TEN-YEAR SERIES · WORKED SOLUTIONS

# 2022 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 4 marks

Integration · Techniques of Integration

Using the substitution $u = \sqrt{x+2}$, so $u^2 = x+2$, giving $x = u^2 - 2$ and $dx = 2u\,du$,

$$\begin{aligned}
&\int \frac{x}{\sqrt{x+2}}\,dx \\
&= \int \frac{u^2-2}{u} \times 2u\,du \\
&= \int 2\left(u^2-2\right)\,du \\
&= 2\left(\frac{u^3}{3} - 2u\right) + c \\
&= \frac{2u^3}{3} - 4u + c
\end{aligned}$$

Substituting back $u = \sqrt{x+2}$,

$$\int \frac{x}{\sqrt{x+2}}\,dx = \frac{2}{3}(x+2)^{3/2} - 4\sqrt{x+2} + c$$

Factoring out $\sqrt{x+2}$,

$$\begin{aligned}
&= \sqrt{x+2}\left[\frac{2}{3}(x+2) - 4\right] + c \\
&= \sqrt{x+2} \times \frac{2(x+2)-12}{3} + c \\
&= \sqrt{x+2} \times \frac{2x-8}{3} + c \\
&= \frac{2}{3}(x-4)\sqrt{x+2} + c
\end{aligned}$$

**To check your answer:** differentiating $\frac{2}{3}(x-4)\sqrt{x+2}$ by the product rule gives $\frac{2}{3}\sqrt{x+2} + \frac{2}{3}(x-4) \times \frac{1}{2\sqrt{x+2}} = \frac{2(x+2)+(x-4)}{3\sqrt{x+2}} = \frac{3x}{3\sqrt{x+2}} = \frac{x}{\sqrt{x+2}}$ ✓

$\therefore \int \frac{x}{\sqrt{x+2}}\,dx = \frac{2}{3}(x-4)\sqrt{x+2} + c.$

## Question 2 7 marks

Integration · Definite integrals

### Part (a) 1 mark

The volume of the container is base area $\times$ height $= 0.9h\ m^3$. Since $1\ m^3 = 1000$ litres, this is $900h$ litres. Water is poured in at a constant rate of $k$ litres per second and fills the container in $72$ seconds, so

$$k \times 72 = 900h$$

$$\Rightarrow k = \frac{900h}{72} = 12.5h \ \text{(shown)}$$

## Part (b) 3 marks

Let $T$ be the time taken to fill the container. Water is now poured in at rate $kt$ litres per second, so the volume poured in by time $T$ is

$$\int_0^T kt\,dt = \left[\frac{kt^2}{2}\right]_0^T = \frac{kT^2}{2}$$

Setting this equal to the container's volume, $900h$, and substituting $k = 12.5h$,

$$\frac{kT^2}{2} = 900h$$
$$\Rightarrow \frac{12.5hT^2}{2} = 900h$$
$$\Rightarrow 6.25T^2 = 900 \ \ (\text{dividing by } h, \text{ since } h \neq 0)$$
$$\Rightarrow T^2 = 144$$
$$\Rightarrow T = 12 \ \ (\text{taking the positive root, since } T > 0)$$

**To check your answer:** with $T = 12$, $\frac{kT^2}{2} = \frac{144k}{2} = 72k = 72 \times 12.5h = 900h$, matching the container's volume ✓

$\therefore$ the container fills in $T = 12\ s$.

## Part (c) 3 marks

Water is now poured in at rate $(kt + 25)$ litres per second and fills the container in $10$ seconds, so the volume poured in is

$$\int_0^{10} (kt + 25)\,dt = \left[\frac{kt^2}{2} + 25t\right]_0^{10} = 50k + 250$$

Setting this equal to $900h$ and substituting $k = 12.5h$,

$$50k + 250 = 900h$$
$$\Rightarrow 50(12.5h) + 250 = 900h$$
$$\Rightarrow 625h + 250 = 900h$$
$$\Rightarrow 250 = 275h$$
$$\Rightarrow h = \frac{250}{275} = \frac{10}{11}$$

**To check your answer:** with $h = \frac{10}{11}$, $k = 12.5 \times \frac{10}{11} = \frac{125}{11}$, so $50k + 250 = \frac{6250}{11} + \frac{2750}{11} = \frac{9000}{11}$, and $900h = 900 \times \frac{10}{11} = \frac{9000}{11}$ — these match ✓

$\therefore h = \frac{10}{11}\ m$.

# Question 3 9 marks

Complex Numbers · Polar and exponential form

## Part (a) 3 marks

Since $z_1 = 3 - i\sqrt{3}$,

$$|z_1| = \sqrt{3^2 + \left(\sqrt{3}\right)^2} = \sqrt{12} = 2\sqrt{3}$$

$z_1$ has positive real part and negative imaginary part, so it lies in the fourth quadrant, giving

$$\arg z_1 = -\tan^{-1}\left(\frac{\sqrt{3}}{3}\right) = -\tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = -\frac{\pi}{6}$$

Also $z_2 = \frac{1}{2}e^{i\frac{2\pi}{5}}$, so $|z_2| = \frac{1}{2}$ and $\arg z_2 = \frac{2\pi}{5}$. Since $z_3 = z_1 \times z_2$,

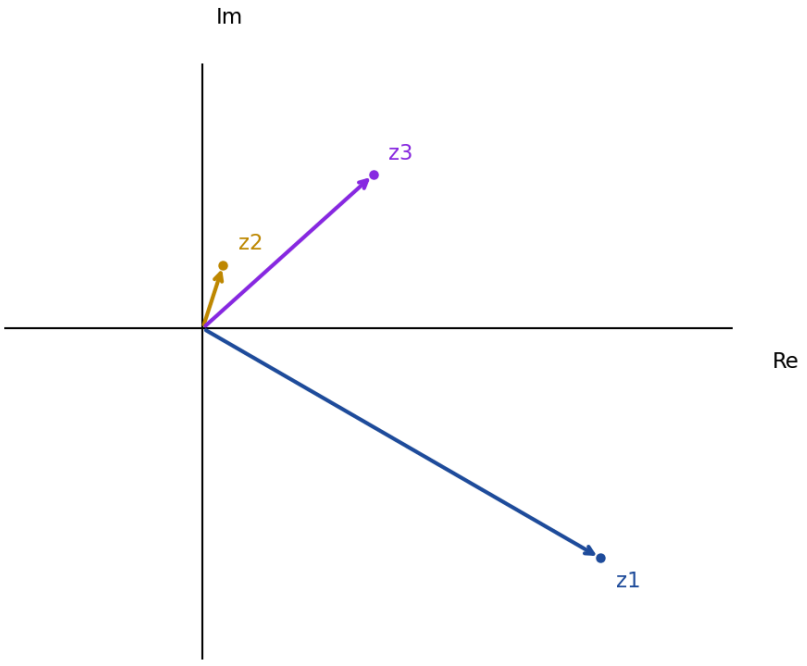
$$|z_3| = |z_1|\,|z_2| = 2\sqrt{3} \times \frac{1}{2} = \sqrt{3}$$

