TEN-YEAR SERIES · WORKED SOLUTIONS

2017 A-Level H2 Mathematics Paper 2

10 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 8 marks

Parametric equations · tangents · area

Part (i) 3 marks

Since A and B lie on both C and the line $y = 2x$, we substitute the parametric coordinates into the line's equation.

$$\begin{aligned} 2t &= 2\left(\frac{3}{t}\right) \\ \Rightarrow 2t &= \frac{6}{t} \\ \Rightarrow t^2 &= 3 \\ \Rightarrow t &= \pm\sqrt{3} \end{aligned}$$

So A and B correspond to $t = \sqrt{3}$ and $t = -\sqrt{3}$:

$$A = \left(\sqrt{3},\ 2\sqrt{3}\right), \quad B = \left(-\sqrt{3},\ -2\sqrt{3}\right)$$

By the distance formula,

$$\begin{aligned} AB^2 &= \left(2\sqrt{3} - \left(-2\sqrt{3}\right)\right)^2 + \left(\sqrt{3} - \left(-\sqrt{3}\right)\right)^2 \\ &= \left(4\sqrt{3}\right)^2 + \left(2\sqrt{3}\right)^2 \\ &= 48 + 12 \\ &= 60 \end{aligned}$$

To check your answer: A and B both satisfy $xy = 6$, the cartesian equation of C: $\sqrt{3} \cdot 2\sqrt{3} = 6$ and $\left(-\sqrt{3}\right)\left(-2\sqrt{3}\right) = 6$. ✓

$\therefore AB = \sqrt{60} = 2\sqrt{15}.$

Part (ii) 5 marks

APPROACH

We find the gradient of C at a general parameter t, write down the tangent at P, and find where it meets each axis. Rather than guessing the shape of the locus, we compute F's coordinates in terms of p and eliminate p directly.

Since $x = \frac{3}{t}$ and $y = 2t$,

$$\frac{dx}{dt} = -\frac{3}{t^2}, \quad \frac{dy}{dt} = 2$$

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = 2 \div \left(-\frac{3}{t^2}\right) = -\frac{2t^2}{3}$$

At $P\left(\frac{3}{p}, 2p\right)$, the gradient is $-\frac{2p^2}{3}$, so the tangent at P is

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

Setting $y = 0$ to find D:

$$\begin{aligned} -2p &= -\frac{2p^2}{3}\left(x - \frac{3}{p}\right) \\ \Rightarrow x - \frac{3}{p} &= \frac{3}{p} \\ \Rightarrow x &= \frac{6}{p} \end{aligned}$$

So $D = \left(\frac{6}{p},\ 0\right)$.

Setting $x = 0$ to find E:

$$\begin{aligned} y - 2p &= -\frac{2p^2}{3}\left(0 - \frac{3}{p}\right) \\ \Rightarrow y - 2p &= 2p \\ \Rightarrow y &= 4p \end{aligned}$$

So $E = (0,\ 4p)$.

F is the midpoint of D and E:

$$F = \left(\frac{1}{2}\left(\frac{6}{p} + 0\right),\ \frac{1}{2}(0 + 4p)\right) = \left(\frac{3}{p},\ 2p\right)$$

So F coincides with P itself for every value of p: for this curve, the midpoint of the two intercepts cut off by a tangent is always the point of tangency. Writing $F = (x, y)$, so $x = \frac{3}{p}$ and $y = 2p$, we eliminate p by multiplying:

$$xy = \frac{3}{p} \cdot 2p = 6$$

To check your answer: at $p = 1$, $P = (3, 2)$, $D = (6, 0)$, $E = (0, 4)$, and the midpoint of D and E is $(3, 2) = P$. ✓

$\therefore$ the curve traced by F has cartesian equation $xy = 6$ (with $x \neq 0,\ y \neq 0$) — the same curve as C.

Question 2 9 marks

Sequences and series · AP and GP

Part (i) 2 marks

Let d be the common difference. Using $S_n = \frac{n}{2}\left(2a + (n-1)d\right)$ with $a = 3$, $n = 13$:

$$\begin{aligned} 156 &= \frac{13}{2}\left(2(3) + 12d\right) \\ \Rightarrow 24 &= 6 + 12d \\ \Rightarrow 12d &= 18 \end{aligned}$$

$$\Rightarrow d = \frac{3}{2}$$

∴ the common difference is $\frac{3}{2}$.

Part (ii) 4 marks

APPROACH

The given equation has r^{13} in it, so it must come from the geometric sum formula $S_n = \frac{a(r^n-1)}{r-1}$, which is only valid for $r \neq 1$. We derive the equation by clearing the denominator, and this is exactly the step that can introduce $r = 1$ as a spurious root, since multiplying both sides by $(r-1)$ is not reversible at $r = 1$.

Since $r \neq 1$, the sum of the first 13 terms of the geometric progression is

$$\frac{3\left(r^{13}-1\right)}{r-1} = 156$$
$$\Rightarrow r^{13} - 1 = 52(r-1)$$
$$\Rightarrow r^{13} - 1 = 52r - 52$$
$$\Rightarrow r^{13} - 52r + 51 = 0 \quad \text{(shown)}$$

Substituting $r = 1$: $1^{13} - 52(1) + 51 = 1 - 52 + 51 = 0$, so $r = 1$ does satisfy this equation algebraically. However, the equation was derived by multiplying the sum formula (valid only for $r \neq 1$) through by $(r-1)$, which introduces $r = 1$ as an extraneous root of the cleared equation. Directly, if $r = 1$ every term of the progression equals 3, so $S_{13} = 13(3) = 39 \neq 156$. So $r = 1$ cannot be the actual common ratio, even though it is a root of $r^{13} - 52r + 51 = 0$.

