TEN-YEAR SERIES · WORKED SOLUTIONS

2019 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

Integration · by parts · substitution

Part (i) 2 marks

Since the integrand is a product of x and $(1-x)^{1/2}$, we integrate by parts, taking $u = x$ (which differentiates to something simpler) and $dv = (1-x)^{1/2}\,dx$.

$$u = x, \qquad \frac{du}{dx} = 1$$
$$dv = (1-x)^{1/2}\,dx, \qquad v = -\frac{2}{3}(1-x)^{3/2}$$

By parts,

$$\begin{aligned} I &= uv - \int v\,\frac{du}{dx}\,dx \\ &= -\frac{2}{3}x(1-x)^{3/2} + \frac{2}{3}\int (1-x)^{3/2}\,dx \\ &= -\frac{2}{3}x(1-x)^{3/2} - \frac{4}{15}(1-x)^{5/2} + C \end{aligned}$$

$\therefore I = -\frac{2}{3}x(1-x)^{3/2} - \frac{4}{15}(1-x)^{5/2} + C.$

Part (ii) 4 marks

Using $u^2 = 1 - x$: differentiating, $2u\,du = -dx$, so $dx = -2u\,du$; also $x = 1 - u^2$ and $(1-x)^{1/2} = u$.

$$\begin{aligned} I &= \int x(1-x)^{1/2}\,dx \\ &= \int \left(1-u^2\right)\, u\,(-2u)\,du \\ &= \int \left(2u^4 - 2u^2\right)\,du \\ &= \frac{2}{5}u^5 - \frac{2}{3}u^3 + C \end{aligned}$$

Since $u = (1-x)^{1/2}$,

$$I = \frac{2}{5}(1-x)^{5/2} - \frac{2}{3}(1-x)^{3/2} + C$$

Part (iii)

APPROACH

Any two antiderivatives of the same function must differ by a constant; to see *what* that constant is here, we subtract the two expressions found in (i) and (ii) (without their constants of integration, which is where the difference would otherwise hide) and simplify by taking out the common factor $(1-x)^{3/2}$.

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

Let $I_1 = -\frac{2}{3}x(1-x)^{3/2} - \frac{4}{15}(1-x)^{5/2}$ and $I_2 = \frac{2}{5}(1-x)^{5/2} - \frac{2}{3}(1-x)^{3/2}$, the two expressions with the constant dropped.

$$\begin{aligned} & I_1 - I_2 \\ &= -\frac{2}{3}x(1-x)^{3/2} + \frac{2}{3}(1-x)^{3/2} - \frac{4}{15}(1-x)^{5/2} - \frac{2}{5}(1-x)^{5/2} \\ &= \frac{2}{3}(1-x)^{3/2}[1-x] - \frac{2}{3}(1-x)^{5/2} \\ &= \frac{2}{3}(1-x)^{5/2} - \frac{2}{3}(1-x)^{5/2} \\ &= 0 \end{aligned}$$

To check your answer: in factored form both (i) and (ii) reduce to $-\frac{2}{15}(3x+2)(1-x)^{3/2}$; differentiating this back gives $x(1-x)^{1/2}$, the original integrand. ✓

$\therefore I_1 - I_2 = 0$, a constant. (shown)

Note: the constant here happens to be 0 — the two methods give exactly the same antiderivative, not merely antiderivatives that differ by a nonzero constant.

Question 2 6 marks

Graphs · rational functions · asymptotes

Part (i) 4 marks

APPROACH

Factorising the denominator locates the vertical asymptotes directly; the degree of the numerator being lower than that of the denominator gives the horizontal asymptote. To fix the shape of the curve between the asymptotes we also find its turning points by differentiating.

Factorising the denominator,

$$3x^2 + 5x - 8 = (3x+8)(x-1)$$

so $y = \frac{2-x}{(3x+8)(x-1)}$ is undefined at $x = -\frac{8}{3}$ and at $x = 1$, giving the vertical asymptotes

$$x = -\frac{8}{3}, \quad x = 1$$

Since $\deg(\text{numerator}) = 1 < 2 = \deg(\text{denominator})$, $y \to 0$ as $x \to \pm\infty$, so

$$y = 0$$

is a horizontal asymptote.

Axis crossings: when $x = 0$, $y = \frac{2}{-8} = -\frac{1}{4}$, giving $\left(0,\ -\frac{1}{4}\right)$; when $y = 0$, $2 - x = 0$, giving $x = 2$, so $(2,\ 0)$.

