TEN-YEAR SERIES · WORKED SOLUTIONS

# 2015 A-Level H2 Mathematics Paper 2

12 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

Differential equations · modelling growth

## Part (i) 1 mark

$$h_{\max} = 32$$

Explanation: the tree stops growing once $\frac{dh}{dt} = 0$, i.e. once $16 - \frac{1}{2}h = 0$, which gives $h = 32$. Since $\frac{dh}{dt} > 0$ for $h < 32$, the height increases towards this value and the model gives no greater height, so the maximum height according to this model is $32\ m$.

## Part (ii) 5 marks

**APPROACH**

the differential equation gives $\frac{dh}{dt}$ as a function of $h$ alone, so we separate variables and integrate directly, using $\int (a+bx)^n\,dx = \frac{(a+bx)^{n+1}}{b(n+1)} + C$.

Since $\frac{dh}{dt} = \frac{1}{10}\sqrt{16 - \frac{1}{2}h}$,

$$\frac{dt}{dh} = \frac{10}{\sqrt{16 - \frac{1}{2}h}}$$

Integrating with respect to $h$,

$$\begin{aligned} t &= \int 10\left(16 - \frac{1}{2}h\right)^{-\frac{1}{2}}\,dh \\ &= 10 \times \frac{\left(16 - \frac{1}{2}h\right)^{\frac{1}{2}}}{-\frac{1}{2} \times \frac{1}{2}} + C \\ &= -40\sqrt{16 - \frac{1}{2}h} + C \end{aligned}$$

Using $h = 0$ when $t = 0$:

$$\begin{aligned} 0 &= -40\sqrt{16} + C \\ C &= 160 \\ t &= 160 - 40\sqrt{16 - \frac{1}{2}h} \end{aligned}$$

Hence, at $h = 16$ (half the maximum height),

$$\begin{aligned} t &= 160 - 40\sqrt{16 - 8} \\ &= 160 - 40\sqrt{8} \\ &= 160 - 80\sqrt{2} \end{aligned}$$

By GC, $t \approx 46.9$ (3 s.f.).

**To check your answer:** at $h = 32$ (the maximum height), the formula gives $t = 160 - 40\sqrt{0} = 160$, consistent with growth slowing to zero as $h \to 32$. ✓

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

$\therefore t = 160 - 40\sqrt{16 - \frac{1}{2}h}$, and the tree takes $160 - 80\sqrt{2} \approx 46.9$ years to reach half its maximum height.

## Question 2 10 marks

Vectors · lines · angles and distances

### Part (i) 2 marks

**APPROACH**

the angle between a line and a coordinate axis is found from the direction vector of the line and the unit vector along that axis, using the acute-angle (absolute-value) form of the angle formula.

$L$ has direction vector $\mathbf{d} = 2\mathbf{i} + 3\mathbf{j} - 6\mathbf{k}$, and the $x$-axis has direction $\mathbf{i}$.

$$|\mathbf{d}| = \sqrt{2^2 + 3^2 + (-6)^2} = \sqrt{49} = 7$$

Let $\theta$ be the acute angle between $L$ and the $x$-axis. Then

$$\cos\theta = \frac{|\mathbf{d} \cdot \mathbf{i}|}{|\mathbf{d}|\,|\mathbf{i}|}$$
$$= \frac{2}{7}$$

By GC, $\theta = 73.4°$ (1 d.p.).

### Part (ii) 5 marks

**APPROACH**

a general point on $L$ is $Q = (1 + 2\lambda,\ -2 + 3\lambda,\ -4 - 6\lambda)$. We set $\left|\overrightarrow{PQ}\right|^2 = 33$ and solve for $\lambda$. Since $\left|\overrightarrow{PQ}\right|^2$ is quadratic in $\lambda$, the closest point (its minimum) occurs at the value of $\lambda$ midway between the two roots, by symmetry of the parabola about its vertex — this is quicker than differentiating afresh.

Since $P = 2\mathbf{i} + 5\mathbf{j} - 6\mathbf{k}$,

$$\overrightarrow{PQ} = (2\lambda - 1)\mathbf{i} + (3\lambda - 7)\mathbf{j} + (2 - 6\lambda)\mathbf{k}$$
$$\left|\overrightarrow{PQ}\right|^2 = (2\lambda - 1)^2 + (3\lambda - 7)^2 + (2 - 6\lambda)^2$$
$$= 49\lambda^2 - 70\lambda + 54$$

Setting this equal to $33$:

$$49\lambda^2 - 70\lambda + 54 = 33$$
$$49\lambda^2 - 70\lambda + 21 = 0$$
$$7\lambda^2 - 10\lambda + 3 = 0$$

$$(7\lambda - 3)(\lambda - 1) = 0$$
$$\lambda = \frac{3}{7} \text{ or } \lambda = 1$$

Substituting back into the coordinates of $Q$:

$$\lambda = \frac{3}{7}: \quad Q = \left(\frac{13}{7},\ -\frac{5}{7},\ -\frac{46}{7}\right)$$
$$\lambda = 1: \quad Q = (3,\ 1,\ -10)$$

$\therefore$ the points on $L$ at distance $\sqrt{33}$ from $P$ are $\left(\frac{13}{7},\ -\frac{5}{7},\ -\frac{46}{7}\right)$ and $(3,\ 1,\ -10)$.