$$\arg z_3 = \arg z_1 + \arg z_2 = -\frac{\pi}{6} + \frac{2\pi}{5} = -\frac{5\pi}{30} + \frac{12\pi}{30} = \frac{7\pi}{30}$$

**To check your answer:** computing $z_1 z_2$ directly on the GC gives $|z_3| \approx 1.732$, matching $\sqrt{3} \approx 1.732$, and $\arg z_3 \approx 0.733$ rad, matching $\frac{7\pi}{30} \approx 0.733$ rad ✓

$\therefore |z_3| = \sqrt{3}$ and $\arg z_3 = \frac{7\pi}{30}$.

## Part (b) 2 marks



## Part (c) 4 marks

**APPROACH**

$z_3^n$ is purely imaginary exactly when $\arg(z_3^n) = n\arg z_3$ is an odd multiple of $\frac{\pi}{2}$. Since $\arg z_3$ is a fraction of $\pi$ with denominator $30$, this condition becomes a linear congruence in $n$ modulo $30$, which we solve using the modular inverse of $7$.

$z_3^n$ is purely imaginary when

$$n \times \frac{7\pi}{30} = \frac{\pi}{2} + k\pi \quad \text{for some integer } k$$

Dividing by $\pi$ and clearing denominators,

$$\frac{7n}{30} = \frac{1}{2} + k$$
$$\Rightarrow 7n = 15 + 30k$$
$$\Rightarrow 7n \equiv 15 \pmod{30}$$

Since $\gcd(7, 30) = 1$, $7$ has an inverse modulo $30$: as $7 \times 13 = 91 = 90 + 1 \equiv 1 \pmod{30}$, the inverse is $13$. Multiplying both sides by $13$,

$$n \equiv 15 \times 13 \pmod{30}$$
$$\Rightarrow n \equiv 195 \pmod{30}$$
$$\Rightarrow n \equiv 15 \pmod{30}$$
$$n = 15$$

**To check your answer:** $7 \times 15 = 105 = 3 \times 30 + 15$, so $105 \equiv 15 \pmod{30}$ ✓, confirming $n = 15$ satisfies $7n \equiv 15 \pmod{30}$

For this value of $n$,

$$\left|z_3^{15}\right| = |z_3|^{15} = \left(\sqrt{3}\right)^{15} = 3^{15/2} = 3^7\sqrt{3} = 2187\sqrt{3}$$
$$\arg\left(z_3^{15}\right) = 15 \times \frac{7\pi}{30} = \frac{7\pi}{2}$$

Reducing to the principal range $(-\pi, \pi]$ by subtracting $2\pi$ twice,

$$\frac{7\pi}{2} - 2\pi = \frac{3\pi}{2}, \quad \frac{3\pi}{2} - 2\pi = -\frac{\pi}{2}$$

**To check your answer:** computing $z_3^{15}$ directly on the GC gives modulus $\approx 3788.0$, matching $2187\sqrt{3} \approx 3788.0$, and argument $\approx -1.5708$ rad, matching $-\frac{\pi}{2} \approx -1.5708$ rad — a purely negative-imaginary number, confirming $z_3^{15}$ is purely imaginary ✓

$\therefore n = 15$, $\left|z_3^{15}\right| = 2187\sqrt{3}$ and $\arg\left(z_3^{15}\right) = -\frac{\pi}{2}$.

# Question 4 9 marks

Sequences and Series · Sigma Notation

## Part (a) 2 marks

We factorise the denominator first. Solving $9r^2+3r-2=0$ gives $r=\frac{-3\pm\sqrt{9+72}}{18}=\frac{-3\pm9}{18}$, i.e. $r=\frac{1}{3}$ or $r=-\frac{2}{3}$, so

$$9r^2+3r-2=(3r-1)(3r+2)$$

Using partial fractions,

$$\frac{1}{(3r-1)(3r+2)}=\frac{A}{3r-1}+\frac{B}{3r+2}$$
$$\Rightarrow 1=A(3r+2)+B(3r-1)$$

Setting $r=\frac{1}{3}$: $1=A(1+2)=3A\Rightarrow A=\frac{1}{3}$.

Setting $r=-\frac{2}{3}$: $1=B(-2-1)=-3B\Rightarrow B=-\frac{1}{3}$.

**To check your answer:** at $r=1$, $\frac{1}{9+3-2}=\frac{1}{10}$, and $\frac{1}{3}\left(\frac{1}{2}-\frac{1}{5}\right)=\frac{1}{3}\times\frac{3}{10}=\frac{1}{10}$ ✓

$\therefore\ \frac{1}{9r^2+3r-2}=\frac{1}{3}\left(\frac{1}{3r-1}-\frac{1}{3r+2}\right)$

## Part (b) 4 marks

**APPROACH**

The sum runs from $r=m$ to $r=3m$, not from $r=1$. We telescope directly over this range, writing out the first and last few brackets so the cancellation pattern is visible, and keep only the surviving first and last terms.