By GC, the real roots of $r^{13} - 52r + 51 = 0$ are $r = 1$ (rejected, as shown) and

$$r = -1.45106\ldots \quad \text{or} \quad r = 1.21002\ldots$$

∴ the possible values of the common ratio are $r = -1.45$ or $r = 1.21$ (3 s.f.).

Part (iii) 3 marks

Since the common ratio is positive, $r = 1.21002397\ldots$ (retaining the unrounded GC value from part (ii)). The nth term of the geometric progression is $3r^{n-1}$; the nth term of the arithmetic progression is $3 + (n-1)\left(\frac{3}{2}\right)$. The given condition is

$$3r^{n-1} > 100\left(3 + (n-1)\left(\frac{3}{2}\right)\right)$$

By GC, tabulating both sides for consecutive integer n:

At $n = 41$: $3r^{40} = 6150$ (3 s.f.), while $100 \times (41\text{st term of A.P.}) = 6300$ — inequality not satisfied.

At $n = 42$: $3r^{41} = 7440$ (3 s.f.), while $100 \times (42\text{nd term of A.P.}) = 6450$ — inequality satisfied.

To check your answer: the geometric term grows like r^n with $r \approx 1.21$ while the arithmetic side grows only linearly in n, so once the geometric term overtakes it near $n \approx 42$ it should stay ahead for all larger n — consistent with $3r^{41}$ exceeding $100\times$(42nd A.P. term) by a growing margin. ✓

$\therefore$ the smallest possible value of n is 42.

Question 3 12 marks

Graphs and transformations · functions and inverses

Part (a)

Since $y = f(x)$ cuts the axes at $(a, 0)$ and $(0, b)$, we have $f(a) = 0$ and $f(0) = b$. Only these two values of f are known, so a point can be stated only when the transformed curve's axis-crossing depends on f at $x = a$ or $x = 0$ alone.

Part (i)

$y = f(2x)$

x-intercept: $f(2x) = 0 \Rightarrow 2x = a \Rightarrow x = \frac{a}{2}$, giving $\left(\frac{a}{2}, 0\right)$.

y-intercept: at $x = 0$, $y = f(0) = b$, giving $(0, b)$.

Part (ii)

$y = f(x - 1)$

x-intercept: $f(x - 1) = 0 \Rightarrow x - 1 = a \Rightarrow x = a + 1$, giving $(a + 1, 0)$.

y-intercept: at $x = 0$, $y = f(-1)$, which is not one of the given values of f, so it is not possible to state this point.

Part (iii)

$y = f(2x - 1)$

x-intercept: $f(2x - 1) = 0 \Rightarrow 2x - 1 = a \Rightarrow x = \frac{a+1}{2}$, giving $\left(\frac{a+1}{2}, 0\right)$.

y-intercept: at $x = 0$, $y = f(-1)$, again not determinable, so it is not possible to state this point.

Part (iv) 4 marks

$y = f^{-1}(x)$

Since f^{-1} exists, its graph is the reflection of $y = f(x)$ in the line $y = x$, so every point (u, v) on $y = f(x)$ gives a point (v, u) on $y = f^{-1}(x)$. Since $(a, 0)$ and $(0, b)$ lie on $y = f(x)$, the points $(0, a)$ and $(b, 0)$ lie on $y = f^{-1}(x)$.

$\therefore y = f^{-1}(x)$ cuts the axes at $(b, 0)$ and $(0, a)$.

Part (b)(i) 2 marks

$$a = 1$$

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

Explanation: g is defined by $1 - \frac{1}{1-x}$, and at $x = 1$ the denominator $1 - x = 0$, so $\frac{1}{1-x}$ is undefined there. Hence $x = 1$ must be excluded from the domain of g.

Part (b)(ii) 4 marks

APPROACH

It is easier to simplify $g(x)$ into a single fraction first, then substitute this simplified form into itself for $g^2(x)$, rather than composing the original unsimplified expression with itself.

$$\begin{aligned} g(x) &= 1 - \frac{1}{1-x} \\ &= \frac{(1-x)-1}{1-x} \\ &= \frac{-x}{1-x} \\ &= \frac{x}{x-1} \end{aligned}$$

Using this simplified form,

$$\begin{aligned} g^2(x) &= g\left(\frac{x}{x-1}\right) \\ &= \frac{\frac{x}{x-1}}{\frac{x}{x-1} - 1} \\ &= \frac{\frac{x}{x-1}}{\frac{1}{x-1}} \\ &= x \end{aligned}$$

So $g^2(x) = x$ for $x \neq 1$: g is its own inverse.

To find $g^{-1}(x)$, let $y = g(x) = \frac{x}{x-1}$ and solve for x in terms of y.

$$\begin{aligned} y &= \frac{x}{x-1} \\ \Rightarrow y(x-1) &= x \\ \Rightarrow yx - x &= y \\ \Rightarrow x(y-1) &= y \\ \Rightarrow x &= \frac{y}{y-1} \end{aligned}$$

Replacing y by x:

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

To check your answer: since $g^2(x) = x$ for all x in the domain, g is self-inverse, so g^{-1} must equal g — and indeed the formula obtained for $g^{-1}(x)$ below is identical to $g(x)$. ✓

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

Part (b)(iii) 2 marks

Using $g^2(x) = x$ and $g^{-1}(x) = \frac{x}{x-1}$ from part (ii), the equation $g^2(b) = g^{-1}(b)$ becomes

$$\begin{aligned} b &= \frac{b}{b-1} \\ \Rightarrow b(b-1) &= b \\ \Rightarrow b^2 - b &= b \\ \Rightarrow b^2 - 2b &= 0 \\ \Rightarrow b(b-2) &= 0 \\ \Rightarrow b = 0 &\text{ or } b = 2 \end{aligned}$$

Both values satisfy $b \neq 1$, so both lie in the domain.