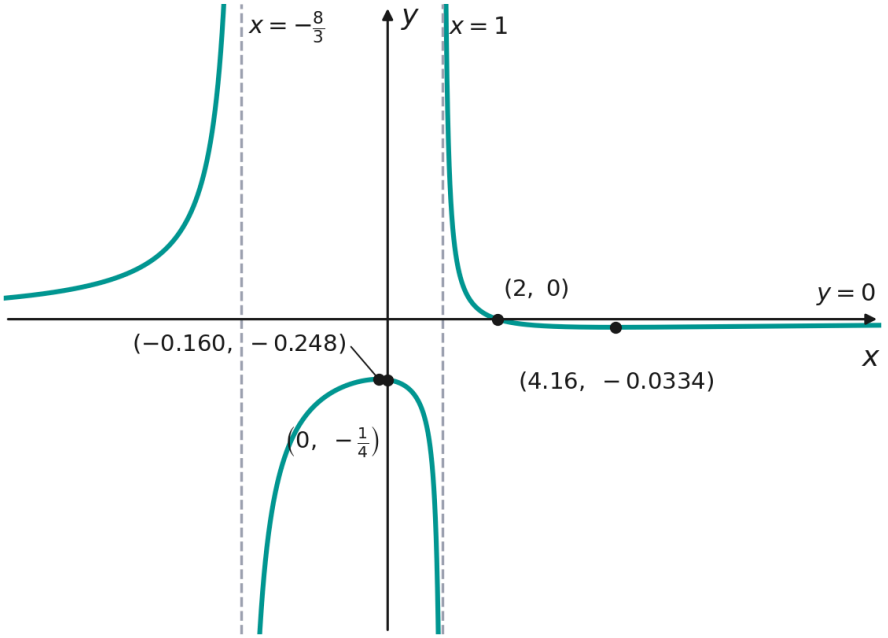
To know which way each branch turns, we differentiate by the quotient rule:

$$\frac{dy}{dx} = \frac{-\left(3x^2+5x-8\right)-(2-x)(6x+5)}{\left(3x^2+5x-8\right)^2}$$

Setting the numerator to 0 and solving (by GC), the turning points are at $x = 2 \pm \frac{\sqrt{42}}{3}$, i.e. at

$$(-0.160,\ -0.248)\ (\text{a maximum}), \qquad (4.16,\ -0.0334)\ (\text{a minimum})$$

both to 3 s.f. This gives three branches: one to the left of $x = -\frac{8}{3}$ rising from near $y = 0$ to $+\infty$; one between the asymptotes, from $-\infty$ up to the maximum $(-0.160, -0.248)$ and back down to $-\infty$; and one to the right of $x = 1$, falling from $+\infty$ through $(2, 0)$ to the minimum $(4.16, -0.0334)$ and rising back towards $y = 0$.



Part (ii) 1 mark

$$x < -\frac{8}{3} \quad \text{or} \quad 1 < x < 2$$

Explanation: reading off the sketch in (i), the curve lies above the x-axis exactly on the branch $x < -\frac{8}{3}$, and on the right-hand branch between the asymptote $x = 1$ and the intercept $x = 2$; everywhere else the curve lies on or below the x-axis.

Part (iii) 1 mark

> **APPROACH**
>
> the new expression is just the negative of the one in (ii), so its sign is reversed everywhere; we do not need to redo the sign analysis.

$$\frac{x-2}{3x^2+5x-8} = -\frac{2-x}{3x^2+5x-8}$$

so $\frac{x-2}{3x^2+5x-8} > 0$ exactly where $\frac{2-x}{3x^2+5x-8} < 0$, i.e. on the complement of the set found in (ii) (excluding the asymptotes and the intercept $x = 2$, where the expression is 0 or undefined).

To check your answer: at $x = 3$, $\frac{x-2}{3x^2+5x-8} = \frac{1}{34} > 0$, and $3 > 2$ is indeed in the solution set below. ✓

$\therefore -\frac{8}{3} < x < 1$ or $x > 2$.

Question 3 7 marks

Differentiation · maximum volume

Let r cm be the radius and h cm the height of the cylinder.

APPROACH

the volume depends on two variables linked by the fixed total surface area, so we use the surface area condition to write h in terms of r, substitute into V to get a function of r alone, and maximise using calculus.

The total surface area consists of the curved surface and the two circular ends:

$$2\pi r^2 + 2\pi rh = 900$$

Making h the subject,

$$h = \frac{900 - 2\pi r^2}{2\pi r} = \frac{450}{\pi r} - r$$

Substituting into $V = \pi r^2 h$,

$$\begin{aligned} & V \\ &= \pi r^2 \left(\frac{450}{\pi r} - r\right) \\ &= 450r - \pi r^3 \end{aligned}$$

Differentiating with respect to r,

$$\frac{dV}{dr} = 450 - 3\pi r^2$$

At a stationary point, $\frac{dV}{dr} = 0$:

$$\begin{aligned} 450 - 3\pi r^2 &= 0 \\ \Rightarrow r^2 &= \frac{150}{\pi} \\ \Rightarrow r &= \sqrt{\frac{150}{\pi}} \quad (r > 0, \text{ rejecting the negative root}) \end{aligned}$$

Since $\frac{d^2V}{dr^2} = -6\pi r < 0$ for all $r > 0$, this stationary point is a maximum.

