ANDERSON SERANGOON JUNIOR COLLEGE

MATHEMATICS **9758**

19 Sept 2022

H2 Math Prelim Paper 2 (100 marks)

3 hours

Additional Material(s): List of Formulae (MF26)

CANDIDATE NAME

CLASS /

READ THESE INSTRUCTIONS FIRST

Write your name and class in the boxes above.
Please write clearly and use capital letters.
Write in dark blue or black pen. HB pencil may be used for graphs and diagrams only.
Do not use staples, paper clips, glue or correction fluid.

Answer **all** the questions and write your answers in this booklet.
Do not tear out any part of this booklet.
Give non-exact numerical answers correct to 3 significant figures, or 1 decimal place in the case of angles in degrees, unless a different level of accuracy is specified in the question.
You are expected to use an approved graphing calculator.
Where unsupported answers from a graphing calculator are not allowed in a question, you are required to present the mathematical steps using mathematical notations and not calculator commands.

All work must be handed in at the end of the examination. If you have used any additional paper, please insert them inside this booklet.
The number of marks is given in brackets [ ] at the end of each question or part question.

Question number	Marks
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
Total	

This document consists of 22 printed pages and 2 blank pages.

[Turn Over

Section A: Pure Mathematics [40 marks]

1 **(i)** The equation $3z^3 - 7z^2 + 17z + m = 0$, where m is a real constant, has a root $z = 1 + 2\mathrm{i}$. Find the value of m.
Hence using an algebraic method, find all the roots of the equation $3z^3 - 7z^2 + 17z + m = 0$. Show your working clearly. [4]

(ii) Hence, solve the equation $\frac{3}{w^3} + \frac{7}{w^2} + \frac{17}{w} - m = 0$, giving your answers in the form $a + b\mathrm{i}$, where $a, b \in \mathbb{R}$. [2]

Commented [KW(W1]: Method
Students should not use the GC to find the roots as the question requires an algebraic method.

Commented [KW(W2]: Method
Students are required to use the answers in (i) to solve (ii).

Solution

(i) Since $z = 1 + 2\mathrm{i}$ is a root,

$3(1+2\mathrm{i})^3 - 7(1+2\mathrm{i})^2 + 17(1+2\mathrm{i}) + m = 0$

$3(-11-2\mathrm{i}) - 7(-3+4\mathrm{i}) + 17 + 34\mathrm{i} + m = 0$

$5 + m = 0$

$m = -5$

$3z^3 - 7z^2 + 17z - 5 = 0$

Since coefficients are all real, $z = 1 - 2\mathrm{i}$ is also a root.

$(z-(1+2\mathrm{i}))(z-(1-2\mathrm{i}))(3z-k) = 3z^3 - 7z^2 + 17z + m$

$(z^2 - 2z + 5)(3z - k) = 3z^3 - 7z^2 + 17z + m$

Commented [KW(W3]: Concepts
Some students did not recognize that the coefficient of z^3 is 3 and wrote z-k instead.

Some others did not know how to express the cubic expression as a product of linear factors.

The highest power is 3 so there should only be 3 roots.

Comparing, $5(-k) = -5$

$k = 1$

$(z-(1+2\mathrm{i}))(z-(1-2\mathrm{i}))(3z-1) = 0$

$\therefore z = 1 + 2\mathrm{i}, 1 - 2\mathrm{i}, \frac{1}{3}$

Alternatively,

Since coefficients are all real, so $z = 1 - 2\mathrm{i}$ is also a root.

$\Rightarrow (z-(1+2\mathrm{i}))(z-(1-2\mathrm{i}))(3z-k) = 3z^3 - 7z^2 + 17z + m = 0$

$(z^2 - 2z + 5)(3z - k) = 3z^3 - 7z^2 + 17z + m$

Comparing coefficients of z

$-2(-k) + 15 = 17$

$k = 1$

$(z^2 - 2z + 5)(3z - 1) = 0$

Comparing, $m = 5(-1) = -5$

$\therefore z = 1 + 2\mathrm{i}, 1 - 2\mathrm{i}, \frac{1}{3}$

(ii) $\frac{3}{w^3} + \frac{7}{w^2} + \frac{17}{w} + 5 = 0$

$-\frac{3}{w^3} - \frac{7}{w^2} - \frac{17}{w} - 5 = 0$

$\frac{3}{(-w)^3} - \frac{7}{(-w)^2} + \frac{17}{(-w)} - 5 = 0$

$3\left(-\frac{1}{w}\right)^3-7\left(-\frac{1}{w}\right)^2+17\left(-\frac{1}{w}\right)-5=0$

Let $z=-\frac{1}{w}$

From (i), $-\frac{1}{w}=1+2i$ or $-\frac{1}{w}=1-2i$ or $-\frac{1}{w}=\frac{1}{3}$

$w=-\frac{1}{1+2i}$ or $w=-\frac{1}{1-2i}$ or $w=-3$

$\therefore w=-\frac{1}{5}(1-2i),\ -\frac{1}{5}(1+2i),\ -3$

Commented [KW(W4]: Concept
Note that the sign for the first and third term should be positive so substitution should be -1/w instead of 1/w.

Commented [KW(W5]: Concept
Some students did not know how to simplify $-\frac{1}{1+2i}$ and $-\frac{1}{1-2i}$.

2 Relative to the origin O, the points A, B and C have position vectors **a**, **b** and **c** respectively. It is given that λ and μ are non-zero numbers such that $\lambda\mathbf{a}+\mu\mathbf{b}-\mathbf{c}=\mathbf{0}$ and $\lambda+\mu=1$.

(i) Show that the points A, B and C are collinear. [3]

The angle between **a** and **b** is known to be obtuse and that $|\mathbf{a}|=2$.