Using the partial fractions from (a),

$$\sum_{r=m}^{3m}\frac{1}{9r^2+3r-2}=\frac{1}{3}\sum_{r=m}^{3m}\left(\frac{1}{3r-1}-\frac{1}{3r+2}\right)$$

Writing out the terms:

$$\frac{1}{3}\left[\left(\frac{1}{3m-1}-\frac{1}{3m+2}\right)+\left(\frac{1}{3m+2}-\frac{1}{3m+5}\right)+\cdots+\left(\frac{1}{9m-1}-\frac{1}{9m+2}\right)\right]$$

Since $-\frac{1}{3r+2}=-\frac{1}{3(r+1)-1}$ cancels with the $+\frac{1}{3(r+1)-1}$ of the next bracket, every middle term cancels, leaving only the first and last term. Combining these into a single fraction,

$$\sum_{r=m}^{3m}\frac{1}{9r^2+3r-2}=\frac{1}{3}\left(\frac{1}{3m-1}-\frac{1}{9m+2}\right)$$
$$=\frac{1}{3}\times\frac{(9m+2)-(3m-1)}{(3m-1)(9m+2)}$$
$$=\frac{1}{3}\times\frac{6m+3}{(3m-1)(9m+2)}$$

$$= \frac{2m+1}{(3m-1)(9m+2)}$$

**To check your answer:** at $m = 1$, direct summation over $r = 1, 2, 3$ gives $\frac{1}{10} + \frac{1}{40} + \frac{1}{88} = \frac{3}{22}$, matching $\frac{2(1)+1}{(3(1)-1)(9(1)+2)} = \frac{3}{22}$ ✓

$\therefore \sum_{r=m}^{3m} \frac{1}{9r^2+3r-2} = \frac{2m+1}{(3m-1)(9m+2)}$

## Part (c) 1 mark

$$\sum_{r=1}^{\infty} \frac{1}{9r^2+3r-2} = \frac{1}{6}$$

Explanation: using the same telescoping as in (b), but starting the sum at $r = 1$ instead of $r = m$,

$$\sum_{r=1}^{N} \frac{1}{9r^2+3r-2} = \frac{1}{3}\left(\frac{1}{2} - \frac{1}{3N+2}\right)$$

As $N \to \infty$, $\frac{1}{3N+2} \to 0$, so the sum tends to $\frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$.

## Part (d) 2 marks

Writing $S(n) = \sum_{r=1}^{n} \frac{1}{9r^2+3r-2}$, part (c) gives $S(n) = \frac{1}{3}\left(\frac{1}{2} - \frac{1}{3n+2}\right) = \frac{n}{2(3n+2)}$. The difference between the sum to infinity and the sum to $n$ terms is

$$\begin{aligned} \frac{1}{6} - S(n) &= \frac{1}{6} - \frac{n}{2(3n+2)} \\ &= \frac{(3n+2) - 3n}{6(3n+2)} \\ &= \frac{2}{6(3n+2)} \\ &= \frac{1}{9n+6} \end{aligned}$$

We require this to be less than $0.004$. Since $\frac{1}{0.004} = 250$,

$$\begin{aligned} &\frac{1}{9n+6} < 0.004 \\ \Rightarrow\ &9n+6 > 250 \\ \Rightarrow\ &n > \frac{244}{9} \end{aligned}$$

Since $\frac{244}{9} = 27.11\ldots$ and $n$ is an integer, the smallest value is $n = 28$.

**To check your answer:** at $n = 27$, the difference is $\frac{1}{9(27)+6} = \frac{1}{249} \approx 0.00402$, which is not less than $0.004$; at $n = 28$, it is $\frac{1}{258} \approx 0.00388$, which is ✓

$\therefore$ the smallest value of $n$ is $28$.

# Question 5 11 marks

Integration · Volume of revolution

## Part (a) 5 marks

Rotating the shaded region between the circle $x^2 + y^2 = r^2$ and the line $y = r - h$ about the $y$-axis, the cross-section at height $y$ is a disc of radius $x$, where $x^2 = r^2 - y^2$. The spherical cap is generated for $y$ from $r - h$ to $r$:

$$\begin{aligned} V &= \int_{r-h}^{r} \pi x^2 \, dy \\ &= \pi \int_{r-h}^{r} \left(r^2 - y^2\right) dy \\ &= \pi \left[ r^2 y - \frac{y^3}{3} \right]_{r-h}^{r} \end{aligned}$$

Evaluating the limits, and expanding $(r-h)^3 = r^3 - 3r^2h + 3rh^2 - h^3$,

$$\begin{aligned} V &= \pi \left( r^3 - \frac{r^3}{3} \right) - \pi \left( r^2(r-h) - \frac{(r-h)^3}{3} \right) \\ &= \frac{2\pi r^3}{3} - \pi r^3 + \pi r^2 h + \frac{\pi (r-h)^3}{3} \\ &= -\frac{\pi r^3}{3} + \pi r^2 h + \frac{\pi r^3}{3} - \pi r^2 h + \pi r h^2 - \frac{\pi h^3}{3} \\ &= \pi r h^2 - \frac{\pi h^3}{3} \\ &= \frac{1}{3}\pi h^2 (3r - h) \end{aligned}$$

**To check your answer:** at $h = r$ (the cap is a full hemisphere), $V = \frac{1}{3}\pi r^2(3r - r) = \frac{2\pi r^3}{3}$, matching the known hemisphere volume ✓

$\therefore\ V = \frac{1}{3}\pi h^2 (3r - h)$ (shown)

## Part (b) 4 marks

**APPROACH**

Apply the part (a) formula twice, once to the top cap (height $p$) and once to the bottom cap (height $3p$), with $r = 15$ throughout, and set (sphere) - (top cap) - (bottom cap) equal to the given ornament volume.

The sphere has volume $\frac{4}{3}\pi(15)^3 = 4500\pi$. The top cap has height $p$, so its volume is $\frac{1}{3}\pi p^2(45 - p)$; the bottom cap has height $3p$, so its volume is $\frac{1}{3}\pi(3p)^2(45 - 3p) = 3\pi p^2(45 - 3p)$. The ornament volume is

$$4500\pi - \frac{1}{3}\pi p^2(45 - p) - 3\pi p^2(45 - 3p) = 3402\pi$$

Dividing by $\pi$ and multiplying by $3$,

$$13500 - p^2(45 - p) - 9p^2(45 - 3p) = 10206$$
$$\Rightarrow 13500 - 45p^2 + p^3 - 405p^2 + 27p^3 = 10206$$
$$\Rightarrow 28p^3 - 450p^2 + 3294 = 0$$

Dividing by 2,

$$14p^3 - 225p^2 + 1647 = 0$$

Testing $p = 3$: $14(27) - 225(9) + 1647 = 378 - 2025 + 1647 = 0$ ✓, so $(p - 3)$ is a factor:

$$14p^3 - 225p^2 + 1647 = (p - 3)\left(14p^2 - 183p - 549\right)$$

The remaining quadratic has discriminant $183^2 + 4(14)(549) = 64233$, which is not a perfect square, giving $p = \frac{183 \pm \sqrt{64233}}{28} \approx -2.52$ or $15.59$. Both are rejected: we need $0 < p$ and $3p < 15$ (so that the bottom cap does not exceed the sphere), i.e. $0 < p < 5$.