Using $\pi r^2 = 150$,

$$\begin{aligned} & V_{max} \\ &= 450r - \pi r^3 \\ &= r\left(450 - \pi r^2\right) \\ &= r(450 - 150) \\ &= 300r \end{aligned}$$

thelearningcave.com.sg
Physics. Math. Chemistry.

Since $r = \sqrt{\frac{150}{\pi}} = \frac{5\sqrt{6}}{\sqrt{\pi}} = \frac{5\sqrt{6\pi}}{\pi}$,

$$V_{max} = \frac{1500\sqrt{6\pi}}{\pi}$$

For the ratio $r : h$, using $\frac{450}{\pi} = 3\left(\frac{150}{\pi}\right) = 3r^2$ (since $r^2 = \frac{150}{\pi}$),

$$h = \frac{450}{\pi r} - r = \frac{3r^2}{r} - r = 3r - r = 2r$$

To check your answer: $2\pi r^2 + 2\pi rh = 2\pi r^2 + 2\pi r(2r) = 6\pi r^2 = 6(150) = 900$, matching the given surface area. ✓

$\therefore$ the maximum possible volume is $V = \frac{1500\sqrt{6\pi}}{\pi}\ cm^3$ (equivalently $1500\sqrt{\frac{6}{\pi}}\ cm^3$, $\approx 2070\ cm^3$), attained when $r : h = 1 : 2$.

Question 4 11 marks

Maclaurin series · repeated differentiation

Part (i) 7 marks

Let $f(x) = \sec 2x$. Differentiating using the chain rule, with $\frac{d}{du}(\sec u) = \sec u \tan u$ and $u = 2x$,

$$f'(x) = 2\sec 2x \tan 2x$$

Differentiating again, using the product rule on $2\sec 2x \tan 2x$,

$$\begin{aligned} & f''(x) \\ &= 2\left[(2\sec 2x \tan 2x)\tan 2x + \sec 2x\left(2\sec^2 2x\right)\right] \\ &= 4\sec 2x \tan^2 2x + 4\sec^3 2x \end{aligned}$$

At $x = 0$: $\sec 0 = 1$ and $\tan 0 = 0$, so

$$f(0) = 1, \qquad f'(0) = 2(1)(0) = 0, \qquad f''(0) = 4(1)(0) + 4(1) = 4$$

Using the Maclaurin expansion $f(x) \approx f(0) + f'(0)x + \frac{f''(0)}{2!}x^2$,

$$\begin{aligned} f(x) &\approx 1 + 0 \cdot x + \frac{4}{2}x^2 \\ &= 1 + 2x^2 \end{aligned}$$

To check your answer: using the standard expansion $\sec u = 1 + \frac{1}{2}u^2 + \ldots$ with $u = 2x$ gives $\sec 2x = 1 + \frac{1}{2}(2x)^2 + \ldots = 1 + 2x^2 + \ldots$, matching the series above. ✓

$\therefore \sec 2x \approx 1 + 2x^2$, up to and including the term in x^2.

Part (ii)

> **APPROACH**
>
> replace $\sec 2x$ by the series just found and integrate the resulting polynomial term by term.

$$\begin{aligned}\int_0^{0.02} \sec 2x \, dx &\approx \int_0^{0.02} \left(1 + 2x^2\right) dx \\ &= \left[x + \frac{2}{3}x^3\right]_0^{0.02} \\ &= 0.02 + \frac{2}{3}(0.02)^3 \\ &= 0.0200053\ldots\end{aligned}$$

$\therefore \int_0^{0.02} \sec 2x \, dx \approx 0.02001$, correct to 5 decimal places.

Part (iii) 3 marks

By GC,

$$\int_0^{0.02} \sec 2x \, dx = 0.02001, \text{ correct to 5 decimal places}$$

Part (iv)

The estimate from (ii) and the calculator value from (iii) round to the same 0.02001 to 5 decimal places; the unrounded values agree to about 8 significant figures, differing only by roughly 2×10^{-9}.

This is because the Maclaurin series was truncated after the x^2 term, so the error in the approximation comes from the next term of the series, $\frac{10}{3}x^4$; integrating this over $[0,\ 0.02]$ contributes only $\frac{2}{3}(0.02)^5 \approx 2.13 \times 10^{-9}$ to the integral, which is negligible. Since $x = 0.02$ is very close to 0, i.e. close to the point about which the series was expanded, the truncated series gives an excellent approximation to the true value.