(ii) If k denotes the area of triangle OAB, show that $(\mathbf{a}\cdot\mathbf{b})^2=4\left(|\mathbf{b}|^2-k^2\right)$. [3]

D is a point on the line segment AB with position vector **d**.

(iii) It is given that area of triangle OAB is 6 units^2, $|\mathbf{b}|=10$ and that AOD is 90^0. By finding the value of $\mathbf{a}\cdot\mathbf{b}$, find **d** in terms of **a** and **b**. [4]

Commented [LT6]: Misconception
Some confused it with the concept of coplanar. A, B and C are **collinear** means 3 points are on the **same straight line** while 3 points being **coplanar** means they are on the **same plane.**

It is important to note that Ratio Theorem **is not a method to prove that 3 points are collinear.** It is a result that works on the basis that the 3 points must be on a line before having the position vector of the 3rd point to be expressed in the form taught in the lecture notes.

Commented [LT7]: Question Reading
Some used the angle OAB or ABO when it should be AOB. Note that the requirement for dot product is to have the vectors to be converging or diverging.

Commented [LT8]: Misconception
Many students did not realize the implication of having the angle to be obtuse means that $\mathbf{a}\cdot\mathbf{b}<0$

Commented [LT9]: Question Reading
Some drew the wrong diagram where it was interpreted as ADO is 90^0

Solution

(i)

$\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$

$\overrightarrow{AC} = \mathbf{c} - \mathbf{a}$

$= \lambda\mathbf{a} + \mu\mathbf{b} - \mathbf{a}$

$= (\lambda - 1)\mathbf{a} + \mu\mathbf{b}$

$= -\mu\mathbf{a} + \mu\mathbf{b}$

$= \mu(\mathbf{b} - \mathbf{a})$

Since $\overrightarrow{AC} = \mu\overrightarrow{AB}$ for some $\mu \in \mathbb{R}$, and A is a common point, therefore A, B, C are collinear.

Commented [LT10]: Presentation of Answer
Many did not make mention of the common point between the two vectors used.

(ii) $k = \frac{1}{2}|\mathbf{a} \times \mathbf{b}|$

Commented [LT11]: Misconception
Incorrect to write it as $k = \frac{1}{2}(\mathbf{a} \times \mathbf{b})$ or $k = \frac{1}{2}\mathbf{ab}$ as both **a** and **b** are vectors.

$k = \frac{1}{2}|\mathbf{a}||\mathbf{b}||\sin\theta|$, where θ is the obtuse angle between **a** and **b**

$k^2 = |\mathbf{b}|^2 \sin^2\theta$

$k^2 = |\mathbf{b}|^2 (1 - \cos^2\theta)$

$k^2 = |\mathbf{b}|^2 \left[1 - \left(\frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}\right)^2\right]$

$k^2 = |\mathbf{b}|^2 - \frac{(\mathbf{a} \cdot \mathbf{b})^2}{4}$

$(\mathbf{a} \cdot \mathbf{b})^2 = 4(|\mathbf{b}|^2 - k^2)$

Commented [LT12]: Misconception
Wrote $(\mathbf{a} \cdot \mathbf{b})^2 = |\mathbf{a}|^2 - 2|\mathbf{a}||\mathbf{b}| + |\mathbf{b}|^2$ when it should be $(\mathbf{a} \cdot \mathbf{b})^2 = |\mathbf{a}|^2|\mathbf{b}|^2\cos^2(\angle AOB)$.

(iii) Since D lies on line AB,

$\mathbf{d} = \mathbf{a} + \lambda(\mathbf{b} - \mathbf{a})$ for some $\lambda \in \mathbb{R}$

Commented [LT13]: Question Reading
Many did not use the fact that point D lines on the line AB.

OD is perpendicular to OA

$\Rightarrow [\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a})] \cdot \mathbf{a} = 0$ for some $\lambda \in \mathbb{R}$

$\Rightarrow (1 - \lambda)|\mathbf{a}|^2 + \lambda(\mathbf{b} \cdot \mathbf{a}) = 0$

$4(1 - \lambda) + \lambda(\mathbf{b} \cdot \mathbf{a}) = 0$

As $(\mathbf{a} \cdot \mathbf{b})^2 = 4(|\mathbf{b}|^2 - k^2)$

$(\mathbf{a} \cdot \mathbf{b})^2 = 4(10^2 - 6^2)$

$\mathbf{a} \cdot \mathbf{b} = -16$ ($\because \theta$ is obtuse)

Commented [LT14]: Misconception
Some thought that dot product must always gi[ve] rise to a positive value which is incorrect. The correct definition should be

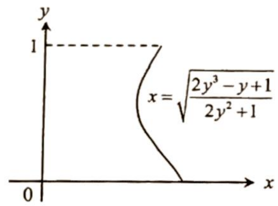
$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta \Leftrightarrow \cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}$

And not $\cos\theta = \left|\frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}\right|$. We only use $\cos\theta = \left|\frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}||\mathbf{b}|}\right|$ if we are told that the angle is acute or when we are trying to find acute ang[le]

$\Rightarrow 4(1 - \lambda) - 16\lambda = 0$

$\lambda = \frac{1}{5}$

$\mathbf{d} = \frac{4}{5}\mathbf{a} + \frac{1}{5}\mathbf{b}$

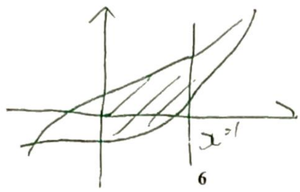
3	**(a)(i)** Use the substitution $u = 1 + x^2$ to find $\int \frac{xe^{1+x^2}}{\sqrt{1+e^{1+x^2}}} \, dx$.	[4]
	(ii) Curves C_1 and C_2 have equations $y = xe^{x^2-2} - \frac{1}{2e}$ and $y = \frac{xe^{1+x^2}}{\sqrt{1+e^{1+x^2}}} - \frac{1}{2e^2}$ respectively. The region bounded by the curves C_1 and C_2, the y-axis and the line $x = 1$ is R. Find the exact area of R.	[3]
	(b) The shape of a vase is formed by rotating the part of the curve $x = \sqrt{\frac{2y^3 - y + 1}{2y^2 + 1}}$ between $y = 0$ and $y = 1$ through 2π radians about the y-axis (see diagram below). Find the exact volume of the vase formed.	[5]
		