Hence, the point on $L$ closest to $P$ occurs at the value of $\lambda$ midway between these two roots:

$$\lambda = \frac{1}{2}\left(\frac{3}{7} + 1\right) = \frac{5}{7}$$
$$Q_{\min} = \left(1 + 2 \times \frac{5}{7},\ -2 + 3 \times \frac{5}{7},\ -4 - 6 \times \frac{5}{7}\right) = \left(\frac{17}{7},\ \frac{1}{7},\ -\frac{58}{7}\right)$$

**To check your answer:** at $\lambda = \frac{5}{7}$, $\left|\overrightarrow{PQ}\right|^2 = 49\left(\frac{5}{7}\right)^2 - 70\left(\frac{5}{7}\right) + 54 = 29$, and $29 < 33$, consistent with this being the closest point on $L$ to $P$. ✓

$\therefore$ the point on $L$ closest to $P$ is $\left(\frac{17}{7},\ \frac{1}{7},\ -\frac{58}{7}\right)$.

## Part (iii) 3 marks

**APPROACH**

a plane containing $L$ and $P$ is fixed by two directions lying in it: the direction of $L$, and the direction from a point on $L$ to $P$. Their cross product gives a normal to the plane.

Let $A = (1,\ -2,\ -4)$, a point on $L$. Then

$$\overrightarrow{AP} = (2-1)\mathbf{i} + (5-(-2))\,\mathbf{j} + (-6-(-4))\,\mathbf{k} = \mathbf{i} + 7\mathbf{j} - 2\mathbf{k}$$

A normal to the plane is

$$\mathbf{n} = \mathbf{d} \times \overrightarrow{AP}$$
$$= (3 \times (-2) - (-6) \times 7)\,\mathbf{i} + ((-6) \times 1 - 2 \times (-2))\,\mathbf{j} + (2 \times 7 - 3 \times 1)\mathbf{k}$$
$$= 36\mathbf{i} - 2\mathbf{j} + 11\mathbf{k}$$

Using $A$ as a point on the plane,

$$\mathbf{n} \cdot \overrightarrow{OA} = 36(1) - 2(-2) + 11(-4) = -4$$

**To check your answer:** substituting $P = (2,\ 5,\ -6)$: $36(2) - 2(5) + 11(-6) = 72 - 10 - 66 = -4$, matching the plane's constant, confirming $P$ lies in the plane. ✓

$\therefore$ a cartesian equation of the plane is $36x - 2y + 11z = -4$.

# Question 3 10 marks

Functions · inverse functions

## Part (a)(i) 2 marks

**APPROACH**

to show f has an inverse, we show f is one-one on its domain $x > 1$, by showing f is strictly monotonic there.

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

Differentiating,

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

For $x > 1$, $2x > 0$ and $(1-x^2)^2 > 0$, so $f'(x) > 0$ for all $x$ in the domain. Thus f is strictly increasing on $x > 1$, so f is one-one, and therefore f has an inverse. (shown)

## Part (a)(ii) 3 marks

Let $y = f(x) = \frac{1}{1-x^2}$.

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

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

$$x = \sqrt{1 - \frac{1}{y}} \quad (\text{taking the positive root, since } x > 1)$$

$\therefore f^{-1}(x) = \sqrt{1 - \frac{1}{x}}$.

The domain of $f^{-1}$ is the range of f. Since f is continuous and strictly increasing on $x > 1$, its range runs between the limiting values

$$\lim_{x \to 1^+} f(x) = -\infty, \quad \lim_{x \to \infty} f(x) = 0^-$$

so the range of f is $(-\infty, 0)$.

$\therefore$ the domain of $f^{-1}$ is $\{x \in \mathbb{R} : x < 0\}$.

**To check your answer:** $f(2) = \frac{1}{1-4} = -\frac{1}{3}$, and $f^{-1}\left(-\frac{1}{3}\right) = \sqrt{1 - \frac{1}{-1/3}} = \sqrt{1+3} = 2$, recovering the original input. ✓

## Part (b) 5 marks

**APPROACH**

to find the range of g algebraically, we set $y = g(x)$, rearrange into a quadratic in $x$, and use the condition that $x$ is real (discriminant $\geq 0$) to find the possible values of $y$.