Part (v) 1 mark

$g(x) = cosec 2x = \frac{1}{\sin 2x}$ is not defined at $x = 0$, since $\sin 0 = 0$ makes the denominator 0; in fact $g(x) \to \pm\infty$ as $x \to 0$. A Maclaurin series requires $g(0)$ (and $g'(0)$, $g''(0)$, ...) to exist, so no Maclaurin series for g can be found.

Question 5 10 marks

Vectors · ratio theorem · areas

Part (i) 4 marks

APPROACH

X is the intersection of the two lines AC and BD. We write $\overrightarrow{OX}$ two ways — once as a point on AC, once as a point on BD — and compare coefficients of $\mathbf{a}$ and $\mathbf{b}$, since $\mathbf{a}$ and $\mathbf{b}$ are non-zero and non-parallel.

Since X lies on AC, for some $t \in \mathbb{R}$,

$$\begin{aligned}\overrightarrow{OX} &= \overrightarrow{OA} + t\overrightarrow{AC} \\ &= \mathbf{a} + t\left[(2\mathbf{a} + 4\mathbf{b}) - \mathbf{a}\right] \\ &= (1+t)\mathbf{a} + 4t\mathbf{b}\end{aligned}$$

Since X also lies on BD, for some $s \in \mathbb{R}$,

$$\begin{aligned}\overrightarrow{OX} &= \overrightarrow{OB} + s\overrightarrow{BD} \\ &= \mathbf{b} + s\left[(5\mathbf{a} + \mathbf{b}) - \mathbf{b}\right] \\ &= 5s\mathbf{a} + \mathbf{b}\end{aligned}$$

Comparing coefficients of $\mathbf{a}$ and $\mathbf{b}$,

$$1 + t = 5s, \qquad 4t = 1$$

Solving: $t = \frac{1}{4}$, $s = \frac{1}{4}$.

$$\therefore \overrightarrow{OX} = \frac{5}{4}\mathbf{a} + \mathbf{b}$$

Part (ii) 6 marks

APPROACH

O, X, Y collinear means $\overrightarrow{OY}$ is a scalar multiple of $\overrightarrow{OX}$. Writing Y on CD gives a second expression for $\overrightarrow{OY}$ in terms of $\mathbf{a}$ and $\mathbf{b}$; comparing coefficients pins down the scalar.

Since Y lies on CD, for some $\mu \in \mathbb{R}$,

$$\begin{aligned}\overrightarrow{OY} &= \overrightarrow{OC} + \mu\overrightarrow{CD} \\ &= (2\mathbf{a} + 4\mathbf{b}) + \mu\left[(5\mathbf{a} + \mathbf{b}) - (2\mathbf{a} + 4\mathbf{b})\right] \\ &= (2 + 3\mu)\mathbf{a} + (4 - 3\mu)\mathbf{b}\end{aligned}$$

Since O, X, Y are collinear, $\overrightarrow{OY} = \lambda\overrightarrow{OX}$ for some $\lambda \in \mathbb{R}$, i.e.

$$(2 + 3\mu)\mathbf{a} + (4 - 3\mu)\mathbf{b} = \frac{5\lambda}{4}\mathbf{a} + \lambda\mathbf{b}$$

Comparing coefficients of $\mathbf{a}$ and $\mathbf{b}$,

$$2+3\mu=\frac{5\lambda}{4}, \qquad 4-3\mu=\lambda$$

Solving simultaneously, $\mu=\frac{4}{9}$, $\lambda=\frac{8}{3}$.

$$\overrightarrow{OY}=\frac{8}{3}\left(\frac{5}{4}\mathbf{a}+\mathbf{b}\right)=\frac{10}{3}\mathbf{a}+\frac{8}{3}\mathbf{b}$$

To check your answer: with $\mu=\frac{4}{9}$, $2+3\mu=\frac{10}{3}$ and $4-3\mu=\frac{8}{3}$, matching the coefficients above. ✓

Since $\overrightarrow{OY}=\lambda\overrightarrow{OX}$ with $\lambda=\frac{8}{3}$, Y lies on ray OX beyond X, and

$$OX:OY=1:\lambda=1:\frac{8}{3}$$
$$\therefore\ OX:OY=3:8$$

Question 6 6 marks

Sampling · sampling methods

Part (i) 1 mark

These 22 clubs form a population.

Explanation: Alice's investigation concerns approaches to training by clubs in Division One specifically, so the group she wants to know about (her population of interest) is exactly those 22 clubs. Since she sends the questionnaire to all 22 of them, she has surveyed the entire population, not a sample of it.

Part (ii) 2 marks

Dilip should take a stratified sample of clubs, with the number of clubs sampled from each of the 4 divisions in proportion to the number of clubs in that division, rather than checking the facilities at all 100 clubs.