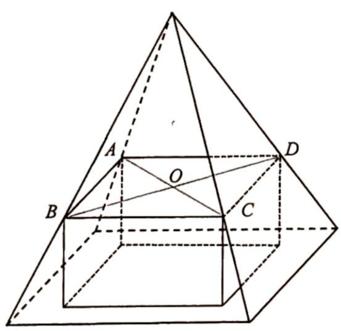
Solution
(ai) $\frac{du}{dx} = 2x$
$\int \frac{xe^{1+x^2}}{\sqrt{1+e^{1+x^2}}} dx = \frac{1}{2} \int \frac{e^{1+x^2}}{\sqrt{1+e^{1+x^2}}} (2x) dx$
$= \frac{1}{2} \int \frac{e^u}{\sqrt{1+e^u}} du$
$= \frac{1}{2} \frac{(1+e^u)^{\frac{1}{2}}}{\frac{1}{2}} + c$
$= (1+e^{1+x^2})^{\frac{1}{2}} + c$

Commented [TCK15]: Presentation
There is no need to replace x by $\pm\sqrt{u-1}$.
Simply replace $2x dx$ by du.
Alternatively,

$$\int \frac{xe^{1+x^2}}{\sqrt{1+e^{1+x^2}}} dx = \int \frac{xe^{1+x^2}}{\sqrt{1+e^{1+x^2}}} \frac{dx}{du} du$$
$$= \int \frac{xe^{1+x^2}}{\sqrt{1+e^{1+x^2}}} \frac{1}{2x} du$$
$$= \int \frac{e^u}{\sqrt{1+e^u}} \frac{1}{2} du$$

Do not forget to include constant of integration.



	(ii) Area $=\int_0^1\left(\frac{xe^{1+x^2}}{\sqrt{1+e^{1+x^2}}}-\frac{1}{2e^2}\right)dx-\int_0^1\left(xe^{x^2-2}-\frac{1}{2e}\right)dx$	
	$=\left[\left(1+e^{1+x^2}\right)^{\frac{1}{2}}\right]_0^1-\left[\frac{e^{x^2-2}}{2}\right]_0^1+\left(\frac{1}{2e}-\frac{1}{2e^2}\right)$	
	$=\sqrt{1+e^2}-\sqrt{1+e}$	
	(b) Volume $=\pi\int_0^1\left(\frac{2y^3-y+1}{2y^2+1}\right)dy$	
	$=\pi\int_0^1 y+\frac{1-2y}{2y^2+1}\,dy$	
	$=\pi\int_0^1\left(y-\frac{2y}{2y^2+1}+\frac{1}{2y^2+1}\right)dy$	
	$=\pi\left[\frac{y^2}{2}-\frac{1}{2}\ln\left(1+2y^2\right)+\frac{1}{\sqrt{2}}\tan^{-1}\left(\sqrt{2}y\right)\right]_0^1$	
	$=\pi\left(\frac{1}{2}-\frac{1}{2}\ln 3+\frac{1}{\sqrt{2}}\tan^{-1}\sqrt{2}\right)$	
4	 The product engineer of a factory crafted the design of a rectangular box, using a right pyramid, that is shown on the diagram above (not drawn to scale). The rectangular box is contained in a right pyramid with a rectangular base such that the upper four corners of the box *A*, *B*, *C* and *D* touch the slant faces of the pyramid, and the bottom four corners lie on the base of the pyramid. *O* is the point of intersection of the two diagonals, *AC* and *BD*. The height of the pyramid is $3\sqrt{2}$ units, the length of the diagonal of its rectangular base is $12\sqrt{2}$ units, the height of the box is b units, where $b<3\sqrt{2}$, and the angle AOB is θ radians. It is given that the box is made of material with negligible thickness.	

Commented [TCK16]: Concept
Area of region is bounded by upper curve C_2 and lower curve C_1 from $x=0$ to $x=1$. Hence the method

$$\int_0^1(C_2-C_1)\,dx=\int_0^1 C_2\,dx-\int_0^1 C_1\,dx$$

Many are not familiar with the technique to integrate xe^{x^2-2}.

Careless mistake
Many took $\frac{1}{2e^2}$ as $\frac{1}{2e}$.

Commented [TCK17]: Misconception
Volume is not $2\pi\int_0^1 x^2dy$ or $\pi\int_0^1 x\,dy$ or $\int_0^1 x^2\,dy$ (many left out π).

Careless mistakes
Many did not reduce $\frac{2y^3-y+1}{2y^2+1}$ to partial fractions correctly.

Many applied the technique in MF26 wrongly as seen below:

$$\int\frac{1}{2y^2+1}dy$$

$$=\int\frac{1}{\left(\sqrt{2}y\right)^2+1}dy$$

$$=\frac{1}{1}\tan^{-1}\left(\frac{\sqrt{2}y}{1}\right)\text{ (wrong)}$$

The correct way is

$$\int\frac{1}{2y^2+1}dy=\frac{1}{2}\int\frac{1}{y^2+\frac{1}{2}}dy$$

$$=\frac{1}{2}\left(\frac{1}{\frac{1}{\sqrt{2}}}\tan^{-1}\frac{y}{\frac{1}{\sqrt{2}}}\right)=\frac{1}{\sqrt{2}}\tan^{-1}\sqrt{2}y$$

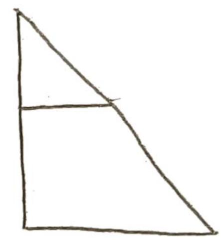
(i) By finding the length of OA in terms of b, show that the volume V of the rectangular box is given by $V = 8b\left(3\sqrt{2}-b\right)^2 \sin\theta$. [3]

For the rest of the question, it is given that $\theta = \frac{\pi}{3}$.