Let $y = g(x) = \frac{2+x}{1-x^2}$.

$$y\left(1-x^2\right) = 2 + x$$
$$yx^2 + x + (2-y) = 0$$

Since $x$ is real, the discriminant of this quadratic in $x$ must be non-negative. (This also covers $y = 0$, where the equation reduces to the linear equation $x = -2$, which does give a real $x$.)

$$1 - 4y(2-y) \geq 0$$
$$4y^2 - 8y + 1 \geq 0$$

Solving $4y^2 - 8y + 1 = 0$,

$$y = \frac{8 \pm \sqrt{64-16}}{8} = \frac{8 \pm 4\sqrt{3}}{8} = \frac{2 \pm \sqrt{3}}{2}$$

Since the coefficient of $y^2$ is positive, $4y^2 - 8y + 1 \geq 0$ outside the roots:

$$y \leq \frac{2-\sqrt{3}}{2} \text{ or } y \geq \frac{2+\sqrt{3}}{2}$$

Note: the excluded values $x = 1$ and $x = -1$ never solve $yx^2 + x + (2-y) = 0$ for any $y$ — substituting $x = 1$ gives $0 = 3$, and $x = -1$ gives $0 = 1$, both false regardless of $y$ — so no further values of $y$ need to be excluded from the range.

$\therefore$ the range of g is $\left\{y \in \mathbb{R} : y \leq \frac{2-\sqrt{3}}{2} \text{ or } y \geq \frac{2+\sqrt{3}}{2}\right\}$.

## Question 4 14 marks

Sequences and series · mathematical induction

### Part (a) 6 marks

Let $P(n)$ be the statement

$$1 \times 3 \times 6 + 2 \times 4 \times 7 + 3 \times 5 \times 8 + \ldots + n(n+2)(n+5) = \frac{1}{12}n(n+1)\left(3n^2 + 31n + 74\right)$$

for $n \in \mathbb{Z}^+$.

When $n = 1$: LHS $= 1 \times 3 \times 6 = 18$. RHS $= \frac{1}{12}(1)(2)(3 + 31 + 74) = \frac{1}{12}(2)(108) = 18$. Since LHS = RHS, $P(1)$ is true.

Assume $P(k)$ is true for some $k \in \mathbb{Z}^+$, i.e.

$$\sum_{r=1}^{k} r(r+2)(r+5) = \frac{1}{12}k(k+1)\left(3k^2 + 31k + 74\right)$$

We want to show $P(k+1)$ is true, i.e.

$$\sum_{r=1}^{k+1} r(r+2)(r+5) = \frac{1}{12}(k+1)(k+2)\left(3(k+1)^2 + 31(k+1) + 74\right)$$

Now,

$$\sum_{r=1}^{k+1} r(r+2)(r+5) = \sum_{r=1}^{k} r(r+2)(r+5) + (k+1)(k+3)(k+6)$$
$$= \frac{1}{12}k(k+1)\left(3k^2+31k+74\right) + (k+1)(k+3)(k+6)$$
$$= \frac{1}{12}(k+1)\left[k\left(3k^2+31k+74\right) + 12(k+3)(k+6)\right]$$
$$= \frac{1}{12}(k+1)\left[3k^3+43k^2+182k+216\right]$$
$$= \frac{1}{12}(k+1)(k+2)\left(3k^2+37k+108\right)$$

Since $3(k+1)^2+31(k+1)+74 = 3k^2+37k+108$, this is

$$\sum_{r=1}^{k+1} r(r+2)(r+5) = \frac{1}{12}(k+1)(k+2)\left(3(k+1)^2+31(k+1)+74\right)$$

so $P(k+1)$ is true whenever $P(k)$ is true.

Since $P(1)$ is true, and $P(k)$ true $\Rightarrow P(k+1)$ true, by mathematical induction $P(n)$ is true for all $n \in \mathbb{Z}^+$. (shown)

## Part (b)(i) 1 mark

$$\frac{2}{4r^2+8r+3} = \frac{2}{(2r+1)(2r+3)}$$

**APPROACH**

write $\frac{2}{(2r+1)(2r+3)} = \frac{A}{2r+1} + \frac{B}{2r+3}$ and compare coefficients after multiplying through.

$$2 = A(2r+3) + B(2r+1)$$

Comparing coefficients of $r$: $2A+2B=0$, so $B=-A$. Comparing constants: $3A+B=2$, so $3A-A=2$, giving $A=1$, $B=-1$.

$\therefore A=1$, $B=-1$, so $\frac{2}{4r^2+8r+3} = \frac{1}{2r+1} - \frac{1}{2r+3}$. (shown)

## Part (b)(ii) 4 marks

**APPROACH**

the partial-fraction form telescopes; we write out the first few and last few terms to see which cancel.

$$S_n = \sum_{r=1}^{n}\left(\frac{1}{2r+1} - \frac{1}{2r+3}\right)$$
$$S_n = \left(\frac{1}{3}-\frac{1}{5}\right) + \left(\frac{1}{5}-\frac{1}{7}\right) + \left(\frac{1}{7}-\frac{1}{9}\right) + \ldots + \left(\frac{1}{2n+1}-\frac{1}{2n+3}\right)$$

All the middle terms cancel, leaving

$$S_n = \frac{1}{3} - \frac{1}{2n+3}$$

$$= \frac{2n+3-3}{3(2n+3)}$$
$$= \frac{2n}{3(2n+3)}$$

$\therefore S_n = \frac{1}{3} - \frac{1}{2n+3} = \frac{2n}{3(2n+3)}$.