Explanation: checking every club's supporter facilities in detail would take far more time and resources than is needed to draw a reliable conclusion, so a sample is appropriate. Facilities are likely to differ systematically between divisions (e.g. clubs in a higher division may have larger budgets and better facilities), so a simple random sample could by chance under-represent a division; stratifying by division guarantees each division is represented in the sample in fair proportion.

Part (iii) 3 marks

> **APPROACH**
> choosing 5 clubs from each of the 4 divisions are 4 independent choices, so we multiply the number of ways of making each choice.

$$\text{No. of samples}=\binom{22}{5}\binom{24}{5}\binom{26}{5}\binom{28}{5}$$

By GC,

$$= 7\,236\,118\,179\,312\,422\,400$$

$$\therefore \text{ the number of different possible samples is } 7\,236\,118\,179\,312\,422\,400 \ \left(\approx 7.24 \times 10^{18}\right).$$

Question 7 11 marks

Binomial distribution · probabilities in context

Part (i) 2 marks

Two assumptions needed to model F by a binomial distribution:

1. Whether any one mug is faulty is independent of whether any other mug is faulty.
2. Every mug has the same, constant probability 0.08 of being faulty (e.g. the production process does not change over the period the sample is collected).

Part (ii) 2 marks

Let F be the number of faulty mugs found in a randomly chosen sample of 50. Then

$$F \sim B(50,\ 0.08)$$

Note: a binomial tail is evaluated by converting to cumulative form first.

$$P(F \geq 7) = 1 - P(F \leq 6)$$

By GC, $P(F \leq 6) = 0.898$ (unrounded 0.898128), so

$$P(F \geq 7) = 1 - 0.898128$$

$$= 0.102 \ (3 \text{ s.f.})$$

Part (iii) 2 marks

> **APPROACH**
>
> each of the 5 days of the week independently either does or does not have "at least 7 faulty mugs found", with probability $P(F \geq 7)$ found in (ii). This is itself a new binomial variable, counting a binomial number of "successful" days.

Let W be the number of days, out of a working week of 5 days, on which at least 7 faulty mugs are found. Since the number of faulty mugs on different days is independent, and each day has probability $P(F \geq 7) = 0.101872$ (unrounded value from (ii)) of being such a day,

$$W \sim B(5,\ 0.101872)$$

$$P(W \leq 2) = 0.991 \ (3 \text{ s.f.})$$

by GC (unrounded 0.990978).

Part (iv) 1 mark

Let S be the number of faulty saucers in a random sample of 10, so $S \sim B(10,\ p)$.

$$P(S=2)=\binom{10}{2}p^2(1-p)^8$$
$$=45p^2(1-p)^8$$

Part (v) 4 marks

APPROACH

a set has 2 mugs and 2 saucers, so the number of faulty items in a set is the sum of a $B(2,\,0.08)$ variable (faulty mugs) and an independent $B(2,\,p)$ variable (faulty saucers). "At most 1 faulty item" is the event that this sum is 0 or 1; we split into the cases 0 faulty mugs & 0 or 1 faulty saucer, or 1 faulty mug & 0 faulty saucers.

Let M be the number of faulty mugs and S' the number of faulty saucers in a set, so $M\sim B(2,\,0.08)$ and $S'\sim B(2,\,p)$ independently.

$$P(M=0)=0.92^2=0.8464,\qquad P(M=1)=2(0.08)(0.92)=0.1472$$
$$P\left(\text{at most 1 faulty item}\right)=P(M=0)\left[P(S'=0)+P(S'=1)\right]+P(M=1)\,P(S'=0)$$

Since $P(S'=0)+P(S'=1)=(1-p)^2+2p(1-p)=(1-p)(1+p)=1-p^2$,

$$0.8464\left(1-p^2\right)+0.1472(1-p)^2=0.97$$

Solving by GC, the equation has two roots, $p=0.0689$ or $p=-0.490$; since p is a probability, $0<p<1$,

$$p=-0.490\ \ (\text{rej. as } p>0)$$
$$\therefore\ p=0.0689\ (3\text{ s.f.})$$

To check your answer: with $p=0.0689$, $0.8464\left(1-p^2\right)+0.1472(1-p)^2=0.970$, matching the given 0.97. ✓

Question 8 9 marks

Probability · selections from a table

We first find the row and column totals of the table (each is the number of items of that character, or of that colour):

	Orange	Yellow	Green	White	Total
Horse	1	1	3	4	9
Rider	1	1	7	5	14
Dog	3	7	1	6	17
Bird	4	5	6	1	16
Total	9	14	17	16	56

Part (i)(a) 1 mark

$$P(\text{Horse or Rider}) = \frac{9+14}{56}$$
$$= \frac{23}{56}$$

Part (i)(b) 1 mark

> **APPROACH**
> "Dog or Bird but not White" removes the White items from the Dog and Bird rows, since a Dog or Bird item that is White is excluded.