(ii) Find the exact value of b which maximises V. Hence find the cost of manufacturing one such box if the material used to make the box cost \$0.03 per unit^2. [6]

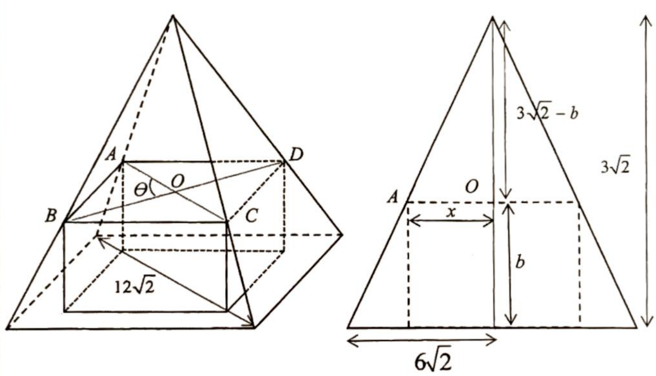
When the height of the box is at half the height of the pyramid, it is reducing at a rate of 2 units per second.

(iii) Determine whether the volume of the box is expanding or shrinking and find the rate at which this is happening. [3]



① similar Δs
② TOA CAH SOH
③ pythagoras thm

Solution



(i) Let $OA = x$ and $V =$ volume of box

By similar triangles, $\dfrac{x}{2(3\sqrt{2})} = \dfrac{3\sqrt{2}-b}{3\sqrt{2}} \Rightarrow x = 2\left(3\sqrt{2}-b\right)$

$V = AB \times BC \times b$

$= \left(2x\sin\frac{\theta}{2}\right)\left(2x\cos\frac{\theta}{2}\right)b$

$= 2b\left[2\left(3\sqrt{2}-b\right)\right]^2 \sin\theta$

$V = 8b\left(3\sqrt{2}-b\right)^2 \sin\theta$ (shown)

Commented [ABK18]: Appproach
For the volume of the box stated, quite a number of students use the following: $V = 4 \times$ Area of $\Delta AOB \times b$. Since this is a **'show' question**, a detailed explaination is needed why they formula is used. Anything based on assumption (even when the result is obtained) is not accepted.

(ii) $V = 4\sqrt{3}b\left(3\sqrt{2}-b\right)^2$

$\dfrac{dV}{db} = 4\sqrt{3}\left[-2b\left(3\sqrt{2}-b\right)+\left(3\sqrt{2}-b\right)^2\right]$

Commented [ABK19]: Strategy
In differentiating V with respect to b, some students expanded the RHS into a cubic equation. This is **not efficient** because when solving dV/db=0, we have to factorise again. Please think about the time constraint of the 3 hour paper.

	$\frac{dV}{db} = 4\sqrt{3}\left(3\sqrt{2}-b\right)\left(3\sqrt{2}-3b\right)$	
	For stationary point, $4\sqrt{3}\left(3\sqrt{2}-b\right)\left(3\sqrt{2}-3b\right)=0$	
	$\Rightarrow b=\sqrt{2}$ or $b=3\sqrt{2}$ (rejected since $b<h$)	
	$\frac{d^2V}{db^2} = 4\sqrt{3}\left[-2b(-1)+\left(3\sqrt{2}-b\right)(-2)+2\left(3\sqrt{2}-b\right)(-1)\right]$ $= 4\sqrt{3}\left[6b-12\sqrt{2}\right]$ $= 24\sqrt{3}\left(b-2\sqrt{2}\right)$	
	$\left.\frac{d^2V}{db^2}\right\|_{b=\sqrt{2}} = -24\sqrt{6} < 0$	
	Thus V is maximised when $b=\sqrt{2}$.	
	$BC = 4\left(3\sqrt{2}-\sqrt{2}\right)\cos\frac{\pi}{6} = 4\sqrt{6}$ $AB = 4\left(3\sqrt{2}-\sqrt{2}\right)\sin\frac{\pi}{6} = 4\sqrt{2}$	
	Cost $= 0.03\times 2\left[4\sqrt{6}\left(4\sqrt{2}\right)+\sqrt{2}\left(4\sqrt{6}\right)+\sqrt{2}\left(4\sqrt{2}\right)\right]$	
	= \$4.64	
	(iii) $\frac{dV}{dt} = \frac{dV}{db}\times\frac{db}{dt}$	
	When $b=\frac{3}{2}\sqrt{2}$,	
	$\left.\frac{dV}{dt}\right\|_{b=\frac{3}{2}\sqrt{2}} = 4\sqrt{3}\left(3\sqrt{2}-\frac{3}{2}\sqrt{2}\right)\left(3\sqrt{2}-\frac{9}{2}\sqrt{2}\right)\times(-2 \text{ units/s})$	
	$= 36\sqrt{3}$ units^3/s	
	Since $\left.\frac{dV}{dt}\right\|_{b=\frac{3}{2}\sqrt{2}} > 0$, the volume of the box is expanding.	
	Section B: Probability and Statistics [60 marks]	
5	Two families, each consisting of an adult couple and three children visited a carnival together.	
	The 10 people went to queue for a ride randomly in one straight line.	
	(i) Find the probability that members of the 2 families stand in alternate positions in that queue.	[2]
	If the ride is made up of two identical circular carriages of five identical seats each.	
	(ii) Find the number of ways the 10 people can be seated if not all the family members are seated together in the same carriage.	[3]
	Solution	

Commented [ABK20]: Misconception
When solving dV/db=0, **there should be 2 values for b** and one of it will be rejected due to the condition given in the question. Some students in the process of solving and factorizing will cancel the factor $\left(3\sqrt{2}-b\right)$. This should not be done. It **should still be considered for part of the solutions** obtained. When needed, it will then be required to be rejected properly.