$\therefore b = 0$ or $b = 2$.

Question 4 11 marks

Integration · area · volume of revolution

Part (a) 4 marks

APPROACH

The area of the plate is the region between the curve and the line, so we first find where they meet, then integrate the difference of the upper and lower expressions between those points.

Writing the line as $y = \frac{x-1}{2}$, we find where it meets the curve.

$$\begin{aligned} x^2 - 6x + 5 &= \frac{x-1}{2} \\ \Rightarrow 2x^2 - 12x + 10 &= x - 1 \\ \Rightarrow 2x^2 - 13x + 11 &= 0 \\ \Rightarrow (x-1)(2x-11) &= 0 \\ \Rightarrow x = 1 &\text{ or } x = \frac{11}{2} \end{aligned}$$

At $x = 3$ (between the two roots), the line gives $y = 1$ and the curve gives $y = 9 - 18 + 5 = -4$, so the line lies above the curve on $\left[1, \frac{11}{2}\right]$. So

$$\begin{aligned} \text{Area} &= \int_1^{11/2} \left[\frac{x-1}{2} - \left(x^2 - 6x + 5\right)\right] dx \\ &= \int_1^{11/2} \left(-x^2 + \frac{13}{2}x - \frac{11}{2}\right) dx \\ &= \left[-\frac{x^3}{3} + \frac{13}{4}x^2 - \frac{11}{2}x\right]_1^{11/2} \end{aligned}$$

$$= \frac{605}{48} - \left(-\frac{31}{12}\right)$$
$$= \frac{243}{16}$$

To check your answer: by GC, numerically integrating $\left[\frac{x-1}{2} - \left(x^2 - 6x + 5\right)\right]$ from 1 to 5.5 gives $15.1875 = \frac{243}{16}$. ✓

∴ the area of the plate is $\frac{243}{16}$ units^2.

Part (b)(i) 7 marks

APPROACH

The container is a solid of revolution about the y-axis, so its volume is $\pi \int x^2\,dy$. Since $x^2 = \frac{y}{(a-y^2)^2}$ has a factor of y multiplying a function of $a - y^2$, the substitution $u = a - y^2$ turns it into a standard power of u.

$$V = \pi \int_0^1 x^2\,dy = \pi \int_0^1 \frac{y}{(a-y^2)^2}\,dy$$

Let $u = a - y^2$, so $du = -2y\,dy$.

$$\int \frac{y}{(a-y^2)^2}\,dy = -\frac{1}{2}\int \frac{du}{u^2}$$
$$= \frac{1}{2u} + C$$
$$= \frac{1}{2(a-y^2)} + C$$
$$V = \pi\left[\frac{1}{2(a-y^2)}\right]_0^1$$
$$= \pi\left(\frac{1}{2(a-1)} - \frac{1}{2a}\right)$$
$$= \frac{\pi}{2}\left(\frac{a-(a-1)}{a(a-1)}\right)$$
$$= \frac{\pi}{2a(a-1)}$$

∴ the volume of the container is $\frac{\pi}{2a(a-1)}$.

Part (b)(ii)

APPROACH

The volume formula from part (i) applies to the second container with a replaced by b, since it was formed the same way. We set this expression equal to 4 times the first volume and solve for b in terms of a.

$$\frac{\pi}{2b(b-1)} = 4\left(\frac{\pi}{2a(a-1)}\right)$$
$$\Rightarrow \frac{1}{b(b-1)} = \frac{4}{a(a-1)}$$
$$\Rightarrow b(b-1) = \frac{a(a-1)}{4}$$
$$\Rightarrow b^2 - b - \frac{a(a-1)}{4} = 0$$

By the quadratic formula,

$$b = \frac{1 \pm \sqrt{1 + a(a-1)}}{2} = \frac{1 \pm \sqrt{a^2 - a + 1}}{2}$$

Since $a > 1$, $a(a-1) > 0$, so $a^2 - a + 1 > 1$ and $\sqrt{a^2 - a + 1} > 1$. So $b = \frac{1-\sqrt{a^2-a+1}}{2} < 0$ is rejected, since b must be greater than 1.

$$b = \frac{1 - \sqrt{a^2 - a + 1}}{2} \text{ (rej. as } b > 1), \quad b = \frac{1 + \sqrt{a^2 - a + 1}}{2}$$

To check your answer: at $a = 2$, this gives $b = \frac{1+\sqrt{3}}{2} \approx 1.37$; the volume with $a = 2$ is $\frac{\pi}{4}$, and with this b the volume is $\frac{\pi}{2b(b-1)} \approx \pi$, which is indeed $4 \times \frac{\pi}{4}$. ✓

$\therefore b = \frac{1+\sqrt{a^2-a+1}}{2}$.