$$P(\text{Dog or Bird, not White}) = \frac{(17-6)+(16-1)}{56}$$
$$= \frac{11+15}{56}$$
$$= \frac{26}{56}$$
$$= \frac{13}{28}$$

Part (ii)(a) 1 mark

> **APPROACH**
> choosing two items at random (without replacement) from the bag, we want both to lie in the "Horse, not Orange" cells, i.e. among the $9-1=8$ items that are a Horse and not Orange.

$$P(\text{both Horses, neither Orange}) = \frac{\binom{8}{2}}{\binom{56}{2}}$$
$$= \frac{28}{1540}$$
$$= \frac{1}{55}$$

Part (ii)(b) 3 marks

> **APPROACH**
> the two chosen items must together contain exactly one Dog and exactly one Yellow item. Since a single item can be both a Dog and Yellow at once (the Dog–Yellow cell, 7 items), we split by whether this double-counting cell is used: either one item is Dog-and-Yellow and the other is neither a Dog nor Yellow (this alone already gives exactly one Dog and one Yellow), or one item is a Dog that is not Yellow and the other is Yellow but not a Dog. Any other combination gives too many Dogs or too many Yellow items.

From the table: Dog-and-Yellow items number 7; Dog-but-not-Yellow items number $17-7=10$; Yellow-but-not-Dog items number $14-7=7$; and the remaining items (neither Dog nor Yellow) number $56-7-10-7=32$.

$$P(\text{exactly 1 Dog and exactly 1 Yellow}) = \frac{\binom{10}{1}\binom{7}{1} + \binom{7}{1}\binom{32}{1}}{\binom{56}{2}}$$
$$= \frac{70 + 224}{1540}$$
$$= \frac{294}{1540}$$
$$= \frac{21}{110}$$

To check your answer: $70 + 224 = 294$ favourable outcomes out of $\binom{56}{2} = 1540$ total, and $294/1540$ simplifies (dividing by 14) to $21/110$. ✓

Part (iii) 3 marks

APPROACH

if Gerri's two favourite combinations are the cells with counts m and n, choosing one item of each favourite (in either order) from the bag of 56 items gives probability $\frac{mn}{\binom{56}{2}}$. Setting this equal to $\frac{1}{77}$ pins down the product mn, which we then match against the counts actually in the table.

$$\frac{mn}{\binom{56}{2}} = \frac{1}{77}$$
$$mn = \frac{1540}{77}$$
$$mn = 20$$

The 16 cell counts in the table are $1, 3, 4, 5, 6, 7$ (with repeats). The only pair of these values with product 20 is 4×5.

The cells with count 4 are Horse–White and Bird–Orange; the cells with count 5 are Rider–White and Bird–Yellow. Gerri's two favourites are one cell from each list:

$$\therefore \text{ Gerri's two favourites are one of:}$$

- Horse (White) and Rider (White)
- Horse (White) and Bird (Yellow)
- Bird (Orange) and Rider (White)
- Bird (Orange) and Bird (Yellow)

Question 9 9 marks

Hypothesis testing · test of a mean

Part (i) 2 marks

Since the manager wants to find out whether the mean resistance of the resistors differs from 750 ohms, without any prior indication of the direction of any difference, a 2-tail test should be used.

Let μ ohms be the population mean resistance of the 750 ohm resistors.

$$H_0 : \mu = 750$$
$$H_1 : \mu \neq 750$$

Part (ii) 5 marks

We first find the sample mean.

$$\begin{aligned}\overline{x} &= \frac{742 + 771 + 768 + 738 + 769 + 752 + 742 + 766}{8} \\ &= \frac{6048}{8} \\ &= 756\end{aligned}$$

$\therefore$ the mean of the sample of 8 resistors is 756 ohms.

> **APPROACH**
>
> The population variance is given as $100\ ohms^2$, so we do not need to estimate it from the sample. Since the individual resistances are normally distributed, $\overline{X}$ is exactly normal for any sample size, and a Z-test is used directly (no Central Limit Theorem needed here).

Under H_0,

$$\overline{X} \sim N\left(750,\ \frac{100}{8}\right).$$

Standardising, the test statistic is

$$Z = \frac{\overline{X} - 750}{\sqrt{100/8}} \sim N(0,1) \text{ under } H_0.$$

Using $\overline{x} = 756$,

$$\begin{aligned}z &= \frac{756 - 750}{\sqrt{100/8}} \\ &= 1.70\ (3\ s.f.)\end{aligned}$$

By GC, carrying out a 1-sample Z-test with $\overline{x} = 756,\ n = 8,\ \sigma^2 = 100$ under $H_0 : \mu = 750$, the p-value is

$$p = 0.0897\ (3\ s.f.)$$

Since $p = 0.0897 > 0.05$, we do not reject H_0 at the 5% significance level.

