

If A fails, the prize is among the remaining $n-2$ boxes, so

$$P(A \text{ fails}) = 1 - \frac{2}{n} = \frac{n-2}{n}$$

Given A fails, B opens 2 of the remaining $n-2$ boxes at random, so

$$P(B \text{ finds prize} \mid A \text{ fails}) = \frac{2}{n-2}$$

Multiplying,

$$\begin{aligned} P(B \text{ finds prize}) &= P(A \text{ fails}) \times P(B \text{ finds prize} \mid A \text{ fails}) \\ &= \frac{n-2}{n} \times \frac{2}{n-2} \\ &= \frac{2}{n} \end{aligned}$$

$\therefore P(B \text{ finds prize}) = \frac{2}{n}$. (shown)

Part (a)(iii) 3 marks

Since A, B, C finding the prize are mutually exclusive and exhaustive,

$$\begin{aligned} P(C \text{ finds prize}) &= 1 - P(A \text{ finds}) - P(B \text{ finds}) \\ &= 1 - \frac{2}{n} - \frac{2}{n} \\ &= \frac{n-4}{n} \end{aligned}$$

The probability that either A or B finds the prize (mutually exclusive events) is

$$P(A \text{ finds}) + P(B \text{ finds}) = \frac{2}{n} + \frac{2}{n} = \frac{4}{n}$$

Requiring $P(C \text{ finds prize})$ to be more than 8 times this,

$$\begin{aligned} \frac{n-4}{n} &> 8 \times \frac{4}{n} \\ &= \frac{32}{n} \end{aligned}$$

Since $n > 4 > 0$, we multiply both sides by n:

$$\begin{aligned} n - 4 &> 32 \\ \Rightarrow n &> 36 \end{aligned}$$

To check your answer: at $n = 37$, $P(C) = \frac{33}{37} \approx 0.892$ and $8 \times \frac{4}{37} = \frac{32}{37} \approx 0.865$; indeed $0.892 > 0.865$ ✓

$\therefore n > 36$.

Part (b) 3 marks

> **APPROACH**
>
> a Venn diagram for Q and R has four regions — $Q \cap R$, $Q \cap R'$, $Q' \cap R$ and $(Q \cup R)'$ — whose probabilities sum to 1. Writing $P(Q) = q$ and $P(R) = r$, each given piece of information becomes one linear equation in q, r and x, and solving the resulting system gives x.

Let $P(Q) = q$ and $P(R) = r$. Since $Q = (Q \cap R) \cup (Q \cap R')$,

$$q = x + 0.1 \quad ...(1)$$

Since $P(Q' \cap R) = P(R) - P(Q \cap R) = r - x$, and this equals $P(Q) = q$,

$$r - x = q$$
$$\Rightarrow r = q + x \quad ...(2)$$

Since $P\left((Q \cup R)'\right) = 1 - P(Q \cup R) = 1 - (q + r - x)$, and this equals $P(Q) + 2P(R) = q + 2r$,

$$1 - q - r + x = q + 2r$$
$$\Rightarrow 1 + x = 2q + 3r \quad ...(3)$$

Substituting (1) into (2),

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

Substituting $q = x + 0.1$ and $r = 2x + 0.1$ into (3),

$$1 + x = 2(x + 0.1) + 3(2x + 0.1)$$
$$= 2x + 0.2 + 6x + 0.3$$
$$= 8x + 0.5$$

Solving,

$$1 + x = 8x + 0.5$$
$$\Rightarrow 0.5 = 7x$$
$$\Rightarrow x = \frac{1}{14}$$

To check your answer: with $x = \frac{1}{14}$, $q = x + 0.1 = \frac{6}{35}$ and $r = 2x + 0.1 = \frac{17}{70}$, so $1 + x = \frac{15}{14}$ and $2q + 3r = \frac{12}{35} + \frac{51}{70} = \frac{15}{14}$ — equal ✓

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

Question 8 8 marks

Hypothesis testing · test of a mean

Part (a) 1 mark

Lee should carry out a **one-tailed test** (lower-tailed).

Explanation: Lee suspects specifically that the mean number of birds has **decreased**, which is a one-directional claim about μ. A two-tailed test is used only when the claim is that the mean has simply *changed*, so here a one-tailed test is appropriate.