Question 5 7 marks

Discrete random variables · expectation and variance

Part (i) 3 marks

APPROACH

T is the number of counters removed when the 2nd red counter is drawn. Since there are only 3 yellow counters in the bag, at most 3 yellows can be drawn before the 2nd red, so T ranges from 2 (both reds drawn first) to 5 (all 3 yellows drawn before the 2nd red). For $T = t$, the first $t - 1$ counters removed must contain exactly 1 red and $t - 2$ yellow (in any order), and the t-th counter removed must be red.

$$P(T = t) = \frac{\binom{6}{1}\binom{3}{t-2}}{\binom{9}{t-1}} \times \frac{5}{10 - t}, \quad t = 2, 3, 4, 5$$

By GC (or direct calculation),

$$P(T = 2) = \frac{6}{9} \times \frac{5}{8} = \frac{5}{12}$$
$$P(T = 3) = \frac{\binom{6}{1}\binom{3}{1}}{\binom{9}{2}} \times \frac{5}{7} = \frac{18}{36} \times \frac{5}{7} = \frac{5}{14}$$
$$P(T = 4) = \frac{\binom{6}{1}\binom{3}{2}}{\binom{9}{3}} \times \frac{5}{6} = \frac{18}{84} \times \frac{5}{6} = \frac{5}{28}$$

thelearningcave.com.sg
Physics. Math. Chemistry.

$$P(T=5)=\frac{\binom{6}{1}\binom{3}{3}}{\binom{9}{4}}\times\frac{5}{5}=\frac{6}{126}\times 1=\frac{1}{21}$$

To check your answer: the four probabilities sum to $\frac{5}{12}+\frac{5}{14}+\frac{5}{28}+\frac{1}{21}=\frac{35+30+15+4}{84}=1$. ✓

$\therefore$

t	2	3	4	5
$P(T=t)$	$\frac{5}{12}$	$\frac{5}{14}$	$\frac{5}{28}$	$\frac{1}{21}$

Part (ii) 2 marks

Using the distribution in (i),

$$E(T)=\sum t\,P(T=t)=\frac{20}{7}$$

$$Var(T)=E\left(T^2\right)-[E(T)]^2=\frac{125}{14}-\left(\frac{20}{7}\right)^2=\frac{75}{98}$$

$\therefore E(T)=\frac{20}{7}=2.86$ (3 s.f.), $Var(T)=\frac{75}{98}=0.765$ (3 s.f.).

Part (iii) 2 marks

APPROACH

"At least 4 counters" means $T\geq 4$, i.e. $T=4$ or $T=5$. Let W be the number of Lee's 15 games (out of 15) in which he has to take at least 4 counters out of the bag. Since each game is an independent repeat of the same experiment, W follows a binomial distribution.

$$P(T\geq 4)=P(T=4)+P(T=5)=\frac{5}{28}+\frac{1}{21}=\frac{19}{84}$$

Let W be the number of Lee's 15 games in which he takes at least 4 counters out of the bag. Then $W\sim B\left(15,\ \frac{19}{84}\right)$.

$$\begin{aligned}P(W\geq 5)&=1-P(W\leq 4)\\&=1-0.762\ldots\\&=0.238\ \text{(3 s.f.)}\ \ \text{(by GC)}\end{aligned}$$

$\therefore$ the probability is 0.238 (3 s.f.).

Question 6 9 marks

Permutations and combinations · circular arrangements

Part (i) 2 marks

APPROACH

Treat each family's 4 cards as a single block, since the 4 cards in each family set must stay next to each other. There are 5 such blocks to arrange, and within each block the 4 family members can be ordered independently.

$$\begin{aligned}\text{no. of ways} &= 5! \times (4!)^5 \\ &= 120 \times 24^5 \\ &= 955\,514\,880\end{aligned}$$

$\therefore$ the cards can be arranged in $955\,514\,880$ ways.

Part (ii) 3 marks

APPROACH

Father Red belongs to both the father block (5 fathers) and the Red family block (4 Red cards), and Father Blue belongs to both the father block and the Blue family block. Since Father Red is the *only* card common to the father block and the Red block, these two blocks can each be a run of consecutive cards only if they overlap at that one shared card — i.e. Father Red sits at one end of the father block, with the other 3 Red family members running on immediately outside that end. The same argument applies to Father Blue at the other end of the father block. So all three conditions force a single mega-block of $5 + 3 + 3 = 11$ consecutive cards (5 fathers, plus the 3 non-father Red cards, plus the 3 non-father Blue cards), with the remaining 9 cards (the non-father members of the Green, Yellow and Orange families) free to sit anywhere else in the row.

Within the mega-block:

- Father Red and Father Blue must occupy the two *ends* of the 5-card father run (they cannot both be at the same end), so there are 2 ways to decide which end is Red and which is Blue.
- The remaining 3 fathers (Green, Yellow, Orange) fill the middle of the father run, in $3!$ ways.
- The 3 non-father Red cards run outward from Father Red, in $3!$ ways.
- The 3 non-father Blue cards run outward from Father Blue, in $3!$ ways.

$$\text{arrangements within the mega-block} = 2 \times 3! \times 3! \times 3! = 432$$

Treating the mega-block as one unit, we now arrange it together with the other 9 free cards: 10 units in total.

$$\begin{aligned}\text{no. of ways} &= 10! \times 432 \\ &= 3\,628\,800 \times 432 \\ &= 1\,567\,641\,600\end{aligned}$$

To check your answer: on a scaled-down version of this problem (3 families of 2 members: father and mother, fathers together, and two of the three family-pairs each together), the same overlapping-block argument gives $2! \times 2 = 4$, which matches a direct brute-force count of all $6! = 720$ arrangements. ✓

∴ the cards can be arranged in $1\,567\,641\,600$ ways.