Commented [ABK21]: Inadequate steps
When proving whether $b=\sqrt{2}$ gives the max/min volume, we can use (1) 2nd derivative test (2) 1st derivative test (sign test). Notice that the **value of the 2nd derivative test need to be quoted** as part of the answer.

Students who did using the 1st derivative test (sign test), a number failed to quote the values. **Values must be quoted** to indicate that the slope is either +ve or -ve.

Commented [ABK22]: Question Reading/ Interpretation
The question asks for the cost of material used to maximise the volume. **Material used is dependent on the SURFACE AREA** of material and NOT the volume. A number of students found the volume and use this to calculate the cost.

(i) Required probability $= \dfrac{2\times 5!\times 5!}{10!}$ or $\dfrac{2\times 5^2\times 4^2\times 3^2\times 2^2\times 1^2}{10!}$

$= \dfrac{1}{126}$

(ii) Number of ways = Total number of ways without restrictions – number of ways where each family sit together

$= \left(\dfrac{{}^{10}C_5\times{}^{5}C_5}{2!}\times(5-1)!\times(5-1)!\right)-(5-1)!\times(5-1)!$

$= 72000$

Method 2:

Case 1: 4 from one family, 1 from other family

${}^{5}C_4\times{}^{5}C_1\times(5-1)!\times{}^{1}C_1\times{}^{4}C_4\times(5-1)!=14400$

Case 2: 3 from one family, 2 from other family

${}^{5}C_3\times{}^{5}C_2\times(5-1)!\times{}^{2}C_2\times{}^{3}C_3\times(5-1)!=57600$

Total = 14400 + 57600 = 72000

> **Commented [KSM23]: Strategy**
> Many resort to slotting – Notice here if slotting is used, you can only slot in consecutive positions ABABABABAB, or BABABABABA. If randomly choose, may end up in situations like _AA_A_A_A

> **Commented [KSM24]:** When randomly choosing members for groups of same size n, you need to divide by n! to to remove multiple counting, as there are n! ways of arranging the n groups of the same compositions

6 In a soccer practice, the coach instructs the players to practise their penalty kicks. A player scores if he successfully kicks a ball into the net of a goal post. The probability that a player scores on the first kick is $\frac{2}{5}$. For all the subsequent kicks, the probability of scoring on that kick will be $\frac{4}{5}$ if the player scores in the preceding kick, and probability of scoring on that kick will be $\frac{1}{6}$ if the player did not score in the preceding kick.

(i) Owen kicked the ball three times consecutively for his practice. Find the probability that he scored on the third kick, given that he scored only twice out of the three kicks. [3]

> **Commented [KSM25]: Question Reading**
> This means that conditional probability should be considered. Plenty ignored that.

(ii) Three players each kicked the ball four times consecutively for their practices. Find the probability that one of the players scored on all four kicks, another player scored on the first kick only, while the remaining player only scored on the second and third kicks. [3]

Solution

(i) P(scored on third kick | scored on only two of the kicks)

$= \dfrac{\text{P(scored on third kick and scored on only two of the kicks)}}{\text{P(scored on only two of the kicks)}}$

$= \dfrac{P(SS'S)+P(S'SS)}{P(SS'S)+P(S'SS)+P(SSS')}$

	$= \dfrac{\left(\frac{2}{5}\times\frac{1}{5}\times\frac{1}{6}\right)+\left(\frac{3}{5}\times\frac{1}{6}\times\frac{4}{5}\right)}{\left(\frac{2}{5}\times\frac{1}{5}\times\frac{1}{6}\right)+\left(\frac{3}{5}\times\frac{1}{6}\times\frac{4}{5}\right)+\left(\frac{2}{5}\times\frac{4}{5}\times\frac{1}{5}\right)}$	
	≈ 0.593 (3 s.f.)	
	(ii) Required probability = P(SSSS) × P(SS'S'S') × P(S'SSS') ×3!	
	$= \left(\frac{2}{5}\times\left(\frac{4}{5}\right)^3\right)\left(\frac{2}{5}\times\frac{1}{5}\times\left(\frac{5}{6}\right)^2\right)\left(\frac{3}{5}\times\frac{1}{6}\times\frac{4}{5}\times\frac{1}{5}\right)\times 3!$	
	$= 0.00109$ (3 s.f.)	
7	Grade A and grade B sugar produced by a company are packed and sold in packets. The mass of both grade A and grade B sugar sold follows independent normal distributions with mean 2.05 kg. The standard deviation for the mass of a randomly chosen packet of grade A and grade B sugar are 0.025 kg and σ kg respectively. If the probability that the mass of a randomly chosen packet of grade B sugar being less than 2 kg is 0.01,	
	(i) show that the value of σ is 0.021493 correct to 5 significant figures.	[2]
	It is given that the profit per kilogram of grade A and B sugar sold is 50 cents and 40 cents respectively.	
	(ii) Find the probability that the total profit of three randomly chosen packets of grade A sugar is higher than three times the profit of a randomly chosen packet of grade B sugar by not more than 65 cents.	[3]
	(iii) Two packets of grade A sugar and n packets of grade B sugar are selected at random. Find the smallest value of n such that the probability that the mean mass of these packets being less than 2.06 kg is at least 0.97.	[3]
	Solution:	
	(i) Let X and Y be the random variable denoting the mass of a packet of grade A and a packet of grade B sugar respectively	
	$Y \sim N(2.05, \sigma^2)$ $P(Y < 2) = 0.01$ $\Rightarrow P(Z < \frac{2-2.05}{\sigma}) = 0.01$ $\Rightarrow \frac{2-2.05}{\sigma} = -2.32635$ $\Rightarrow \sigma = 0.021493$	
	(ii) Let $C = (50)(X_1+X_2+X_3) - 3(40)Y$ $E(C) = 3(50)(2.05) - 3(40)(2.05) = 61.5$ $Var(C) = 3(50)^2(0.025^2) + (3\times 40)^2(0.021493^2) = 11.33957$ Thus $C \sim N(61.5, 11.33957)$	