Part (b) 2 marks

Let μ be the population mean number of birds visiting the new feeder in the 20-minute interval while Lee eats his breakfast.

$$H_0 : \mu = 17.3$$
$$H_1 : \mu < 17.3$$

Part (c) 3 marks

We first find unbiased estimates of the population mean and variance from the sample.

$$\begin{aligned}\overline{x} &= \frac{\sum x}{n} \\ &= \frac{512}{32} \\ &= 16\end{aligned}$$

$$\begin{aligned}s^2 &= \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right) \\ &= \frac{1}{31}\left(8702 - \frac{512^2}{32}\right) \\ &= \frac{1}{31}(8702 - 8192) \\ &= \frac{510}{31}\end{aligned}$$

Since $n = 32$ is large, we do not need to assume the population is normal: by the Central Limit Theorem, $\overline{X}$ is approximately normal, and we use the test statistic

$$\begin{aligned}z &= \frac{\overline{x} - \mu_0}{\sqrt{s^2/n}} \\ &= \frac{16 - 17.3}{\sqrt{\frac{510/31}{32}}} \\ &= -1.813 \ \ (4 \text{ s.f., by GC})\end{aligned}$$

For H_0 to be rejected at the $\alpha\%$ level (one-tailed, lower), the $\alpha\%$ significance level must exceed the p-value:

$$\begin{aligned}p\text{-value} &= P(Z < -1.813) \\ &= 1 - \Phi(1.813) \\ &= 0.0349 \ \ (3 \text{ s.f., by GC})\end{aligned}$$

H_0 is rejected when $\frac{\alpha}{100} > p\text{-value}$, so

$$\begin{aligned}\frac{\alpha}{100} &> 0.03491 \\ \Rightarrow \alpha &> 3.491\end{aligned}$$

Since α is a positive integer, the least value satisfying this is $\alpha = 4$.

To check your answer: at $\alpha = 4$, $p = 0.0349 < 0.04$, so H_0 is rejected ✓; at $\alpha = 3$, $p = 0.0349 > 0.03$, so H_0 would not be rejected — consistent with $\alpha_{\min} = 4$. ✓

$\therefore \overline{x} = 16$, $s^2 = \frac{510}{31}$ (≈ 16.5), and the least possible value of α is 4.

Part (d) 1 mark

Since H_0 is rejected, there is sufficient evidence, at the 4% significance level, to conclude that the population mean number of birds visiting the new feeder during the 20-minute interval while Lee eats his breakfast has decreased (is less than 17.3).

Part (e) 1 mark

No.

Explanation: the population distribution of the number of birds visiting the feeder is not stated (or known) to be normal, and $n = 10$ is too small for the Central Limit Theorem to guarantee that $\overline{X}$ is approximately normal. Without either of these, the test in (c) could not be carried out validly on only 10 mornings.