Part (iii) 4 marks

APPROACH

Arranging the 15 non-father cards around the circle first leaves 15 gaps between them (a circular arrangement of n objects has n gaps). Placing the 5 distinct father cards into 5 different gaps then guarantees no two father cards are adjacent, since each gap holds at most one father.

There are 20 distinct cards, so the total number of circular arrangements is

$$(20-1)! = 19!$$

Arrange the 15 non-father cards in a circle first:

$$(15-1)! = 14! \text{ ways}$$

This creates 15 gaps between consecutive non-father cards. Choosing and arranging the 5 distinct father cards into 5 of these 15 gaps (order matters, since the fathers are distinct cards):

$$\frac{15!}{10!} \text{ ways}$$

$$\begin{aligned} P(\text{no two father cards adjacent}) &= \frac{14! \times \frac{15!}{10!}}{19!} \\ &= \frac{14 \times 13 \times 12 \times 11}{19 \times 18 \times 17 \times 16} \\ &= \frac{1001}{3876} \\ &= 0.258 \ \ (3 \text{ s.f.}) \end{aligned}$$

To check your answer: by GC, $1001/3876 = 0.25826\ldots$, consistent with the exact fraction. ✓

∴ the probability that no two father cards are next to each other is $\frac{1001}{3876} = 0.258$ (3 s.f.).

Question 7 10 marks

Hypothesis testing · central limit theorem

Part (i) 1 mark

A sample is random in this context if every biscuit bar produced that day has an equal chance of being selected for the sample, and bars are selected independently of one another.

Part (ii) 2 marks

Since $n = 40$, $\Sigma(x - 32) = -7.7$ and $\Sigma(x - 32)^2 = 11.05$,

$$\begin{aligned}\overline{x} &= 32 + \frac{\Sigma(x - 32)}{n} \\ &= 32 + \frac{-7.7}{40} \\ &= 31.8075\end{aligned}$$

$$\begin{aligned}s^2 &= \frac{1}{n-1}\left[\Sigma(x-32)^2 - \frac{(\Sigma(x-32))^2}{n}\right] \\ &= \frac{1}{39}\left[11.05 - \frac{(-7.7)^2}{40}\right] \\ &= \frac{9.5675}{39} \\ &= 0.245 \ (3 \text{ s.f.})\end{aligned}$$

$\therefore$ the unbiased estimate of the population mean is $31.8075\ g$, and the unbiased estimate of the population variance is $0.245\ g^2$ (3 s.f.).

Part (iii) 5 marks

Let μ be the population mean mass, in grams, of biscuit bars produced that day.

We test

$$H_0 : \mu = 32 \qquad H_1 : \mu \neq 32$$

at the 1% significance level.

APPROACH

The population distribution of masses is not given, but $n = 40$ is large, so by the Central Limit Theorem $\overline{X}$ is approximately normal even though the population variance is unknown (and is estimated by s^2). This justifies a z-test.

Under H_0, since $n = 40$ is large, by the Central Limit Theorem

$$\overline{X} \sim N\left(32,\ \frac{s^2}{n}\right) \text{ approximately, where } s^2 = 0.24533 \ (5 \text{ s.f.})$$

Using a z-test,

$$\begin{aligned}z &= \frac{\overline{x} - 32}{\sqrt{s^2/n}} \\ &= \frac{31.8075 - 32}{\sqrt{0.24533/40}} \\ &= -2.458 \ (4 \text{ s.f.}) \quad (\text{by GC})\end{aligned}$$

By GC, the p-value is 0.0140 (3 s.f.).

Since $p\text{-value} = 0.0140 > 0.01$, we do not reject H_0.

$\therefore$ at the 1% level of significance, there is insufficient evidence that the mean mass of biscuit bars differs from 32 grams, i.e. no evidence against the claim on the packets.

Part (iv) 2 marks

Explanation: No need to know the population distribution because $n = 40$ is large, so by the Central Limit Theorem the sample mean $\overline{X}$ is approximately normally distributed whatever the underlying distribution of the individual bar masses is. This means the test in (iii) is valid (approximately) regardless of the population distribution.

Question 8 10 marks

Correlation and regression · scatter diagrams

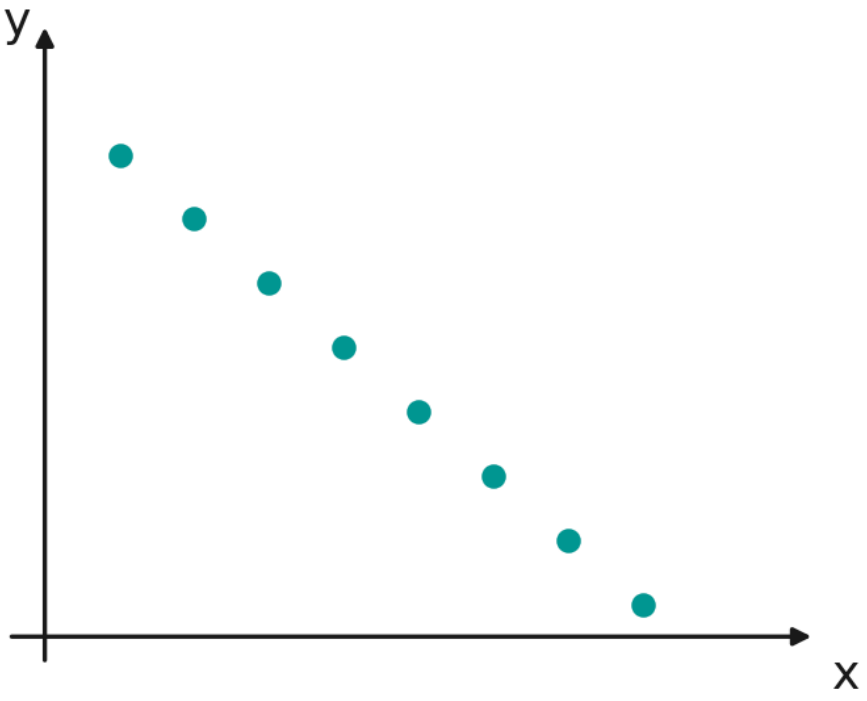
Part (a)