Commented [KSM26]: Misconception
The 4 kicks are executed consecutively by the 3 players, with 3 different outcomes. Many add up the individual probabilities for each player instead of multiplying, and without considering the random matching of 3 outcomes to the 3 players.

Commented [ABK27]: Inadequate working
The **problem of NOT DEFINING VARIABLES** clearly still persists. Students are to take note that defining variables clearly is **not only a requirement** but also serves **to provide clarity** for themselves in solving such a question.

$$P((50)(X_1 + X_2 + X_3) - 3(40)Y) \le 65)$$
$$= P(C \le 65)$$
$$= 0.85068$$
$$= 0.851 \text{ (3 s.f.)}$$

(iii) Let T be the mean mass of two packets of grade A sugar and n packets of grade B sugar.

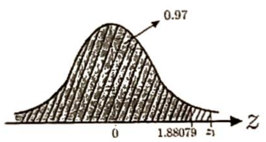
$$T = \frac{X_1 + X_2 + Y_1 + Y_2 + \ldots + Y_n}{n+2}$$

$$E(T) = 2.05 \text{ and } Var(T) = \frac{2(0.025^2) + n(0.021493^2)}{(n+2)^2}$$

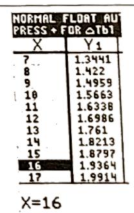
$$P(T < 2.06) \ge 0.97$$

$$\Rightarrow P\left(Z < \frac{2.06 - 2.05}{\sqrt{Var(T)}}\right) \ge 0.97$$

$$\Rightarrow P(Z < z_1) \ge 0.97$$
$$\Rightarrow z_1 > 1.88079$$



$$\Rightarrow \frac{0.01}{\sqrt{\left[\frac{2(0.025^2) + n(0.021493^2)}{(n+2)^2}\right]}} > 1.88079$$



NORMAL FLOAT AU
PRESS + FOR △Tbl

X	Y_1
7	1.3441
8	1.422
9	1.4959
10	1.5663
11	1.6338
12	1.6986
13	1.761
14	1.8213
15	1.8797
16	1.9364
17	1.9914

X=16

Using GC, $n \ge 16$.

Therefore, the smallest possible value of n is 16.

Alternatively,

n	$P(T < 2.06)$
15	0.9699($<$ 0.97)
16	0.9736($>$ 0.97)
17	0.9768($>$ 0.97)

Therefore, the smallest possible value of n is 16.

8 In a public swimming centre, the time spent by a randomly chosen user in using its facilities is T minutes, is known to be normally distributed. The centre manager claims that its users spend an average of 50 minutes to use its facilities. To check this claim, time spent by a random sample of 60 users were recorded. The data recorded has an average of 47 minutes and a standard deviation of 16.4 minutes.

	(i) Find an unbiased estimate of the population variance, giving your answer correct to 2 decimal places.	[1]
	(ii) Test, at the 5% significance level, whether the centre manager overstated the average time spent.	[4]
	(iii) Another sample of size n ($n > 30$) that was collected independently is now used to test, at the 5% significance level, whether the centre manager's claim is valid. For this sample, the mean time taken is 46 minutes. If the result of the test using this information and the unbiased estimate of the population variance in part **(i)** is that the null hypothesis is rejected, find the least possible value that n can take.	[4]
	Solution	
	(i) Unbiased estimate of the population variance $s^2 = \frac{60}{59}(16.4^2) = 273.5186441 \approx 273.52$ (2 decimal places)	
	(ii) Let the random variable T denote the time spent in minutes using the pool facilities and μ denote the population mean time spent in minutes using the pool facilities.	
	To test H_0: $\mu = 50.0$ Against H_1: $\mu < 50.0$ (Centre manager is overstating the claim) Conduct a one-tail test at 5% level of significance, i.e., $\alpha = 0.05$ Under H_0, $\bar{T} \sim N\left(50.0, \frac{273.5186441}{60}\right)$	
	$\bar{t} = 47$ Using GC, p-value = 0.0799976609 $\approx$ 0.0800 (3 sf)	
	Since p-value = 0.0800 > 0.05, we do not reject H_0. There is insufficient evidence at 5% level of significance to conclude that that the centre manager is overstating the average time spent.	
	(iii) Using two-tailed test at 5% significance level, to reject null hypothesis, z_{calc} must lie inside the critical region.	
	To test H_0: $\mu = 50.0$ against H_1: $\mu \neq 50.0$ (Centre manager's claim is valid) Critical Region: $z \le -1.959963986$ or $z \ge 1.959963986$	
	Test Statistics, $Z = \frac{\bar{T} - 50.0}{\sqrt{\frac{273.5186441}{n}}} \sim N(0, 1)$	
	$\therefore z_{calc} = \frac{46.0 - 50.0}{\sqrt{\frac{273.5186441}{n}}} \le -1.959963986$ or $\frac{46.0 - 50.0}{\sqrt{\frac{273.5186441}{n}}} \ge 1.959963986$	

Commented [CKJ28]: Question Reading
Many students did not leave their final answer in 2 decimal places as required in the question.