Question 9 11 marks

Binomial distribution · probabilities in context

Part (a) 2 marks

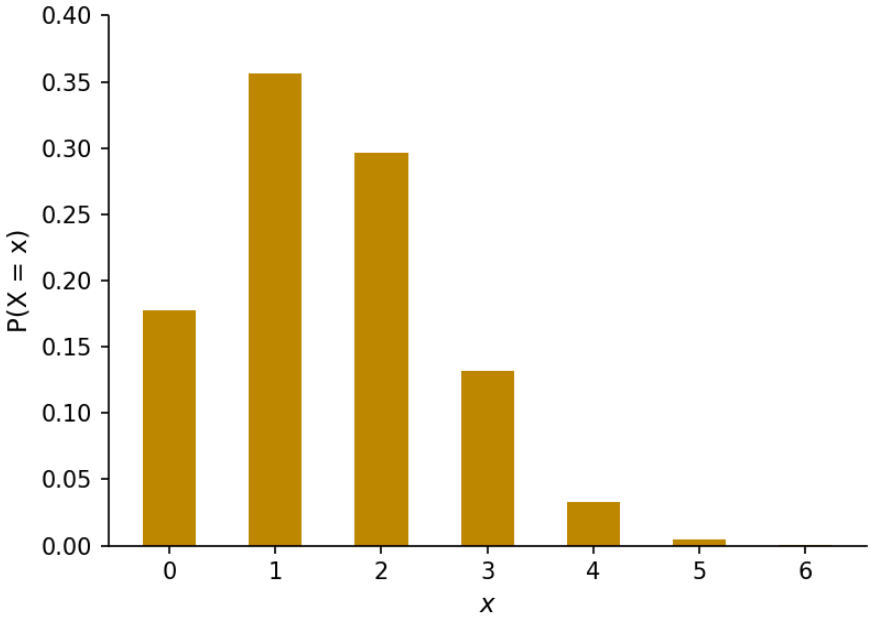
$X \sim B(6, 0.25)$, so

$$P(X = x) = \binom{6}{x}(0.25)^x(0.75)^{6-x}, \quad x = 0, 1, \ldots, 6$$

By GC, the probabilities are:

x	0	1	2	3	4	5	6
$P(X = x)$	0.178	0.356	0.297	0.132	0.0330	0.00439	0.000244

THE LEARNING CAVE thelearningcave.com.sg Physics. Math. Chemistry.



Part (b) 2 marks

Over the 80 days, the total number of pies tested is $80 \times 50 = 4000$.

$$\begin{aligned}\text{Total number unsatisfactory} &= 0(57) + 1(9) + 2(6) + 3(4) + 4(3) + 5(1) + 6(0) \\ &= 9 + 12 + 12 + 12 + 5 \\ &= 50\end{aligned}$$

Estimating the probability by the observed proportion,

$$\begin{aligned}P\left(\text{unsatisfactory}\right) &= \frac{50}{4000} \\ &= \frac{1}{80} \\ &= 0.0125\end{aligned}$$

$\therefore P\left(\text{unsatisfactory}\right) = \frac{1}{80} = 0.0125.$

Part (c)(i) 2 marks

1. Each burger has the same, constant probability of having insufficient cheese, regardless of which burger it is or which day it is tested (the probability does not vary from burger to burger or day to day).
2. Whether any one burger has insufficient cheese is independent of whether any other burger does (burgers are filled/tested independently of one another).

Part (c)(ii) 1 mark

$Y \sim B(120, 0.03)$. Converting the tail to cumulative form for the GC,

$$\begin{aligned}P(Y < 3) &= P(Y \leq 2) \\ &= 0.298 \ \ (3 \text{ s.f., by GC})\end{aligned}$$

thelearningcave.com.sg
Physics. Maths. Chemistry.

To check your answer: adding the three terms directly, $0.97^{120} + 120(0.03)(0.97)^{119} + \binom{120}{2}(0.03)^2(0.97)^{118} = 0.02586 + 0.09597 + 0.17661 = 0.298$. ✓

$\therefore P(Y < 3) \approx 0.298.$

Part (c)(iii) 2 marks

APPROACH

"Fewer than 3 burgers with insufficient cheese" happening on a given day is itself a Bernoulli event, with probability found in (ii); we count how many of the 28 (independent) days this happens on, which is binomial.

Let W be the number of days, out of 28, on which fewer than 3 burgers have insufficient cheese. Since the days are independent, each with the probability found in (ii),

$$W \sim B(28,\ 0.29844)$$
$$E(W) = 28 \times 0.29844$$
$$= 8.36 \ \ (3 \text{ s.f.})$$

$\therefore E(W) \approx 8.36$ days.

Part (c)(iv) 2 marks

APPROACH

Over 28 days the total number of burgers tested is fixed (28×120), each independently having insufficient cheese with probability 0.03, so the total count is itself binomial, and the GC gives its cumulative probabilities directly.

Let T be the total number of burgers, out of $28 \times 120 = 3360$ tested over the 28 days, found to have insufficient cheese. Since the days are independent with the same $p = 0.03$,

$$T \sim B(3360,\ 0.03)$$

Converting the tail to cumulative form for the GC,

$$P(T > 100) = 1 - P(T \leq 100)$$
$$= 1 - 0.49422$$
$$= 0.506 \ \ (3 \text{ s.f., by GC})$$

To check your answer: $E(T) = 3360(0.03) = 100.8$, so asking for more than 100 is asking for roughly "more than the mean", and an answer close to 0.5 is what we expect. ✓

$\therefore P(T > 100) \approx 0.506.$

Question 10 12 marks

Normal distribution · sums of normal variables

Part (a) 1 mark

Let X be the mass, in grams, of a plate before drilling: $X \sim N\left(200,\ 1.6^2\right)$.