APPROACH

The product moment correlation coefficient measures the strength of a *linear* relationship. For (ii) and (iii) we need patterns that are not simply "more scatter" versions of a straight line: (ii) needs a clear non-linear pattern with zero net linear trend, and (iii) needs a positive linear trend with enough scatter to pull r below 1 but not below 0.5.

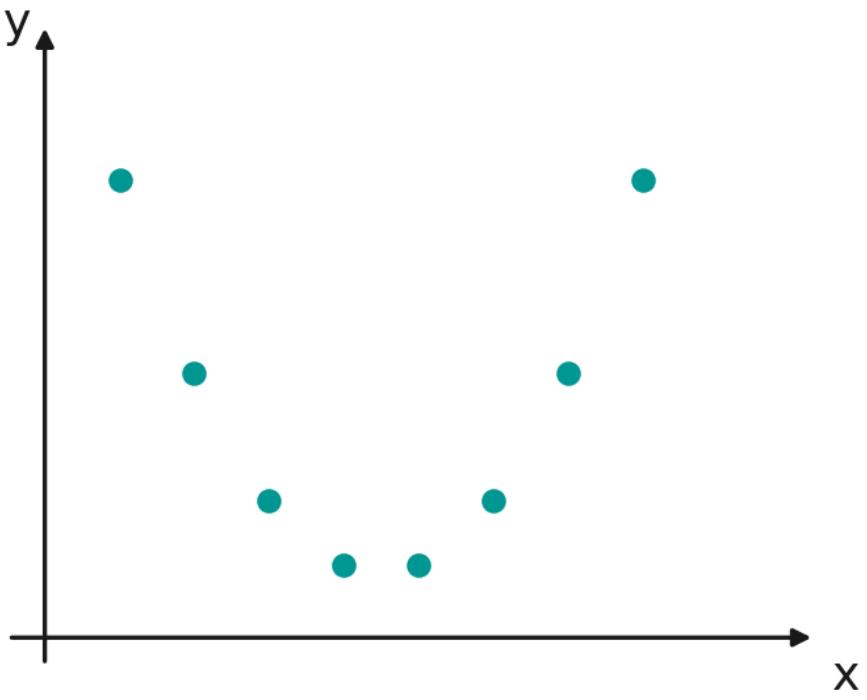
Part (i)

$r = -1$: eight points lying exactly on a straight line of negative gradient.



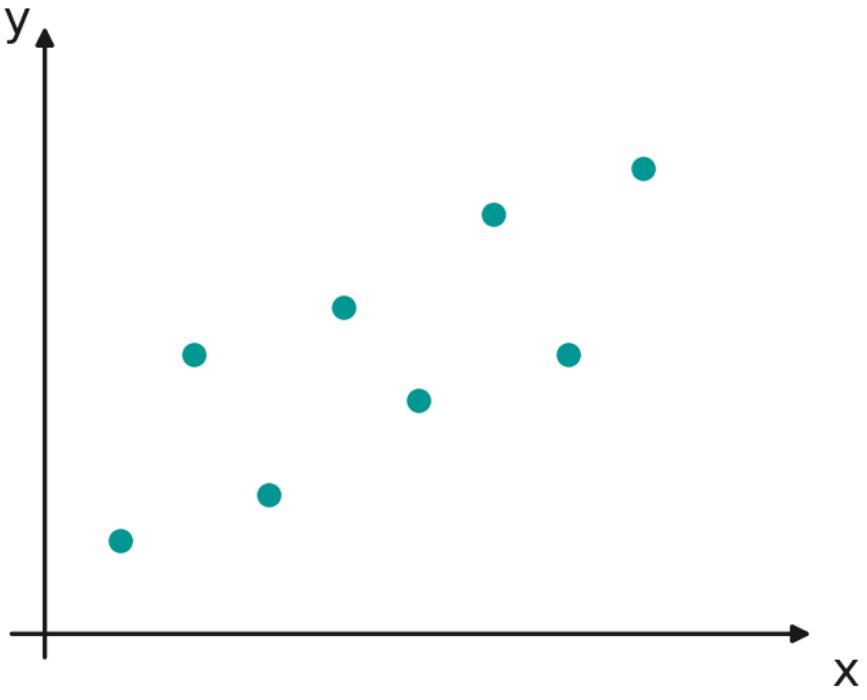
Part (ii)

$r = 0$: eight points in a symmetric curved (non-linear) pattern, so that increases in x are matched equally by increases and decreases in y, giving no net linear trend.