**To check your answer:** with $p = 3$, top cap $= \frac{1}{3}\pi(9)(42) = 126\pi$, bottom cap $= \frac{1}{3}\pi(81)(18) = 486\pi$; ornament $= 4500\pi - 126\pi - 486\pi = 3402\pi$ ✓

$\therefore$ $14p^3 - 225p^2 + 1647 = 0$, and $p = 3$.

## Part (c) 2 marks

**APPROACH**

The two ornaments share both face radii but not the same volume. Each face radius corresponds to two possible cutting heights ($y = \pm\text{value}$), so the same pair of radii can be produced by cutting planes on either the same or opposite sides of the centre. The part (b) ornament uses opposite sides (spanning most of the sphere); the smaller one must use the same side instead.

With $p = 3$: the top face of the part (b) ornament is at $y = 15 - 3 = 12$ (radius $\sqrt{15^2 - 12^2} = 9$). The bottom cap has height $3p = 9$, so the bottom face is at $y = -(15 - 9) = -6$ (radius $\sqrt{15^2 - 6^2} = 3\sqrt{21}$). This ornament spans $y = -6$ to $y = 12$, a height of $18$.

A face of radius $9$ occurs only at $y = 12$ or $y = -12$; a face of radius $3\sqrt{21}$ occurs only at $y = 6$ or $y = -6$. For a *smaller* solid with a top above a bottom, the only remaining valid pairing is $y = 12$ (top) and $y = 6$ (bottom), giving a height of just $6$:

$$V = \int_6^{12} \pi\left(15^2 - y^2\right) dy$$
$$= \pi\left[225y - \frac{y^3}{3}\right]_6^{12}$$
$$= \pi\left[(2700 - 576) - (1350 - 72)\right]$$
$$= \pi(2124 - 1278)$$
$$= 846\pi$$

**To check your answer:** $846\pi < 3402\pi$ ✓, confirming a smaller volume, as required.

$\therefore\ V = 846\pi\ cm^3.$

## Question 6 6 marks

Probability · Conditional Probability

### Part (a) 3 marks

**APPROACH**

Anil's first throw only sets a target number; it cannot itself win. From then on the two players alternate attempts to match the number thrown immediately before them: Babs's 1st attempt, then Anil's 1st attempt, then Babs's 2nd, and so on. Each attempt is independent, succeeding with probability $\frac{1}{6}$ and failing with probability $\frac{5}{6}$. Anil wins exactly when the first success in this alternating sequence falls on one of *his* attempts.

Anil wins on his $n$th attempt ($n = 1, 2, 3, \ldots$) if Babs's 1st through $n$th attempts all fail, Anil's 1st through $(n-1)$th attempts all fail, and Anil's $n$th attempt succeeds — a total of $2n-1$ failures followed by one success:

$$P\left(\text{Anil wins on his } n\text{th attempt}\right) = \left(\frac{5}{6}\right)^{2n-1}\frac{1}{6}$$

Summing over all $n$ (a geometric series with first term $\frac{5}{6}\times\frac{1}{6} = \frac{5}{36}$ and common ratio $\left(\frac{5}{6}\right)^2 = \frac{25}{36}$),

$$\begin{aligned} P\left(\text{Anil wins}\right) &= \sum_{n=1}^{\infty}\left(\frac{5}{6}\right)^{2n-1}\frac{1}{6} \\ &= \frac{5/36}{1-25/36} \\ &= \frac{5/36}{11/36} \\ &= \frac{5}{11} \end{aligned}$$

**To check your answer:** by the same method, $P\left(\text{Babs wins}\right) = \sum_{m=1}^{\infty}\left(\frac{5}{6}\right)^{2(m-1)}\frac{1}{6} = \frac{1/6}{1-25/36} = \frac{6}{11}$, and $\frac{5}{11}+\frac{6}{11} = 1$, confirming the game ends with probability 1 and the two results are consistent ✓

$\therefore\ P\left(\text{Anil wins}\right) = \frac{5}{11}$

### Part (b) 3 marks

**APPROACH**

We need $P\left(\text{Babs wins on her 2nd throw} \mid \text{Babs wins}\right)$, so we need both the probability that Babs wins specifically on her 2nd attempt, and $P\left(\text{Babs wins}\right)$ from (a).

thelearningcave.com.sg
Physics. Math. Chemistry.

Babs wins on her $m$th attempt ($m = 1, 2, 3, \ldots$) if Babs's 1st through $(m-1)$th attempts all fail, Anil's 1st through $(m-1)$th attempts all fail, and Babs's $m$th attempt succeeds — a total of $2(m-1)$ failures followed by one success:

$$P\left(\text{Babs wins on her } m\text{th attempt}\right) = \left(\frac{5}{6}\right)^{2(m-1)} \frac{1}{6}$$

For $m = 2$ (Babs's second throw is her second attempt),

$$\begin{aligned} P\left(\text{Babs wins on her 2nd attempt}\right) &= \left(\frac{5}{6}\right)^{2} \times \frac{1}{6} \\ &= \frac{25}{216} \end{aligned}$$

From (a), $P\left(\text{Babs wins}\right) = \frac{6}{11}$. By conditional probability,

$$\begin{aligned} P\left(\text{Babs wins on 2nd throw} \mid \text{Babs wins}\right) &= \frac{P\left(\text{Babs wins on 2nd attempt}\right)}{P\left(\text{Babs wins}\right)} \\ &= \frac{25/216}{6/11} \\ &= \frac{25}{216} \times \frac{11}{6} \\ &= \frac{275}{1296} \end{aligned}$$

**To check your answer:** $\frac{275}{1296} \approx 0.212$; since Babs's chance of winning on a given attempt decreases geometrically, this should be a little less than $P\left(\text{Babs wins on 1st attempt} \mid \text{Babs wins}\right) = \frac{1/6}{6/11} = \frac{11}{36} \approx 0.306$ — and indeed it is ✓

$\therefore\ P\left(\text{Babs wins on her 2nd throw} \mid \text{Babs wins}\right) = \frac{275}{1296}$

## Question 7 8 marks

Permutations and Combinations · Basic counting

### Part (a) 1 mark

The 53 employees who replied form a sample, not the population.