### Part (b)(iii) 3 marks

**APPROACH**

as $n \to \infty$, $\frac{1}{2n+3} \to 0$, so the sum to infinity is $\frac{1}{3}$. We need the smallest $n$ for which $S_n$ is within $10^{-3}$ of this value, i.e. for which $S_\infty - S_n < 10^{-3}$.

$$S_\infty = \frac{1}{3}$$
$$S_\infty - S_n = \frac{1}{2n+3}$$

We require

$$\frac{1}{2n+3} < 10^{-3}$$
$$2n+3 > 1000$$
$$n > 498.5$$

By GC, checking the integers near this boundary: at $n = 498$, $\frac{1}{2(498)+3} = \frac{1}{999} \approx 0.001001 > 0.001$; at $n = 499$, $\frac{1}{2(499)+3} = \frac{1}{1001} \approx 0.000999 < 0.001$.

$\therefore$ the smallest value of $n$ is $499$.

## Question 5 4 marks

Sampling · sampling methods

### Part (i) 1 mark

Stratified sampling requires a sampling frame — a complete list of all the customers of the supermarket, divided into strata by age, from which a random sample can be drawn within each stratum. The manager has no such list, since customers are not registered members and simply walk in and out of the supermarket, so stratified sampling cannot be carried out.

### Part (ii) 2 marks

The manager first divides the customers into age groups (the strata), for example under 20, 20–39, 40–59, 60 and over. Using known information about the proportion of the supermarket's customers in each age group (e.g. from past records), the manager sets a quota — the number of

customers to be interviewed — for each age group. The manager then stands at the supermarket and interviews customers as they are encountered, asking each one's age group and cola preference, continuing within each age group until its quota is filled.

### Part (iii) 1 mark

Note: within each quota, the customers interviewed are not chosen at random — the manager (or interviewer) is free to choose whom to approach.

The interviewer is likely to approach customers who look more approachable or willing to stop and answer, rather than choosing at random from that age group. This introduces bias, so the customers interviewed within each age group need not be representative of all customers of that age group in the supermarket.

## Question 6 7 marks

Binomial distribution · normal approximation

### Part (i) 2 marks

Let $X$ be the number of red sweets in a small packet of 10 Droppers. Since each sweet is independently red with probability $0.25$,

$$X \sim B(10,\ 0.25).$$

We want $P(X \geq 4)$. Converting to cumulative form,

$$\begin{aligned} & P(X \geq 4) \\ \Rightarrow\ & 1 - P(X \leq 3) \\ = \ & 1 - 0.776 \\ = \ & 0.224 \ (3 \text{ s.f., by GC}). \end{aligned}$$

$\therefore P(X \geq 4) = 0.224.$

### Part (ii) 3 marks

Let $Y$ be the number of red sweets in a large packet of 100 Droppers, so $Y \sim B(100,\ 0.25)$.

> **APPROACH**
>
> With $n = 100$ trials, evaluating the binomial probability term by term is impractical, so we approximate $Y$ by a normal distribution. Since $n$ is large and $np = 25 > 5$ and $nq = 75 > 5$, the approximation is valid.

$$Y \sim N(25,\ 18.75) \text{ approximately,}$$

since $E(Y) = np = 25$ and $Var(Y) = \text{npq} = 100(0.25)(0.75) = 18.75$.

Note: $Y$ is discrete and we are approximating it by a continuous distribution, so a continuity correction is needed.

$$P(Y \geq 30)$$
$$\approx P(Y > 29.5)$$

Standardising, with $Z = \frac{Y-25}{\sqrt{18.75}}$,

$$= P\left(Z > \frac{29.5 - 25}{\sqrt{18.75}}\right)$$
$$= P(Z > 1.039)$$
$$= 0.149 \text{ (3 s.f., by GC)}.$$

**To check your answer:** the exact binomial value is $P(Y \geq 30) = 0.1495$ (by GC), close to the normal approximation $0.1493$. ✓

$\therefore P(Y \geq 30) = 0.149.$

### Part (iii) 2 marks

Let $W$ be the number of large packets, out of the 15 Yip buys, that contain at least 30 red sweets. Using the unrounded probability from part (ii), $p = P(Y \geq 30) = 0.14935$, and since the packets are independent of each other,

$$W \sim B(15,\ 0.14935).$$

We want $P(W \leq 3)$:

$$P(W \leq 3) = 0.825 \text{ (3 s.f., by GC)}.$$

$\therefore P(W \leq 3) = 0.825.$

## Question 7 7 marks

Poisson distribution · assumptions · probabilities

### Part (i) 2 marks

Two assumptions needed, in context:

1. Errors occur independently of one another, i.e. an error occurring at one point on a page does not affect the probability of an error occurring at any other point on that page (or on another page).
2. Errors occur at a constant average rate of $1.3$ per page throughout the newspaper, i.e. the mean number of errors per page is uniform across all pages and does not vary, for example, from the front page to the later pages.