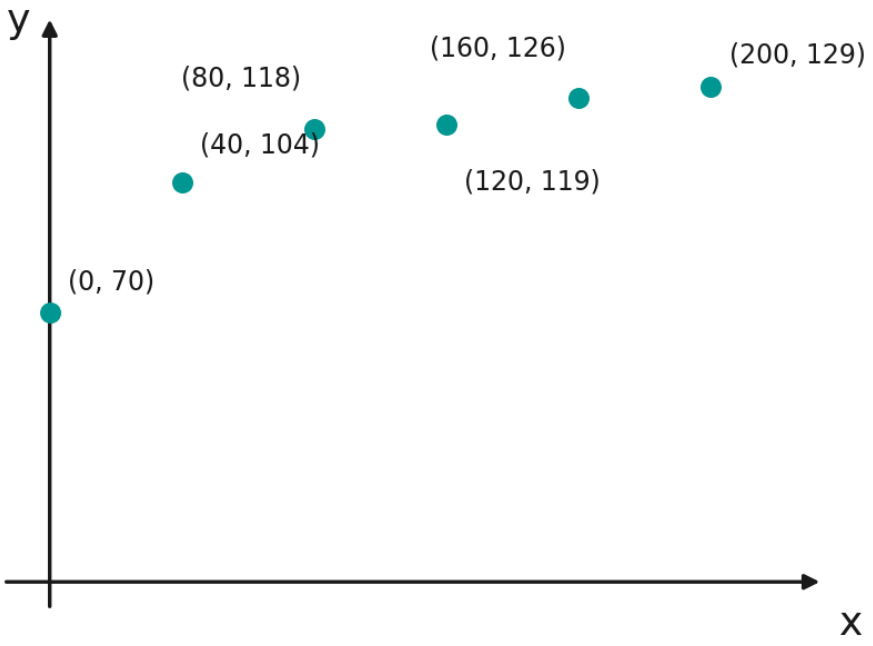
Part (iii)

$0.5 < r < 0.9$: eight points showing a clear increasing trend, but with enough scatter about the underlying line that the fit is not near-perfect.



Part (b)

Part (i)



The scatter diagram shows y increasing with x, but at a decreasing rate, levelling off as x grows, with no sign of turning downward.

Note: $x = 0$ is one of the given data points. Models **(B)** $y = \frac{a}{x^2} + b$ and **(C)** $y = a \ln 2x + b$ are both undefined at $x = 0$, so neither can represent this data set at all. This leaves **(A)** $y = ax^2 + b$ and **(D)** $y = a\sqrt{x} + b$ to compare.

By GC, the product moment correlation coefficients for the two remaining linearised models are

$$r_{x^2,\,y} = 0.735 \ \ (3 \text{ s.f.}), \qquad r_{\sqrt{x},\,y} = 0.981 \ \ (3 \text{ s.f.})$$

Since $\left|r_{\sqrt{x},\,y}\right|$ is much closer to 1, and $y = ax^2 + b$ would be increasing at an *increasing* rate (the opposite of the levelling-off pattern seen), model **(D)** $y = a\sqrt{x} + b$ provides the most accurate model.

Part (ii)

Using model **(D)** with x and y as given, by GC (linear regression of y on $\sqrt{x}$),

$$y = 4.18\sqrt{x} + 74.0 \ \ (3 \text{ s.f.})$$

The product moment correlation coefficient for this model is

$$r = 0.981 \ \ (3 \text{ s.f.})$$

Part (iii)

1. $x = 189$ lies within the range of the given data ($0 \le x \le 200$), so this is interpolation, not extrapolation, which is more reliable.
2. The product moment correlation coefficient $r = 0.981$ is close to 1, showing a strong linear relationship between $\sqrt{x}$ and y, so the model fits the data well.

Question 9 12 marks

Binomial distribution · assumptions · probabilities

Part (i) 2 marks

For the number of faulty lights in a box to be well modelled by a binomial distribution, we need, in this context:

1. Whether any one light is faulty is independent of whether any other light in the box is faulty.
2. Each light has the same constant probability, 0.08, of being faulty.

Part (ii) 1 mark

Let X be the number of faulty lights in a box of 12. Then $X \sim B(12,\ 0.08)$.

$$\begin{aligned} & P(X \geq 1) \\ \Rightarrow\ & 1 - P(X = 0) \\ = \ & 1 - (0.92)^{12} \\ = \ & 0.632 \ \ (3 \text{ s.f.}) \ \ \ (\text{by GC}) \end{aligned}$$

Part (iii) 2 marks

Let W be the number of boxes, out of the 20 in a carton, that contain at least one faulty light. Since the boxes are independent, each box has probability $P(X \geq 1) = 0.63233$ (5 s.f.) of containing at least one faulty light, so we require every one of the 20 boxes to satisfy this.

$$\begin{aligned} & P\left(\text{every box has} \geq 1 \text{ faulty light}\right) \\ \Rightarrow\ & [P(X \geq 1)]^{20} \\ = \ & (0.63233\ldots)^{20} \\ = \ & 0.000104 \ \ (3 \text{ s.f.}) \ \ \ (\text{by GC}) \end{aligned}$$

Part (iv) 2 marks

> **APPROACH**
>
> A carton contains $20 \times 12 = 240$ lights in total, and each light is independently faulty with probability 0.08, so the total number of faulty lights in the carton is itself binomial.

Let Y be the number of faulty lights in a randomly selected carton of 240 lights. Then $Y \sim B(240,\ 0.08)$.

$$\begin{aligned} & P(Y \geq 20) \\ \Rightarrow\ & 1 - P(Y \leq 19) \\ = \ & 1 - 0.54167\ldots \\ = \ & 0.458 \ \ (3 \text{ s.f.}) \ \ \ (\text{by GC}) \end{aligned}$$

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

Part (v)

Explanation: The event in (iv) only requires the total count of faulty lights across the whole carton to reach 20; it does not matter how those faulty lights are distributed among the 20 boxes — they could all be concentrated in a handful of boxes. The event in (iii) is far more restrictive: it requires faulty lights to be spread out so that *every single one* of the 20 boxes has at least one. Since (iii)'s event is a much stronger requirement than (iv)'s, $P\left((\text{iv})\right) > P\left((\text{iii})\right)$.