Explanation: the population is the entire group under study — here, all $75$ employees in the Staffing department, since the questionnaire was sent to every one of them. The $53$ who replied are only part of that group (a self-selecting subset, since not everyone asked chose to respond), so they are a sample of the department, not the department itself.

### Part (b) 2 marks

She should take a random sample of employees — for example a stratified sample, grouping employees by department and then taking a simple random sample within each department, in proportion to that department's size.

Explanation: she wants to know the views of employees across the *whole* company but cannot survey everyone, so a random sample is needed to give every employee a known, non-zero chance of selection and so avoid bias. Stratifying by department further ensures each department is fairly represented, so the results can be reliably generalised to the whole company's workforce.

## Part (c) 5 marks

> **APPROACH**
>
> Let $a, p, m, s$ be the numbers chosen from Administration, Production, Marketing and Staffing respectively, so $a+p+m+s=8$ with $1 \le a \le 7,\ 1 \le p \le 6,\ 1 \le m \le 4,\ 1 \le s \le 3$ (at least one from each department) and $a > p,\ a > m,\ a > s$ (Administration strictly the largest). Since $p+m+s \ge 3$, we have $a \le 5$; since $a$ must exceed each of $p, m, s \ge 1$, $a \ge 2$. We work through $a = 5, 4, 3, 2$ in turn.

**Case $a=5$:** then $p+m+s=3$ with each $\ge 1$ and $<5$, forcing $p=m=s=1$:

$$\binom{7}{5}\binom{6}{1}\binom{4}{1}\binom{3}{1} = 21 \times 6 \times 4 \times 3$$
$$= 1512$$

**Case $a=4$:** then $p+m+s=4$ with each $\ge 1$ and $<4$, so $\{p, m, s\}$ is $2, 1, 1$ in some order:

$$\binom{6}{2}\binom{4}{1}\binom{3}{1} + \binom{6}{1}\binom{4}{2}\binom{3}{1} + \binom{6}{1}\binom{4}{1}\binom{3}{2}$$
$$= 180 + 108 + 72$$
$$= 360$$

Multiplying by $\binom{7}{4} = 35$ for the choice of Administration members,

$$35 \times 360 = 12600$$

**Case $a=3$:** then $p+m+s=5$ with each $\ge 1$ and $\le 2$, so $\{p, m, s\}$ is $2, 2, 1$ in some order:

$$\binom{6}{2}\binom{4}{2}\binom{3}{1} + \binom{6}{2}\binom{4}{1}\binom{3}{2} + \binom{6}{1}\binom{4}{2}\binom{3}{2}$$
$$= 270 + 180 + 108$$
$$= 558$$

Multiplying by $\binom{7}{3} = 35$,

$$35 \times 558 = 19530$$

**Case $a=2$:** then $p+m+s=6$ with each $\ge 1$ and $=1$ only (since each $<2$), but $1+1+1=3 \neq 6$, so this case is impossible and contributes $0$.

Adding all cases,

$$1512 + 12600 + 19530 + 0 = 33642$$

**To check your answer:** a direct enumeration over every valid $(a, p, m, s)$ with $a+p+m+s=8$ in the stated ranges and $a > p,\ a > m,\ a > s$, summing $\binom{7}{a}\binom{6}{p}\binom{4}{m}\binom{3}{s}$ over each, gives $33642$, matching this calculation ✓

∴ the number of ways is $33642$.

## Question 8 9 marks

Normal Distribution · Properties of a Normal Distribution

### Part (a) 2 marks

$$aX - bY \sim N\left(ap - bs,\ a^2q^2 + b^2t^2\right)$$

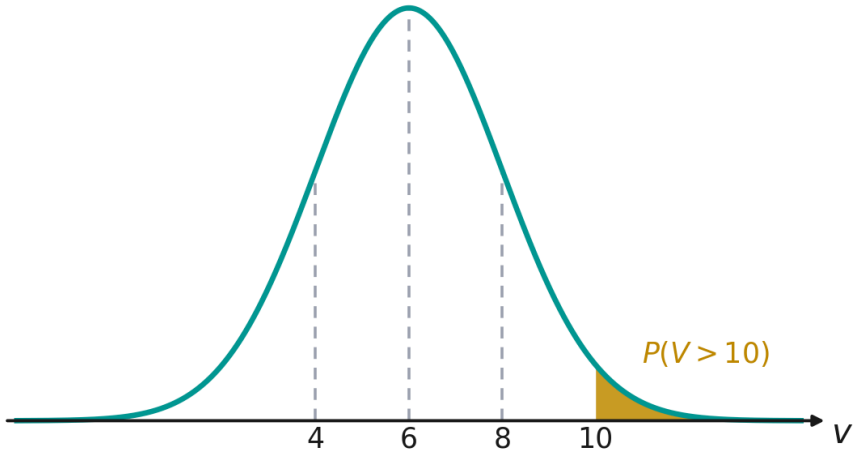
Explanation: since $X$ and $Y$ are independent normal random variables, any linear combination of them is also normal. The mean of $aX - bY$ is $a\,E(X) - b\,E(Y) = ap - bs$, and since $X$ and $Y$ are independent, $Var(aX - bY) = a^2Var(X) + b^2Var(Y) = a^2q^2 + b^2t^2$.

### Part (b)(i) 2 marks

Let $V \sim N(\mu, 4)$, since the standard deviation is $2$ gives variance $2^2 = 4$. We are told $P(V > 8) = P(V < 4)$. By symmetry of the normal curve about its mean $\mu$, the values $4$ and $8$ must be equidistant from $\mu$, so $\mu$ is their midpoint:

$$\mu = \frac{4 + 8}{2} = 6$$

So $V \sim N(6, 4)$.




∴ $V \sim N(6, 4)$, sketched above with the shaded region for **(b)(ii)**.