### Part (ii) 3 marks

Let $T$ be the total number of errors on the first 6 pages of the newspaper on a given day. Since the number of errors per page is $Po(1.3)$ and errors on different pages are independent, the sum of 6 independent $Po(1.3)$ variables is itself Poisson, with mean $6 \times 1.3$:

$$T \sim Po(7.8).$$

We want $P(T > 10)$. Converting to cumulative form,

$$\begin{aligned} & P(T > 10) \\ \Rightarrow\ & 1 - P(T \leq 10) \\ = \ & 1 - 0.835 \\ = \ & 0.165 \ (3 \text{ s.f., by GC}). \end{aligned}$$

$\therefore P(T > 10) = 0.165.$

### Part (iii) 2 marks

Let $T_n$ be the total number of errors on the first $n$ pages of the newspaper, so $T_n \sim Po(1.3n)$.

We are given $P\left(T_n < 2\right) < 0.05$, i.e.

$$\begin{aligned} P\left(T_n = 0\right) + P\left(T_n = 1\right) &< 0.05 \\ e^{-1.3n} + 1.3n\, e^{-1.3n} &< 0.05 \\ e^{-1.3n}(1 + 1.3n) &< 0.05. \end{aligned}$$

**APPROACH**

this inequality cannot be solved for $n$ algebraically, since $n$ appears both in the exponent and outside it. As $n$ must also be a positive integer (a number of pages), we instead evaluate the left side by GC for successive integer values of $n$ and find where it first drops below $0.05$ .

By GC:

$$\begin{aligned} n = 3: \ e^{-3.9}(1 + 3.9) &= 0.0992 > 0.05 \\ n = 4: \ e^{-5.2}(1 + 5.2) &= 0.0342 < 0.05. \end{aligned}$$

Since $e^{-1.3n}(1 + 1.3n)$ is decreasing in $n$, the least value of $n$ satisfying the inequality is $n = 4$.

$\therefore$ the least possible value of $n$ is $4$.

## Question 8 7 marks

Hypothesis testing · one-tailed test of a mean

Let $X$ be the mass, in kg, of a pineapple from the stall, and let $n = 8$ be the sample size.

**Finding the unbiased estimates**

Summarising the sample: $\sum x = 7.06$ and $\sum x^2 = 6.2696$.

$$\overline{x} = \frac{\sum x}{n} = \frac{7.06}{8} = 0.8825$$

$$s^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right) = \frac{1}{7}\left(6.2696 - \frac{7.06^2}{8}\right) = 0.00559 \text{ (3 s.f.)}$$

$\therefore$ the unbiased estimate of the population mean is $\overline{x} = 0.8825\ kg$, and the unbiased estimate of the population variance is $s^2 = 0.00559\ kg^2$ (3 s.f.).

**Setting up the test**

Let $\mu$ be the population mean mass, in kg, of pineapples sold at the stall. The stall owner's claim is $\mu \geq 0.9$, so we test whether the sample gives evidence against this at the $10\%$ level of significance.

$$H_0 : \mu = 0.9$$
$$H_1 : \mu < 0.9$$

where the test is carried out at the $10\%$ significance level.

Note: since the population variance is unknown and is estimated from a small sample ($n = 8 < 30$), we need the population of pineapple masses to be normally distributed (as given) to use a $t$-test.

> **APPROACH**
>
> the alternative hypothesis is one-tailed ("doubt the claim that the mean is at least $0.9$" means checking whether the mean is actually less than $0.9$), so this is a lower-tailed test.

Under $H_0$, the test statistic is

$$T = \frac{\overline{X} - 0.9}{S/\sqrt{n}} \sim t(n-1),$$

where $\overline{X}$ is the sample mean and $S^2$ is the unbiased estimate of the population variance.

$$t = \frac{0.8825 - 0.9}{\sqrt{0.00559/8}}$$
$$= -0.662 \text{ (3 s.f., by GC)}.$$

By GC, the corresponding one-tailed $p$-value is

$$p\text{-value} = P(T \leq -0.662) = 0.265 \text{ (3 s.f.)}.$$

**To check your answer:** the critical value for a one-tailed test at the $10\%$ level with 7 degrees of freedom is $t = -1.415$; since $-0.662 > -1.415$, the test statistic does not fall in the critical region, consistent with $p = 0.265 > 0.10$. ✓

Since $p = 0.265 > 0.10$, we do not reject $H_0$.

$\therefore$ there is insufficient evidence, at the $10\%$ level of significance, to conclude that the mean mass of pineapples is less than $0.9\ kg$, so there is no evidence to doubt the stall owner's claim.