To check your answer: $z = 1.70 < 1.960$, the 2-tail critical value at the 5% level, which agrees with $p > 0.05$. ✓

$\therefore$ there is insufficient evidence, at the 5% significance level, to conclude that the mean resistance of the 750 ohm resistors differs from 750 ohms.

Part (iii) 2 marks

Since nothing is known about the distribution of the resistances of the 1250 ohm resistors, the manager cannot assume they are normally distributed, so the test would need to differ in two ways.

A large sample (e.g. $n \geq 30$) would need to be taken, so that, by the Central Limit Theorem, the sample mean $\overline{X}$ is approximately normally distributed regardless of the underlying distribution of the individual resistances.

Since the population variance of the 1250 ohm resistors is also unknown, the unbiased estimate of the population variance,

$$s^2 = \frac{1}{n-1}\sum(x-\overline{x})^2,$$

computed from the sample, would be used in place of the population variance in the test statistic.

Question 10 13 marks

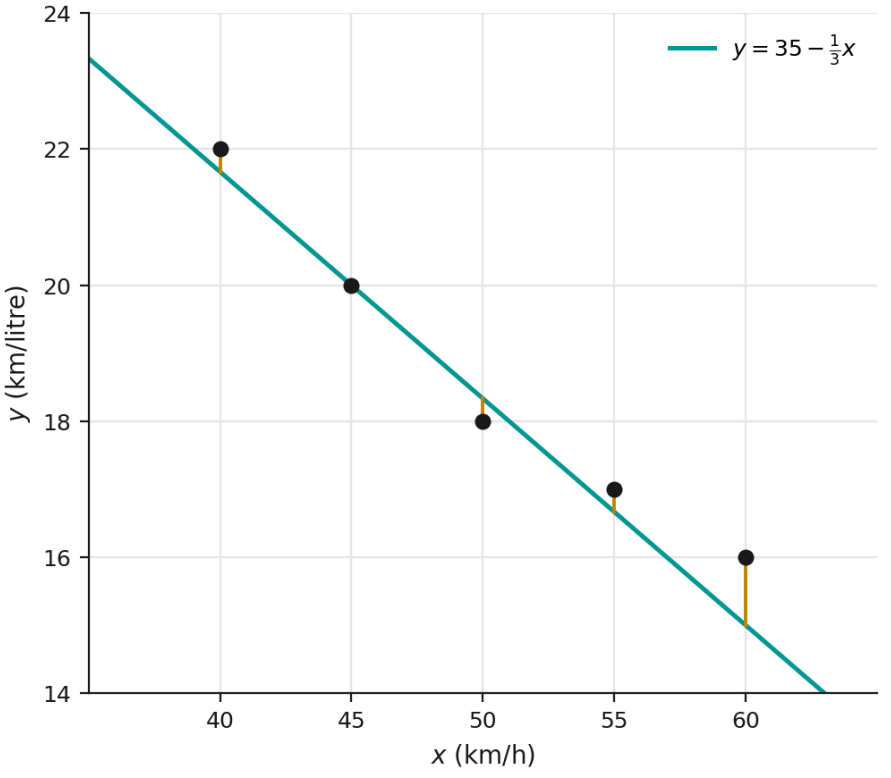
Correlation and regression · transformed models

Part (i)(a) 2 marks

The five data points and Abi's line $y = 35 - \frac{1}{3}x$ are plotted below.

Part (i)(b) 1 mark

For each point, the vertical segment down to (or up to) the line is the residual for that point, also shown on the diagram.



Part (i)(c) 1 mark

For each x, the value predicted by Abi's line is $\hat{y} = 35 - \frac{1}{3}x$, and each residual is $y - \hat{y}$.

x	40	45	50	55	60
$\hat{y}$ $y - \hat{y}$	$65/3$ $1/3$	20 0	$55/3$ $-1/3$	$50/3$ $1/3$	15 1

Sum of the squares of the residuals:

$$\left(\frac{1}{3}\right)^2 + 0^2 + \left(-\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + 1^2$$
$$= \frac{4}{3}$$

$\therefore$ the sum of the squares of the residuals for Abi's line is $\frac{4}{3}$.

Part (i)(d) 1 mark

Explanation: Residuals above the line are positive and residuals below the line are negative, so in the sum of the residuals these can cancel out, giving a small (or even zero) total even when individual points lie far from the line. Squaring first removes the signs, so the sum of squares only shrinks when every point genuinely lies close to the line, making it a meaningful measure of the overall goodness of fit.

Part (ii) 1 mark

Since Bhani's sum of squares of residuals, 1, is less than Abi's, $\frac{4}{3}$, Bhani's line gives a better fit to the data.

Part (iii) 1 mark

$\therefore$ the least squares regression line must pass through $(\overline{x},\ \overline{y})$.