### Part (b)(ii) 2 marks

$$P(V > 10) = 0.0228 \ \ (3\text{ s.f.})$$

Explanation: standardising, with $\mu = 6$ and $\sigma = 2$,

$$Z = \frac{10 - 6}{2} = 2$$

$$P(V > 10) = P(Z > 2) = 1 - \Phi(2)$$

By GC, $\Phi(2) = 0.97725$, so $1 - \Phi(2) = 0.02275 \approx 0.0228$ (3 s.f.).

**To check your answer:** $10$ is $2$ standard deviations above the mean of $6$, so $P(V > 10)$ should be small — $0.0228$ is consistent with this. ✓

## Part (c) 3 marks

**APPROACH**

We first use the given relationship between the mean and variance of $W$ to find $p$, then use this value of $p$ to compute the required probability directly from the two individual terms of the binomial distribution.

$W \sim B(8, p)$, so $E(W) = 8p$ and $Var(W) = 8p(1-p)$. Since $E(W) = 1.2\,Var(W)$:

$$8p = 1.2 \times 8p(1-p)$$

Dividing both sides by $8p$ (since $p \neq 0$),

$$\begin{aligned} 1 &= 1.2(1-p) \\ \Rightarrow 1 - p &= \frac{1}{1.2} = \frac{5}{6} \\ \Rightarrow p &= \frac{1}{6} \end{aligned}$$

So $W \sim B\left(8, \frac{1}{6}\right)$.

$$\begin{aligned} P(W < 2) &= P(W = 0) + P(W = 1) \\ &= \binom{8}{0}\left(\frac{5}{6}\right)^8 + \binom{8}{1}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^7 \\ &= \frac{1015625}{1679616} \\ &= 0.605 \ \ (3\text{ s.f.}) \end{aligned}$$

**To check your answer:** by GC, cumulative binomial $P(W \leq 1)$ with $n = 8$, $p = \frac{1}{6}$ gives $0.60468\ldots$, matching $\frac{1015625}{1679616} = 0.60468\ldots$ ✓

$\therefore p = \frac{1}{6}$ and $P(W < 2) = \frac{1015625}{1679616} = 0.605$ (3 s.f.).

# Question 9 10 marks

Probability · Conditional Probability

## Part (a) 2 marks

**APPROACH**

From the diagram, the counter makes exactly $5$ moves in total from $S$ to a top endpoint: the first move (from $S$ to $L$ or $R$) is equally likely either way, and each of the remaining $4$ moves independently goes left with probability $p$ or right with probability $q$. Tracking net horizontal displacement (each left move $-1$, each right move $+1$), the six endpoints $A, \ldots, F$ correspond to the six possible totals $-5, -3, -1, 1, 3, 5$, with $B$ (second from the left) corresponding to a total of $-3$. There are exactly two ways to reach a total of $-3$, and we find the probability of each.

**Route 1:** first move left, then exactly $1$ right move among the remaining $4$ (net $-2$ over those $4$, combined with the first move's $-1$, giving $-3$). There are $\binom{4}{1} = 4$ ways to place the right move among the $4$ moves:

$$P(\text{Route 1}) = \frac{1}{2} \times \binom{4}{1} p^3 q = 2p^3 q$$

**Route 2:** first move right, then all $4$ remaining moves left (net $+1 - 4 = -3$):

$$P(\text{Route 2}) = \frac{1}{2} \times p^4 = \frac{1}{2} p^4$$

Since these are the only two ways to reach $B$, and they are mutually exclusive,

$$P(\text{arrives at } B) = 2p^3 q + \frac{1}{2} p^4 \quad \text{(shown)}$$

**To check your answer:** summing $P(\text{arrives at each endpoint})$ by the same method for every one of $A, \ldots, F$ gives $p^4 + 4p^3 q + 6p^2 q^2 + 4pq^3 + q^4 = (p+q)^4 = 1$, since $p + q = 1$, confirming all six probabilities are correctly accounted for. ✓

$\therefore P(\text{arrives at } B) = 2p^3 q + \frac{1}{2} p^4$ (shown).

## Part (b) 4 marks

**APPROACH**

There are $5$ distinct routes from $S$ to $B$: the $4$ routes of "Route 1" type from part (a), each with probability $\frac{1}{2} p^3 q$, and the $1$ route of "Route 2" type, with probability $\frac{1}{2} p^4$. Since Jon's and Kath's counters move independently, we sum, over each of the $5$ routes, the probability that *both* take that specific route, then divide by $P(\text{both at } B)$.

$$P(\text{same route} \mid \text{both at } B) = \frac{P(\text{same route and both at } B)}{P(\text{both at } B)}$$

Since each player's route is chosen independently, the probability both take one specific route is the square of that route's probability:

$$\begin{aligned} &P\left(\text{same route and both at } B\right) \\ &= 4 \times \left(\frac{1}{2}p^3q\right)^2 + \left(\frac{1}{2}p^4\right)^2 \\ &= p^6q^2 + \frac{1}{4}p^8 \end{aligned}$$

Since Jon's and Kath's counters are independent, $P\left(\text{both at } B\right) = P(B)^2$:

$$\begin{aligned} P\left(\text{both at } B\right) &= \left(2p^3q + \frac{1}{2}p^4\right)^2 \\ &= \left[\frac{1}{2}p^3(4q+p)\right]^2 \\ &= \frac{1}{4}p^6(4q+p)^2 \end{aligned}$$

So:

$$\begin{aligned} P\left(\text{same route} \mid \text{both at } B\right) &= \frac{p^6q^2 + \frac{1}{4}p^8}{\frac{1}{4}p^6(4q+p)^2} \\ &= \frac{4q^2+p^2}{(4q+p)^2} \end{aligned}$$

Substituting $q = 1-p$:

$$\begin{aligned} 4q + p &= 4(1-p) + p = 4 - 3p \\ 4q^2 + p^2 &= 4(1-p)^2 + p^2 = 5p^2 - 8p + 4 \end{aligned}$$

**To check your answer:** at $p = q = \frac{1}{2}$, $\frac{4q^2+p^2}{(4q+p)^2} = \frac{1+\frac{1}{4}}{\left(2+\frac{1}{2}\right)^2} = \frac{5/4}{25/4} = \frac{1}{5}$, a valid probability strictly between $0$ and $1$. ✓

$\therefore P\left(\text{same route} \mid \text{both at } B\right) = \frac{5p^2-8p+4}{(4-3p)^2}$.