# Question 9 8 marks

Probability · independent and combined events

## Part (i) 1 mark

Since $A$ and $B$ are independent,

$$P(B \mid A) = P(B) = 0.4$$

## Part (ii) 3 marks

**APPROACH**

we want $P(A' \cap B' \cap C')$, the probability that none of the three events occur. It is easiest to find this as $1 - P(A \cup B \cup C)$, and $P(A \cup B \cup C)$ is found by inclusion-exclusion. Since $A, B$ are independent, $A, C$ are independent, and now $B, C$ are independent too, every pairwise intersection is just a product of probabilities.

$$P(A \cap B) = P(A)P(B) = 0.45 \times 0.4 = 0.18$$
$$P(A \cap C) = P(A)P(C) = 0.45 \times 0.3 = 0.135$$
$$P(B \cap C) = P(B)P(C) = 0.4 \times 0.3 = 0.12$$

By inclusion-exclusion,

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C) \\ &= 0.45 + 0.4 + 0.3 - 0.18 - 0.135 - 0.12 + 0.1 \\ &= 0.815 \end{aligned}$$

Note: even though every pair of events is independent here, $P(A \cap B \cap C) = 0.1 \neq P(A)P(B)P(C) = 0.054$, so the three events are not mutually independent — the given value of $0.1$ has to be used as it stands, not replaced by a product.

$$\begin{aligned} P(A' \cap B' \cap C') &= 1 - P(A \cup B \cup C) \\ &= 1 - 0.815 \\ &= 0.185 \end{aligned}$$

$\therefore P(A' \cap B' \cap C') = 0.185.$

## Part (iii) 4 marks

**APPROACH**

with $B$ and $C$ no longer independent, $P(B \cap C)$ is not pinned down by $P(B)$ and $P(C)$ alone; it can take a range of values, and $P(A)$, $P(A \cap B \cap C) = 0.1$, and the independence of $A$ with $B$ and with $C$ stay fixed throughout. We split the sample space into the seven regions of the Venn diagram, express each region's probability in terms of $x = P(B \cap C)$, and require every region to be non-negative — this pins down the range of $x$, and hence of $P(A' \cap B' \cap C')$.

Since $A, B$ are independent and $A, C$ are independent, $P(A \cap B) = 0.18$ and $P(A \cap C) = 0.135$ as before, and $P(A \cap B \cap C) = 0.1$ still. So

$$P(A \cap B \cap C') = P(A \cap B) - P(A \cap B \cap C) = 0.18 - 0.1 = 0.08$$

$$P(A \cap B' \cap C) = P(A \cap C) - P(A \cap B \cap C) = 0.135 - 0.1 = 0.035$$

$$P\,(A \text{ only}) = P(A) - P(A \cap B) - P(A \cap C) + P(A \cap B \cap C) = 0.45 - 0.18 - 0.135 + 0.1 = 0.235$$

Let $x = P(B \cap C)$. Then

$$P(A' \cap B \cap C) = x - 0.1$$

$$P\,(B \text{ only}) = P(B) - P(A \cap B) - x + P(A \cap B \cap C) = 0.4 - 0.18 - x + 0.1 = 0.32 - x$$

$$P\,(C \text{ only}) = P(C) - P(A \cap C) - x + P(A \cap B \cap C) = 0.3 - 0.135 - x + 0.1 = 0.265 - x$$

Each of these seven regional probabilities must be non-negative. $P(A' \cap B \cap C) = x - 0.1 \geq 0$ gives $x \geq 0.1$; $P\,(B \text{ only}) = 0.32 - x \geq 0$ gives $x \leq 0.32$; $P\,(C \text{ only}) = 0.265 - x \geq 0$ gives $x \leq 0.265$. Together,

$$0.1 \leq x \leq 0.265$$

Summing all seven regions,

$$\begin{aligned} P(A \cup B \cup C) &= 0.235 + (0.32 - x) + (0.265 - x) + 0.08 + 0.035 + (x - 0.1) + 0.1 \\ &= 0.935 - x \end{aligned}$$

so

$$\begin{aligned} P(A' \cap B' \cap C') &= 1 - (0.935 - x) \\ &= 0.065 + x \end{aligned}$$

This is increasing in $x$, so it is greatest when $x$ is greatest and least when $x$ is least.

At $x = 0.265$: $P(A' \cap B' \cap C') = 0.065 + 0.265 = 0.33$. At $x = 0.1$: $P(A' \cap B' \cap C') = 0.065 + 0.1 = 0.165$.

**To check your answer:** at $x = 0.1$, $P(A' \cap B \cap C) = 0$, i.e. $B \cap C$ occurs only together with $A$ — a valid boundary case; at $x = 0.265$, $P\,(C \text{ only}) = 0$, another valid boundary case. Both extremes correspond to one region vanishing, as expected at the edge of a feasible range. ✓