Commented [CKJ29]: Presentation of Working
Many students did not define the symbols used in the question.

Commented [CKJ30]: Common Mistakes
Many students quoted Central Limit Theorem to be applied in their working. As the distribution of T is normally distribution, this will also imply $\bar{T}$ follow normal distribution.

Commented [CKJ31]: Presentation of Working
Students should define H_0 and H_1 clearly at the start of their working.

	$\frac{-4\sqrt{n}}{\sqrt{273.5186441}} \le -1.959963986$ or $\frac{-4\sqrt{n}}{\sqrt{273.5186441}} \ge 1.959963986$ $4\sqrt{n} \ge 32.41466658$ or $4\sqrt{n} \le -32.41466658$ (rejected) $\sqrt{n} \ge 8.103666645$ $n \ge 65.669$	
	Since n is an integer, the least possible value of n it can take is 66.	
9	**(a)** A random variable X has a binomial distribution with $n = 10$ and probability of success p, where $p < 0.5$. **(i)** Given that $P(X = 3 \text{ or } 4) = 0.2$, write down an equation for the value of p, and find this value numerically.	[2]
	It is given that $p = \frac{1}{5}$.	
	(ii) The mean and standard deviation of X are denoted by μ and σ respectively. Find $P(\mu - \sigma < X < \mu + \sigma)$, correct to 2 decimal places.	[3]
	(b) Mr Chua attempts an online sudoku puzzle each day. The probability that he manages to solve a puzzle on any given day is 0.75, independently of any other day.	
	(i) Find the probability that he solves his third puzzle on the eighth day of his attempt.	[2]
	(ii) Find the probability that, over a period of 8 weeks, Mr Chua manages to solve at least 4 puzzles each week.	[2]
	Solution	
	(a)(i) $X \sim B(10, p)$	
	$P(X = 3 \text{ or } 4) = 0.2$	
	$P(X = 3) + P(X = 4) = 0.2$	
	${}^{10}C_3 p^3(1-p)^7 + {}^{10}C_4 p^4(1-p)^6 = 0.2$	
	$120p^3(1-p)^7 + 210p^4(1-p)^6 = 0.2$	
	Using GC, $p = 0.570$ (rejected $\because p < 0.5$) or $p = 0.163$	
	(a) (ii) $X \sim B(10, \frac{1}{5})$	
	$\mu = E(X) = 10\left(\frac{1}{5}\right) = 2, \ \sigma^2 = 10\left(\frac{1}{5}\right)\left(\frac{4}{5}\right) = \frac{8}{5}$	
	$P(\mu - \sigma < X < \mu + \sigma)$	
	$= P(2 - \sqrt{\frac{8}{5}} < X < 2 + \sqrt{\frac{8}{5}})$	
	$= P(0.73509 < X < 3.2649)$	
	$= P(1 \le X \le 3)$	
	$= P(X \le 3) - P(X = 0)$	
	$= 0.77175$	
	$= 0.77$	
	(b)(i) Let X be the random variable denoting "the number of days in which Mr Chua solves the puzzle out of 7 days"	
	$X \sim B(7, 0.75)$	

Commented [CKJ32]: Interpretation of Question
Many students wrongly interpreted that P(X=3) = 0.2 and P(X=4) = 0.2. The lack of practice on binomial distribution questions in TYS was evident from the students' working.

Commented [CKJ33]: This question was well attempted. Students were familiar in solving this type of question.

Commented [CKJ34]: Common Mistake
Some students attempted to solve the equation algebraically. They failed to realise that they were supposed to GC to solve the equation graphically.

Commented [CKJ35]: Common Mistake
Many students attempted to standardize to find the probability. They failed to realise this is question on Binomial Distribution, not Normal Distribution.

	Required probability $= P(X=2)\times 0.75$	
	$= 0.00865$	
	(ii) Let Y be the random variable denoting "the number of weeks in which Mr Chua solves the puzzle at least 4 times out of 8 weeks".	
	$Y \sim B(8, P(X \geq 4))$	
	$Y \sim B(8, 0.92944)$	
	$P(Y=8) = 0.55690 = 0.557$ Or $(0.92944)^8 = 0.55690 = 0.557$	
10	A bag contains nine numbered discs. Three discs are numbered 3, four discs are numbered 4 and two discs are numbered −1. Two discs are drawn simultaneously. The sum of numbers on them, denoted by X, is recorded. (i) Find the probability distribution for X.	[3]
	(ii) Find E(X) and Var(X).	[2]
	(iii) Two independent observations of X are taken. Find the probability that the difference between these two values is at most 5.	[3]
	(iv) Fifty independent observations of X are taken. Find the approximate probability that the sum of these fifty observations is between 250 and 260.	[3]

(i) Probability Distribution of X:

x	−2	2	3	6	7	8
$P(X=x)$	$\frac{2}{9}\times\frac{1}{8}$ $=\frac{1}{36}$	$2\times\frac{3}{9}\times\frac{2}{8}$ $=\frac{1}{6}$	$2\times\frac{4}{9}\times\frac{2}{8}$ $=\frac{2}{9}$	$\frac{3}{9}\times\frac{2}{8}$ $=\frac{1}{12}$	$2\times\frac{4}{9}\times\frac{3}{8}$ $=\frac{1}{3}$	$\frac{4}{9}\times\frac{3}{8}$ $=\frac{1}{6}$

(ii) $E(X) = \left(-2\times\frac{1}{36}\right)+\left(2\times\frac{1}{6}\right)+\left(3\times\frac{2}{9}\right)+\left(6\times\frac{1}{12}\right)+\left(7\times\frac{1}{3}\right)+\left(8\times\frac{1}{6}\right)$