## Part (c) 4 marks

By the same method as **(a)**, $C$ (third from the left) corresponds to a net displacement of $-1$, reached either by first move left with $2$ right moves among the remaining $4$, or first move right with $1$ right move among the remaining $4$:

$$P\left(\text{arrives at } C\right) = \frac{1}{2}\binom{4}{2}p^2q^2 + \frac{1}{2}\binom{4}{1}p^3q = 3p^2q^2 + 2p^3q$$

Setting $P(B) = P(C)$:

$$\begin{aligned} 2p^3q + \frac{1}{2}p^4 &= 3p^2q^2 + 2p^3q \\ \Rightarrow \frac{1}{2}p^4 &= 3p^2q^2 \end{aligned}$$

Dividing both sides by $p^2$ (since $p \neq 0$),

$$\frac{1}{2}p^2 = 3q^2$$

$$\Rightarrow p^2 = 6q^2$$

Since $p, q > 0$, taking the positive square root,

$$p = \sqrt{6}\,q$$

Substituting into $p + q = 1$:

$$\sqrt{6}\,q + q = 1$$
$$\Rightarrow q\left(\sqrt{6} + 1\right) = 1$$
$$\Rightarrow q = \frac{1}{\sqrt{6}+1} = \frac{\sqrt{6}-1}{5}$$
$$p = \sqrt{6}\,q = \frac{6-\sqrt{6}}{5}$$

**To check your answer:** $p = \frac{6-\sqrt{6}}{5} \approx 0.7101$ and $q = \frac{\sqrt{6}-1}{5} \approx 0.2899$, and $p + q \approx 1$. ✓ Substituting back, $P(B) \approx 0.3347$ and $P(C) \approx 0.3347$, which match. ✓

$\therefore p = \frac{6-\sqrt{6}}{5}$.

## Question 10 14 marks

Correlation and Regression · Regression Lines

### Part (a) 2 marks

There is a fairly strong negative linear correlation between distance driven, $d$, and price, $p$.

Explanation: the product moment correlation coefficient $r = -0.78$ is negative and reasonably close to $-1$, so cars in this sample that have been driven further tend to have a lower price, and this decreasing relationship is reasonably well described by a straight line.

**To check your answer:** on the scatter diagram, the point with the largest price ($p = 22500$) has the smallest distance ($d = 5055$), and prices generally fall as distance increases across the rest of the points, consistent with a negative correlation. ✓

### Part (b) 1 mark

The PMCC would be unchanged; it would remain $r = -0.78$.

Explanation: converting units is a linear scaling of each variable — $d_{\text{km}} = 1.6\,d_{\text{miles}}$ and $p_{\text{SGD}} = 1.7\,p_{\text{GBP}}$, i.e. each variable is multiplied by a positive constant. The PMCC is unaffected by this kind of linear (origin-preserving, positive-scale) transformation of either variable.

### Part (c) 1 mark

This car should not be used in any analysis of the relationship between $d$ and $p$ for the standard model.

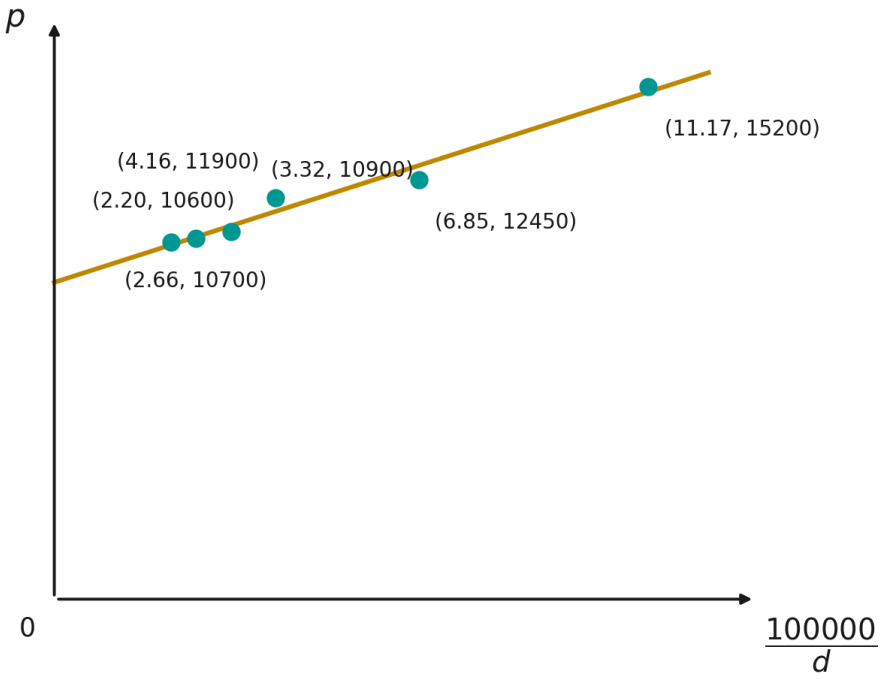
Explanation: being a special version with a bigger engine and extra features, it is not part of the same population as the rest of the data. Its price is likely to be higher for reasons unrelated to distance driven, so including it would bias any analysis of the relationship for the standard model.

## Part (d) 5 marks

We exclude the special car and compute $x = \frac{100\,000}{d}$ for the remaining 6 cars.

| $d$ | 45452 | 37551 | 30100 | 24030 | 14588 | 8954 |
|---|---|---|---|---|---|---|
| $x = \frac{100\,000}{d}$ | 2.20 | 2.66 | 3.32 | 4.16 | 6.85 | 11.17 |
| $p$ | 10600 | 10700 | 10900 | 11900 | 12450 | 15200 |

We plot $p$ against $x$ and draw a line of best fit by eye.




∴ scatter diagram of $p$ against $\frac{100\,000}{d}$, with line of best fit as sketched above.

## Part (e)

The distances used in the least squares regression line of $p$ on $\frac{100\,000}{d}$ are the **vertical** distances from each plotted point to the line — that is, the difference between each observed price $p_i$ and the price predicted by the line at the corresponding $x_i = \frac{100\,000}{d_i}$.

These distances are squared so that points above and below the line (positive and negative deviations) do not cancel out when summed, and so that points further from the line are penalised more heavily than points close to it. The line is chosen to make the sum of these squared vertical distances as small as possible, which is why the procedure is called the **method of least squares**.

∴ vertical distances from each point to the line are squared and summed; the line minimising this sum is the least squares regression line.