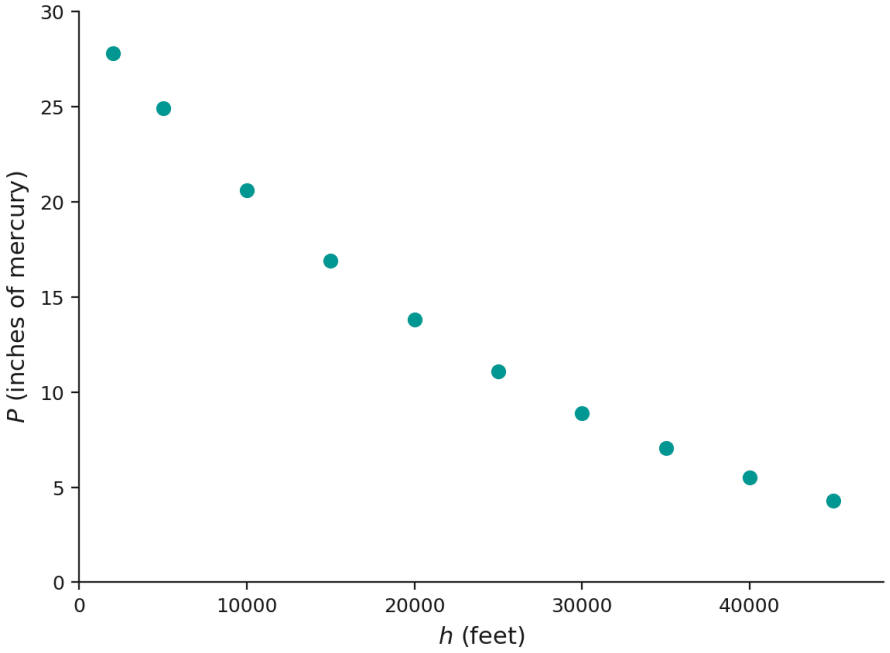
∴ the greatest possible value of $P(A' \cap B' \cap C')$ is $0.33$, and the least possible value is $0.165$.

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

# Question 10 9 marks

Correlation and regression · transformed models

## Part (i) 1 mark




## Part (ii) 3 marks

By GC, entering the ten pairs of values and computing the product moment correlation coefficient for each transformation of $h$:

## Part (a)

$r = -0.9807$

## Part (b)

$r = -0.9748$

## Part (c) 3 marks

$r = -0.9986$

## Part (iii) 2 marks

**APPROACH**

the "most appropriate" model is the one whose PMCC is closest to $-1$, since a value of $r$ nearer $\pm 1$ indicates a stronger linear relationship between the transformed variable and $P$ — and only a variable that is (nearly) linearly related to $P$ should be used for linear regression.

Physics. Math. Chemistry.

Since $|-0.9986|$ is closest to $1$ among the three values in part **(ii)**, the model using $\sqrt{h}$ and $P$ is the most appropriate.

By GC, the regression line of $P$ on $\sqrt{h}$ is

$$P = 34.789 - 0.146865\sqrt{h}$$

∴ the equation which best models air pressure at different heights is $P = 34.8 - 0.147\sqrt{h}$ (3 s.f.), where $h$ is in feet and $P$ is in inches of mercury.

### Part (iv)

**APPROACH**

the equation in part **(iii)** takes $h$ in feet. Since $1$ metre $=3.28$ feet, a height of $h_m$ metres corresponds to $h = 3.28h_m$ feet, and we substitute this into the equation from part **(iii)**.

Let $h_m$ be the height in metres, so $h = 3.28h_m$.

$$\begin{aligned} P &= 34.789 - 0.146865\sqrt{3.28h_m} \\ &= 34.789 - 0.146865\sqrt{3.28}\sqrt{h_m} \\ &= 34.789 - 0.265984\sqrt{h_m} \end{aligned}$$

∴ $P = 34.8 - 0.266\sqrt{h_m}$ (3 s.f.), where $h_m$ is the height in metres.

## Question 11 9 marks

Permutations and combinations · repeated letters

The word CABBAGES has $8$ letters: $A$ (twice), $B$ (twice), and $C$, $G$, $E$, $S$ (once each).

### Part (i) 2 marks

$$\begin{aligned} \text{no. of arrangements} &= \frac{8!}{2!\,2!} \\ &= 10080 \end{aligned}$$

∴ the number of different arrangements is $10080$.

### Part (ii) 1 mark

$$10080$$

Explanation: among all $10080$ arrangements, exactly one has its letters in alphabetical order — namely AABBCEGS itself. Every other arrangement has the letters out of order, so the number of arrangements not in alphabetical order is $10080 - 1 = 10079$.

∴ the number of arrangements not in alphabetical order is $10079$.

## Part (iii) 2 marks

**APPROACH**

treat the two A's as a single block $[AA]$ and the two B's as a single block $[BB]$. This leaves $6$ distinguishable entities to arrange in a row: $[AA]$, $[BB]$, $C$, $G$, $E$, $S$, with no repeats among them.

$$\begin{aligned}\text{no. of arrangements} &= 6! \\ &= 720\end{aligned}$$

$\therefore$ the number of arrangements with both A's together and both B's together is $720$.

## Part (iv) 4 marks

**APPROACH**

we want to exclude arrangements where the two A's are adjacent, and arrangements where the two B's are adjacent (C, G, E, S each occur only once, so they can never repeat adjacently). This is a "not both" count, done by inclusion-exclusion on the two events "A's together" and "B's together".