Standardising,

$$\begin{aligned} Z &= \frac{197.5 - 200}{1.6} \\ &= -1.5625 \\ P(X > 197.5) &= P(Z > -1.5625) \\ &= \Phi(1.5625) \\ &= 0.941 \ \ (3\text{ s.f.}) \end{aligned}$$

$\therefore P(X > 197.5) \approx 0.941.$

Part (b) 4 marks

APPROACH

Drilling scales every plate's mass by the same factor, so the drilled mass is a linear transformation of X; we find its distribution first, then use it to find the probability that one plate falls in $[190, 192]$, and finally treat "at least 5 of 8 plates" as a binomial count.

Drilling reduces mass by 5%, so the drilled mass is $Y = 0.95X$, where

$$Y \sim N\left(0.95(200),\ 0.95^2\left(1.6^2\right)\right) = N\left(190,\ 1.52^2\right)$$

Let $p = P(190 < Y < 192)$, the probability that one drilled plate has mass between 190 g and 192 g.

$$Z_1 = \frac{190 - 190}{1.52} = 0, \qquad Z_2 = \frac{192 - 190}{1.52} = 1.3158$$

$$\begin{aligned} p &= \Phi(1.3158) - \Phi(0) \\ &= 0.9059 - 0.5 \\ &= 0.4059 \end{aligned}$$

Let W be the number of the 8 plates, out of 8 chosen independently, with mass between 190 g and 192 g:

$$W \sim B(8,\ 0.4059)$$

Converting the tail to the cumulative form the GC uses,

$$\begin{aligned} P(W \geq 5) &= 1 - P(W \leq 4) \\ &= 1 - 0.8171 \\ &= 0.183 \ \ (3\text{ s.f., by GC}) \end{aligned}$$

THE LEARNING CAVE thelearningcave.com.sg Physics. Math. Chemistry.

To check your answer: $E(W) = 8(0.4059) = 3.25$, so $W \geq 5$ is an upper-tail event; a probability of about 0.18 for it is a reasonable size. ✓

$\therefore P(W \geq 5) \approx 0.183.$

Part (c) 3 marks

Let S be the total mass of 20 randomly chosen drilled plates. Since each drilled plate is $N\left(190,\ 1.52^2\right)$ (from (b)) and the plates are chosen independently,

$$S \sim N\left(20(190),\ 20\left(1.52^2\right)\right) = N(3800,\ 46.208)$$

$$Z = \frac{3805 - 3800}{\sqrt{46.208}}$$

$$= 0.7358$$

$$P(S < 3805) = \Phi(0.7358)$$

$$= 0.769 \ \ (3 \text{ s.f.})$$

To check your answer: 3805 is only 0.74 standard deviations above the mean 3800, so a probability just above 0.5 (here ≈ 0.77) is a reasonable size. ✓

$\therefore P(S < 3805) \approx 0.769.$

Part (d) 4 marks

APPROACH

A Value pack's total mass is the sum of the pack of plates and all its screws and bolts, all independent; we find E and Var of this sum, then find the value m that leaves 5% of the normal distribution above it.

Each plate needs 4 screws and 2 bolts (given with the question), so a pack of 20 drilled plates needs $20 \times 4 = 80$ screws and $20 \times 2 = 40$ bolts.

Let M be the total mass of a Value pack: $M = S + (80 \text{ screws}) + (40 \text{ bolts})$, where each screw is $N\left(10,\ 0.3^2\right)$, each bolt is $N\left(44,\ 1.1^2\right)$, and S is the mass of the 20 plates from (c); all these masses are independent of one another.

$$E(M) = 3800 + 80(10) + 40(44)$$

$$= 3800 + 800 + 1760$$

$$= 6360$$

$$Var(M) = Var(S) + 80\left(0.3^2\right) + 40\left(1.1^2\right)$$

$$= 46.208 + 7.2 + 48.4$$

$$= 101.808$$

So $M \sim N(6360,\ 101.808)$, with standard deviation $\sqrt{101.808} = 10.090$.