Part (vi) 3 marks

Let F be the event that a light is faulty, and T be the event that the quick test identifies a light as faulty.

$$P(F) = 0.08, \qquad P(T \mid F) = 0.95, \qquad P(T \mid F') = 0.06$$

> **APPROACH**
> We want $P(F' \mid T)$, the reverse of the conditional probabilities given, so we use Bayes' theorem, finding $P(T)$ first.

$$\begin{aligned} &P(T) \\ &= P(F)P(T \mid F) + P(F')P(T \mid F') \\ &= (0.08)(0.95) + (0.92)(0.06) \\ &= 0.1312 \\ &P(F' \mid T) \\ &= \frac{P(F')P(T \mid F')}{P(T)} \\ &= \frac{(0.92)(0.06)}{0.1312} \\ &= 0.421 \ \ (3 \text{ s.f.}) \end{aligned}$$

Part (vii) 1 mark

The test correctly identifies a light either by (correctly) flagging a faulty light as faulty, or by (correctly) not flagging a non-faulty light as faulty.

$$\begin{aligned} &P\left(\text{correct identification}\right) \\ &= P(F)P(T \mid F) + P(F')P(T' \mid F') \\ &= (0.08)(0.95) + (0.92)(0.94) \\ &= 0.941 \ \ (3 \text{ s.f.}) \end{aligned}$$

Part (viii) 1 mark

The test correctly identifies lights 94.1% of the time, which sounds good, but a large fraction — 42.1% — of lights the test flags as faulty are in fact not faulty, so good lights would be needlessly discarded. Whether the test is worthwhile depends on whether the cost of discarding these good lights is outweighed by the benefit of catching most of the genuinely faulty ones.

Question 10 12 marks

Normal distribution · sums of normal variables

Part (i) 1 mark

Let X be the mass, in grams, of a randomly selected sphere. Then $X \sim N\left(20,\ 0.5^2\right)$.

Standardising,

$$\begin{aligned} & P(X > 20.2) \\ \Rightarrow\ & P\left(Z > \frac{20.2 - 20}{0.5}\right) \\ \Rightarrow\ & P(Z > 0.4) \\ =\ & 0.345\ \ (3\text{ s.f.})\ \ (\text{by GC}) \end{aligned}$$

Part (ii) 3 marks

> **APPROACH**
> Coating increases the mass of *each* sphere by 10%, so the coated mass is a linear transformation, $1.1X$, of the uncoated mass, not $X + 10\%$ of a fixed amount.

Let Y be the mass, in grams, of a coated sphere, so $Y = 1.1X$. Since $X \sim N\left(20,\ 0.5^2\right)$,

$$Y \sim N\left(1.1(20),\ (1.1 \times 0.5)^2\right) = N\left(22,\ 0.55^2\right)$$

$$P(21.5 < Y < 22.45) = 0.612\ \ (3\text{ s.f.})\ \ (\text{by GC})$$

$\therefore Y \sim N(22,\ 0.3025)$, and $P(21.5 < Y < 22.45) = 0.612$ (3 s.f.).

Part (iii) 4 marks

Let B be the mass, in grams, of a randomly selected metal bar, and let $B \sim N\left(\mu,\ \sigma^2\right)$.

$$P(B > 12.2) = 0.6 \qquad \text{and} \qquad P(B < 12) = 0.25$$

By GC (inverse normal),

$$\frac{12.2 - \mu}{\sigma} = -0.253347\ldots \qquad \text{and} \qquad \frac{12 - \mu}{\sigma} = -0.674490\ldots$$

$$12.2 = \mu - 0.253347\sigma \qquad \ldots(1)$$

$$12 = \mu - 0.674490\sigma \qquad \ldots(2)$$

Solving (1) and (2) simultaneously (by GC),

$$\mu = 12.3\ \ (3\text{ s.f.}), \qquad \sigma = 0.475\ \ (3\text{ s.f.})$$

To check your answer: with $\mu = 12.320$ and $\sigma = 0.47490$, $P(B > 12.2) = 0.600$ and $P(B < 12) = 0.250$, which match. ✓

$\therefore$ the mean mass of the bars is $12.3\ g$ and the standard deviation is $0.475\ g$ (3 s.f.).

Part (iv) 4 marks

> **APPROACH**
>
> The total mass of a component is a sum of three independent normal random variables (two coated spheres and one bar), so it is itself normal, with mean and variance found by adding the individual means and variances.

Let $T = Y_1 + Y_2 + B$, where Y_1, Y_2 are the masses of the two coated spheres and B is the mass of the bar, be the total mass, in grams, of the component. Since $Y_1, Y_2 \sim N(22,\ 0.3025)$ independently and $B \sim N\left(12.320,\ 0.47490^2\right)$ independently of Y_1, Y_2,

$$T \sim N\left(2(22) + 12.320,\ 2(0.3025) + 0.47490^2\right) = N(56.320,\ 0.83053)\ (5\text{ s.f.})$$

Since $P(T > k) = 0.75$, by GC (inverse normal),

$$k = 55.7\ (3\text{ s.f.})$$

thelearningcave.com.sg
Physics. Math. Chemistry.