If the A's are together, treat them as one block $[AA]$: this leaves 7 entities — $[AA]$, $B$, $B$, $C$, $G$, $E$, $S$ — with the two B's still repeated.

$$\begin{aligned}\text{no. with A}'\text{s together} &= \frac{7!}{2!} \\ &= 2520\end{aligned}$$

By the same argument with the roles of $A$ and $B$ swapped,

$$\begin{aligned}\text{no. with B}'\text{s together} &= \frac{7!}{2!} \\ &= 2520\end{aligned}$$

From part **(iii)**, the number of arrangements with both pairs together is $720$.

By inclusion-exclusion,

$$\begin{aligned}\text{no. with A}'\text{s together or B}'\text{s together} &= 2520 + 2520 - 720 \\ &= 4320 \\ \text{no. with no two adjacent letters the same} &= 10080 - 4320 \\ &= 5760\end{aligned}$$

**To check your answer:** a direct computer enumeration of all $10080$ distinct arrangements of CABBAGES, counting those with no two equal letters adjacent, also gives $5760$. ✓

$\therefore$ the number of arrangements with no two adjacent letters the same is $5760$.

# Question 12 9 marks

Normal distribution · sums of normal variables

## Part (i) 2 marks

Let $X$ be the mass, in grams, of a randomly chosen apple, so $X \sim N\left(300, 20^2\right)$.

Let $T = X_1 + X_2 + X_3 + X_4 + X_5$ be the total mass of 5 randomly chosen apples, where $X_1, \ldots, X_5$ are independent and each distributed as $X$.

$$E(T) = 5(300) = 1500$$
$$Var(T) = 5\left(20^2\right) = 2000$$

so $T \sim N(1500, 2000)$.

Standardising,

$$P(T > 1600) = P\left(Z > \frac{1600 - 1500}{\sqrt{2000}}\right)$$
$$= 0.0127 \text{ (3 s.f., by GC)}.$$

$\therefore P(T > 1600) = 0.0127.$

## Part (ii) 3 marks

Let $Y$ be the mass, in grams, of a randomly chosen pear, so $Y \sim N\left(200, 15^2\right)$.

Let $S = Y_1 + \cdots + Y_8$ be the total mass of 8 randomly chosen pears, where $Y_1, \ldots, Y_8$ are independent and each distributed as $Y$.

$$E(S) = 8(200) = 1600$$
$$Var(S) = 8\left(15^2\right) = 1800$$

so $S \sim N(1600, 1800)$.

**APPROACH**

the question asks for $P(T > S)$, a comparison between two sums. We handle this by working with the single random variable $D = T - S$, whose distribution follows from $T$ and $S$ both being normal.

Let $D = T - S$, the total mass of the 5 apples minus the total mass of the 8 pears.

$$E(D) = E(T) - E(S) = 1500 - 1600 = -100$$
$$Var(D) = Var(T) + Var(S) = 2000 + 1800 = 3800$$

Note: $Var(D) = Var(T) + Var(S)$, not $Var(T) - Var(S)$, since $T$ and $S$ are independent.

so $D \sim N(-100, 3800)$.

$$P(T > S) = P(D > 0)$$
$$= 0.0524 \text{ (3 s.f., by GC)}.$$

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

$\therefore P(T > S) = 0.0524.$

## Part (iii) 4 marks

**APPROACH**

peeling and coring an apple scales its mass by a constant factor $0.85$ (a reduction of $15\%$), and a pear's mass by $0.9$ (a reduction of $10\%$). Since a linear transformation of a normal random variable is normal, the prepared masses are still normally distributed, with mean and variance scaled accordingly — the variance scales by the square of the factor.

Let $X' = 0.85X$ be the mass, in grams, of a prepared apple, and $Y' = 0.9Y$ be the mass, in grams, of a prepared pear.

$$E(X') = 0.85(300) = 255, \qquad Var(X') = 0.85^2\left(20^2\right) = 289$$

so $X' \sim N(255, 289)$.

$$E(Y') = 0.9(200) = 180, \qquad Var(Y') = 0.9^2\left(15^2\right) = 182.25$$

so $Y' \sim N(180, 182.25)$.

Let $U = X_1' + \cdots + X_5' + Y_1' + \cdots + Y_8'$ be the total prepared mass of $5$ apples and $8$ pears, where all $13$ masses are independent.

$$E(U) = 5(255) + 8(180) = 1275 + 1440 = 2715$$
$$Var(U) = 5(289) + 8(182.25) = 1445 + 1458 = 2903$$

so $U \sim N(2715, 2903)$.

$$P(U < 2750) = P\left(Z < \frac{2750 - 2715}{\sqrt{2903}}\right)$$
$$= 0.742 \text{ (3 s.f., by GC)}.$$

**To check your answer:** $E(U) = 2715$ is already close to and below $2750$, and $2750$ is under one standard deviation ($\sqrt{2903} \approx 53.9$) above the mean, so a probability just above $0.5$ (here $0.742$) is a sensible size. ✓

$\therefore P(U < 2750) = 0.742.$