We need m such that $P(M > m) = 0.05$, i.e. m is z standard deviations above the mean, where $P(Z > z) = 0.05 \Rightarrow z = 1.6449$ (by GC).

$$\begin{aligned} m &= E(M) + z \cdot sd(M) \\ &= 6360 + 1.6449(10.090) \\ &= 6376.6 \end{aligned}$$

To check your answer: at $m = 6376.6$, $Z = \frac{6376.6 - 6360}{10.090} = 1.645$, and $P(Z > 1.645) = 0.05$. ✓

$\therefore$ the mass exceeded by just 5% of Value packs is $m \approx 6377\ g$ (nearest gram).

Question 11 15 marks

Correlation and regression · transformed models

Part (a)(i) 1 mark

As m increases, p increases, but not at a constant rate: p rises quickly for small m and then more slowly as m gets larger, so the rate of increase of p decreases as m increases.

Part (a)(ii) 6 marks

APPROACH

Both proposed models are linear in a transformed variable ($\ln m$ or $\sqrt{m}$), so we find the product moment correlation coefficient for each transformed scatter plot and take the one closer to ± 1 as the better fit.

For the model $p = a + b \ln m$, using the pairs $(\ln m,\ p)$, by GC:

$$r = 0.9285\ \ (4\text{ d.p.}), \qquad a = -703.8, \qquad b = 203.0$$

For the model $p = d + e\sqrt{m}$, using the pairs $(\sqrt{m},\ p)$, by GC:

$$r = 0.9914\ \ (4\text{ d.p.}), \qquad d = 32.95, \qquad e = 17.27$$

To check your answer: both correlation coefficients are close to 1, consistent with the scatter diagram showing a strong relationship; the $\sqrt{m}$ model's is markedly closer to 1, which matches the diagram levelling off much as $\sqrt{m}$ does. ✓

Since r for the $p = d + e\sqrt{m}$ model is closer to 1, this model gives the better fit.

$\therefore p = 32.95 + 17.27\sqrt{m}$, with $r = 0.9914$.

Part (a)(iii)

Using $p = 32.95 + 17.27\sqrt{m}$ from (a)(ii) with $m = 750$,

$$p = 32.95 + 17.27\sqrt{750}$$

$$= 32.95 + 472.96$$
$$= 505.9$$

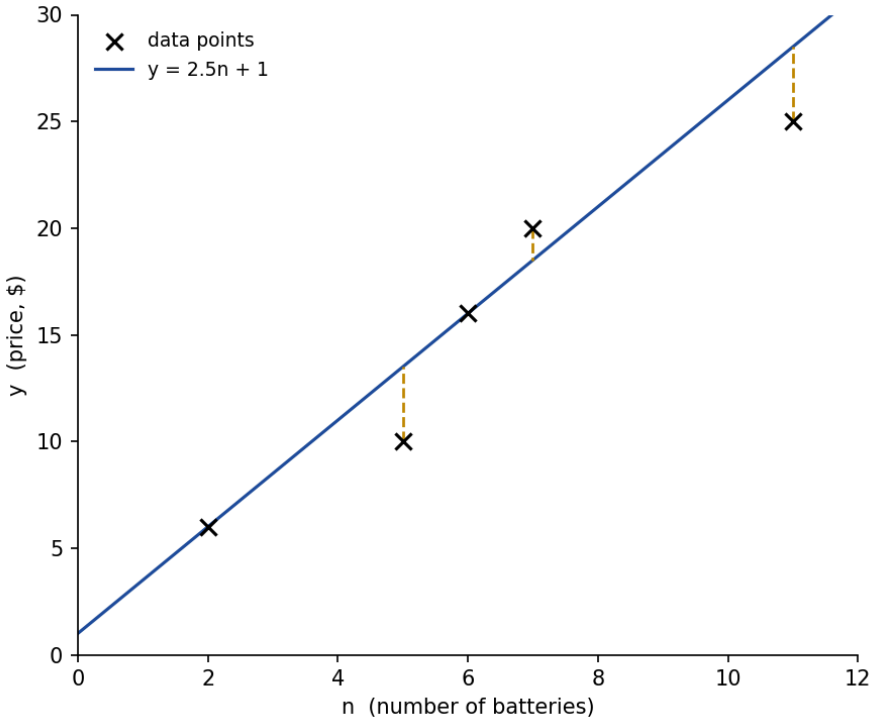
Correct to the nearest 10 cents, $p \approx 510$ cents.