### Part (f) 3 marks

By GC, performing linear regression of $p$ on $x = \frac{100\,000}{d}$ for the 6 remaining cars,

$$p = 9398 + 506x, \quad r = 0.987 \ (3 \text{ s.f.})$$

**To check your answer:** with $\bar{x} \approx 5.062$ and $\bar{p} \approx 11958.3$, $S_{xx} = \sum (x_i - \bar{x})^2 \approx 73.06$ and $S_{xp} = \sum (x_i - \bar{x})(p_i - \bar{p}) \approx 36949$, giving gradient $\approx \frac{36949}{73.06} \approx 505.7$ and intercept $\approx 11958.3 - 505.7(5.062) \approx 9398.5$, matching the GC output. ✓

∴ $p = 9398 + 506\left(\frac{100\,000}{d}\right)$ (3 s.f. coefficients), $r = 0.987$ (3 s.f.).

### Part (g) 2 marks

For $d = 5055$:

$$x = \frac{100\,000}{5055} = 19.8 \ (3 \text{ s.f.})$$

Substituting into the regression equation from **(f)**:

$$p = 9398 + 506(19.78) = 19403$$

This gives an estimated price of about £19400.

**Explanation:** this estimate is **not** likely to be reliable. The 6 points used to fit the line have $x$ ranging only from about 2.20 to 11.17, and $x = 19.78$ lies well outside this range, so using the equation here is an extrapolation: the linear relationship may not hold beyond the data. Moreover, $d = 5055$ is exactly the special car excluded in **(c)**, which is not of the standard specification the line was fitted to, so the estimate is doubly unreliable.

∴ estimated price $\approx £19400$, but this is unreliable — it is an extrapolation, and the car is not of the standard specification.

## Question 11 13 marks

Hypothesis Testing · Conducting a hypothesis test

### Part (a) 3 marks

Let $Z \sim N\left(80, 2^2\right)$ be Zhou's time and $T \sim N\left(79, 3^2\right)$ be Tan's time, independent. Zhou wins the race if his time is shorter, i.e. if $Z < T$, equivalently if $D = Z - T < 0$.

Since $Z$ and $T$ are independent normal random variables,

$$D = Z - T \sim N\left(80 - 79,\ 2^2 + 3^2\right) = N(1, 13)$$

Standardising,

$$
\begin{aligned}
&P(D < 0) \\
&= P\left(Y < \frac{0-1}{\sqrt{13}}\right), \quad Y \sim N(0,1) \\
&= \Phi\left(-\frac{1}{\sqrt{13}}\right) \\
&= 0.391 \ \text{(3 s.f., by GC)}
\end{aligned}
$$

**To check your answer:** $-\frac{1}{\sqrt{13}} \approx -0.2774$, and $\Phi(-0.2774) \approx 0.391$, a value just under $0.5$ — this makes sense, since Zhou's mean time ($80$) is only slightly slower than Tan's ($79$), so his probability of winning should be just under one-half. ✓

$\therefore P\,(\text{Zhou wins}) = \Phi\left(-\frac{1}{\sqrt{13}}\right) = 0.391$ (3 s.f.).

## Part (b) 2 marks

$$\overline{x} = 79.21, \quad s^2 = 14.7 \ \text{(3 s.f.)}$$

Explanation: with $n = 30$, $\sum x = 2376.3$, $\sum x^2 = 188653.7$, the unbiased estimate of the population mean is

$$\overline{x} = \frac{\sum x}{n} = \frac{2376.3}{30} = 79.21$$

and the unbiased estimate of the population variance is

$$s^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right) = \frac{1}{29}\left(188653.7 - \frac{2376.3^2}{30}\right) = 14.7 \ \text{(3 s.f.)}$$

**To check your answer:** $2376.3^2 = 5646787.69$, so $\frac{2376.3^2}{30} = 188226.2563$, giving $188653.7 - 188226.2563 = 427.4437$, and $\frac{427.4437}{29} \approx 14.7233$. ✓

## Part (c) 5 marks

Let $\mu$ be the population mean time, in seconds, for Zhou's 100-metre freestyle swims after his exercise regime. We test:

$$H_0 : \mu = 80 \qquad H_1 : \mu < 80$$

at the $5\%$ level of significance (a one-tail test, since we are testing specifically whether his times have **reduced**).

Since the population variance is unknown and $n = 30$ is reasonably large, we use the unbiased estimate $s^2 = 14.7233$ from **(b)** and, by the Central Limit Theorem, the test statistic

$$
\begin{aligned}
Z &= \frac{\overline{x} - \mu_0}{s/\sqrt{n}} = \frac{79.21 - 80}{\sqrt{14.7233/30}} \\
&\approx -1.128
\end{aligned}
$$

**To check your answer:** $\sqrt{14.7233/30} = \sqrt{0.49078} \approx 0.7006$, so $Z \approx \frac{-0.79}{0.7006} \approx -1.128$. ✓

The critical value for a one-tail test at the $5\%$ level is $z = -1.6449$, so we reject $H_0$ if $Z < -1.6449$.

Since $-1.128 > -1.6449$, $Z$ does **not** fall in the critical region, so we do **not** reject $H_0$.

**To check your answer:** the $p$-value is $\Phi(-1.128) \approx 0.130$, which is greater than $0.05$ — confirming, via the equivalent $p$-value approach, that we do not reject $H_0$. ✓

$\therefore$ there is insufficient evidence, at the $5\%$ level of significance, that Zhou's mean swim time has reduced.

### Part (d) 1 mark

A $2$-tail test.

Explanation: Tan is testing whether his mean time has **changed**, which could mean either an increase or a decrease with no direction stated in advance, so the alternative hypothesis is non-directional.

### Part (e) 2 marks

Since Tan's sample size is small ($n = 6$), the Central Limit Theorem cannot be relied on to justify treating the sample mean as approximately normal, so Tan must assume:

**Assumption 1:** Tan's individual swim times are normally distributed, so that the sample mean itself is exactly normal even for this small sample size.

**Assumption 2:** the population variance of Tan's times is unchanged, still $3^2 = 9$ (its previously known value), so that a $Z$-test using this known variance can be carried out, since only the sample mean — not a sample variance — was recorded.

$\therefore$ (1) Tan's times are normally distributed; (2) the population variance is still $3^2 = 9$.

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