Explanation: Here $\overline{x} = 50$ and $\overline{y} = 18.6$, so the line passes through $(50,\ 18.6)$.

Part (iv) 3 marks

By GC, the equation of the least squares regression line of y on x is

$$y = -0.3x + 33.6.$$

The product moment correlation coefficient is

$$r = -0.985\ (3\ s.f.).$$

Part (v) 2 marks

Using the regression line, at $x = 30$,

$$y = -0.3(30) + 33.6$$
$$= 24.6.$$

$\therefore$ the estimated fuel consumption at a speed of $30\ km/h$ is $24.6\ km/\text{litre}$.

Explanation: This value is not expected to be reliable, since $x = 30$ lies outside the range of the given data ($40 \leq x \leq 60$), so the estimate requires extrapolation, and there is no evidence that the linear model still holds at that speed.

Part (vi) 1 mark

$\therefore$ all of Cerie's data points lie exactly on her regression line, i.e. the points are perfectly collinear.

Explanation: A sum of squares of residuals of 0 means every residual is 0, so every data point lies exactly on the line.

Question 11 12 marks

Normal distribution · sums of normal variables

Let W grams be the mass of a randomly chosen white ball, so $W \sim N\left(110,\ 4^2\right)$.

Let B grams be the mass of a randomly chosen black ball, so $B \sim N\left(55,\ 2^2\right)$.

Part (i) 2 marks

Let $S = W_1 + W_2 + W_3 + W_4$ be the total mass of 4 randomly and independently chosen white balls.

$$E(S) = 4(110) = 440$$
$$Var(S) = 4\left(4^2\right) = 64$$
$$S \sim N(440,\ 64)$$

By GC,

$$P(S > 425) = 0.970\ (3\ s.f.)$$

Part (ii) 2 marks

Let $T = W + B$ be the total mass of a randomly chosen white ball and a randomly chosen black ball, chosen independently.

$$E(T) = 110 + 55 = 165$$
$$Var(T) = 4^2 + 2^2 = 20$$
$$T \sim N(165,\ 20)$$

By GC,

$$P(161 < T < 175) = 0.802\ (3\ s.f.)$$

Part (iii) 4 marks

Let $U = W_1 + W_2 + B_1 + B_2 + B_3$ be the total mass of 2 randomly and independently chosen white balls and 3 randomly and independently chosen black balls.

$$E(U) = 2(110) + 3(55) = 385$$
$$Var(U) = 2\left(4^2\right) + 3\left(2^2\right) = 44$$

$$U \sim N(385,\ 44)$$

We are given $P(U < M) = 0.271$. Standardising,

$$P\left(Z < \frac{M - 385}{\sqrt{44}}\right) = 0.271, \text{ where } Z \sim N(0, 1).$$

By GC (inverse normal),

$$\frac{M - 385}{\sqrt{44}} = -0.610\ (3\ s.f.)$$
$$M = 385 - 0.6098\sqrt{44}$$
$$M = 381\ (3\ s.f.)$$

To check your answer: since $0.271 < 0.5$, we expect $M < 385 = E(U)$, and indeed $M = 381 < 385$. ✓

$\therefore M = 381$ (to 3 s.f.).

Part (iv) 4 marks

APPROACH

Drilling the holes scales each ball's mass by a constant factor (not a fixed reduction), so the drilled ball masses are found by multiplying the original normal random variables W and B by constants; this keeps them normal, but scales both the mean and the variance.

Drilling reduces the mass of each black ball by 10%, so a drilled black ball has mass $0.9B$, where

$$E(0.9B) = 0.9(55) = 49.5$$
$$Var(0.9B) = 0.9^2\left(2^2\right) = 3.24$$
$$0.9B \sim N(49.5,\ 3.24)$$

Drilling reduces the mass of each white ball by 30%, so a drilled white ball has mass $0.7W$, where

$$E(0.7W) = 0.7(110) = 77$$
$$Var(0.7W) = 0.7^2\left(4^2\right) = 7.84$$
$$0.7W \sim N(77,\ 7.84)$$

Let R grams be the mass of a randomly chosen connecting rod, so $R \sim N\left(20,\ 0.9^2\right)$.

The model is made from 4 drilled black balls, 1 drilled white ball and 4 connecting rods, all chosen independently. Let X be the total mass, in grams, of the model.

$$X = (0.9B_1 + 0.9B_2 + 0.9B_3 + 0.9B_4) + 0.7W + (R_1 + R_2 + R_3 + R_4)$$
$$E(X) = 4(49.5) + 77 + 4(20)$$
$$= 355$$
$$Var(X) = 4(3.24) + 7.84 + 4(0.81)$$
$$= 24.04$$
$$X \sim N(355,\ 24.04)$$

By GC,

$$P(X > 350) = 0.846 \ (3 \ s.f.)$$

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