Since $m = 750$ lies within the range of the given data ($24 \leq m \leq 5000$), this estimate is an **interpolation**, so it is likely to be reliable.

$\therefore$ the estimated price is 510 cents, and the estimate is reliable.

Part (b)(i) 1 mark

The line $y = 2.5n + 1$, together with the residual (dashed) for each data point, is:



Part (b)(ii)(A) 1 mark

The residual for each point is $y - f(n)$, where $f(n) = 2.5n + 1$:

n	2	5	6	7	11
y	6	10	16	20	25
$f(n)$	6	13.5	16	18.5	28.5
residual	0	-3.5	0	1.5	-3.5

These are the dashed vertical segments marked in the sketch in (b)(i).

Part (b)(ii)(B) 1 mark

Residuals can be positive or negative, so simply summing them could give a small (or even zero) total even when the model fits poorly, because positive and negative residuals cancel each other out. Squaring the residuals before summing removes this cancellation, so the sum reflects the true overall size of the deviations.

Part (b)(ii)(C) 2 marks

$$\begin{aligned}\text{Sum of squares} &= 0^2 + (-3.5)^2 + 0^2 + 1.5^2 + (-3.5)^2 \\ &= 0 + 12.25 + 0 + 2.25 + 12.25 \\ &= 26.75\end{aligned}$$

To check your answer: the sum of the (unsquared) residuals is $0 - 3.5 + 0 + 1.5 - 3.5 = -5.5 \neq 0$, as expected since $y = 2.5n + 1$ was not obtained by least squares (a least-squares line always has residuals summing to 0). ✓

$\therefore$ the sum of squares of residuals is 26.75.

Part (b)(iii)

A line that fits the data better has a **smaller** sum of squares of residuals than 26.75: the smaller the sum of squares, the better the fit.

Part (b)(iv) 3 marks

APPROACH

All lines $y = 2.5n + c$ share the same slope, so shifting c shifts every residual by the same constant. Writing each residual for $y = 2.5n + c$ in terms of the $c = 1$ residuals lets us express the whole sum of squares as a single quadratic in c, which we then compare against 26.75.

Let $d_i = y_i - 2.5n_i$, the residual when $c = 0$:

n	2	5	6	7	11
$d_i = y_i - 2.5n_i$	1	-2.5	1	2.5	-2.5

For general c, the residual is $d_i - c$, so the sum of squares is

$$\begin{aligned}S(c) &= \sum (d_i - c)^2 \\ &= \sum d_i^2 - 2c\sum d_i + 5c^2\end{aligned}$$

$$\sum d_i = 1 - 2.5 + 1 + 2.5 - 2.5 = -0.5, \qquad \sum d_i^2 = 1 + 6.25 + 1 + 6.25 + 6.25 = 20.75$$

$$S(c) = 20.75 + c + 5c^2$$

To check your answer: $S(1) = 20.75 + 1 + 5 = 26.75$, matching (b)(ii)(C) exactly, since $c = 1$ gives back the original line $y = 2.5n + 1$. ✓

We require $S(c) < S(1) = 26.75$:

$$\begin{aligned} 20.75 + c + 5c^2 &< 26.75 \\ \Rightarrow 5c^2 + c - 6 &< 0 \\ c &= \frac{-1 \pm \sqrt{1 + 120}}{10} \\ &= \frac{-1 \pm 11}{10} \\ &= 1 \text{ or } -1.2 \end{aligned}$$

Since the graph of $5c^2 + c - 6$ against c is an upward parabola, it is negative between its roots.

$\therefore -1.2 < c < 1.$

THE LEARNING CAVE
thelearningcave.com.sg
Physics. Math. Chemistry.