$= \frac{46}{9}$ or $5.1111 \approx 5.11 (3s.f.)$

$E(X^2) = \left((-2)^2\times\frac{1}{36}\right)+\left(2^2\times\frac{1}{6}\right)+\left(3^2\times\frac{2}{9}\right)+\left(6^2\times\frac{1}{12}\right)+\left(7^2\times\frac{1}{3}\right)+\left(8^2\times\frac{1}{6}\right)$

$= \frac{295}{9}$

$\text{Var}(X) = E(X^2) - (E(X))^2$

$= \frac{295}{9} - \left(\frac{46}{9}\right)^2$

$= \frac{539}{81}$

(iii) $P(|X_1 - X_2| \leq 5) = 1 - P(|X_1 - X_2| \geq 6)$

$= 1 - (2P(-2, 6) + 2P(-2, 7) + 2P(-2, 8) + 2P(2, 8))$

Commented [CKJ36]: Presentation of Working
Many students did not know how to define the random variable for binomial distribution. Inappropriate letters such as Z, N were used in definition of random variables.

Commented [KSX37]: Question Reading

This means there is no replacement of disc. Hence total number of outcomes is NOT 81.

Commented [KSX38]: Many students failed to consider two cases (-1,3) and (3,-1).

Strategy
Total probability should add up to 1.

Commented [KSX39]: Careless Mistakes
Some students knew the formula but did not find this value correctly.

Commented [KSX40]: Strategy

Listing down all the 28 possible cases is not recommended. For those who use this method, few obtain the correct answer.

Using the complementary cases to find the answer is a better strategy.

	$= 1 - 2 \times \frac{1}{36} \times \left(\frac{1}{12} + \frac{1}{3} + \frac{1}{6} \right) - 2 \left(\frac{1}{6} \right)^2$	
	$= \frac{197}{216}$	

	(iv) Since $n = 50$ is large, by Central Limit Theorem,	
	Let $T = X_1 + X_2 + ... + X_{50} \sim N(50 \times \frac{46}{9}, 50 \times \frac{539}{81})$ approximately	
	$T \sim N\left(\frac{2300}{9}, \frac{26950}{81}\right)$ approximately	
	$P(250 < T < 260) \approx 0.216$ (3 s.f.)	
11	Research is being carried out to study the degradation of a herbicide in soil. The concentration (in percentage) of the herbicide in the soil measured over a period of 120 days is recorded. The observations are listed in the table below. It is given that one of the observations has been recorded wrongly.	

Number of days (d)	20	40	60	80	100	120
Concentration (c)	60	57	41	36	33	31

	(i) Draw a scatter diagram to illustrate the data and circle the incorrect observation.	[3]
	For the rest of the question, you should exclude the incorrect observation.	
	(ii) Comment on whether a linear model would be appropriate, referring both to the scatter diagram and the context of the question.	[2]
	It is thought that this set of data can be modelled by one of the following formulae after removing the incorrect observation.	
	Model A: $c^2 = a + bd$	
	Model B: $c = ae^{bd}$	

Commented [KSX41]: Misconception

Students assume X follow normal distribution.

Strategy

Students know that $\overline{X}$ can be approximated to normal distribution but do not know how to proceed. Students could have considered finding $P(5 < \overline{X} < 5.2)$.

By central limit theorem, if n is large, when X follows non-normal distribution, $X_1 + X_2 + ... + X_{50}$ can be approximated to normal distribution as well.

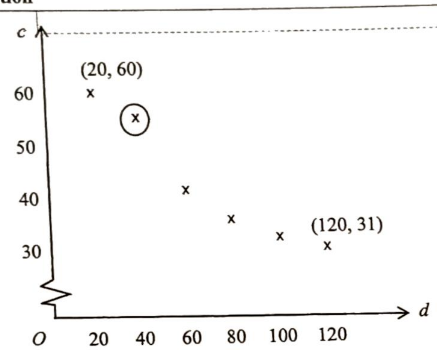
Commented [KSX42]: Conceptual Understanding

There is not need to do this: $P(T < 260) - P(T \le 250)$ because on the GC, the function allows you to key in lower limit and upper limit, since T is a continuous random variable.

(iii)	By calculating the product moment correlation coefficients, explain clearly which of the above models is a more appropriate model for this set of data.	[3]
(iv)	Use the model you identified in (iii) to find the equation of a suitable regression line, and use your equation to estimate the concentration of the herbicide in the soil after 140 days.	[2]
(v)	Comment on the reliability of the estimate obtained in (iv).	[1]
(vi)	Give an interpretation of the vertical intercept of the regression line obtained in (iv) in the context of the question.	[1]

Solution

(i)



Commented [SH43]: Recommendation
1. Appropriate scale and labeling of the axes
2.Correct plot with coordinate of the end points must be labelled.
3.Correct identification of the outlier – by circling as per question requirement

Student who did not draw/ use the space well (¾ of the space provided) to draw a well scaled diagram faced difficulty in identifying the outlier correctly.
For this topic, good graphing skills is important to minimize potential mistakes early.

This must be done well. Easiest way to score marks.

(ii) From the scatter diagram (after removing the outlier), as d increases, c decreases at a decreasing rate.
Also, the concentration of the herbicide will not decrease indefinitely and become a negative percentage.
Hence a linear model should not be used to model this set of data.

Commented [SH44]: Conceptual understanding
Referring to the scatter diagram

1.Describe the trend of the scatter diagram to establish the difference when it is a linear model.
Many students described the scatter diagram as "as d increases, c decreases.". A linear model is also described in the same manner (one with negative gradient). So, to show distinction between the 2. We should describe it as ***decreasing in a decreasing rate.***

Referring to the context of the question
2. Do not predict the trend of the scatter diagram as we do not have the data points out of the data range. So, we should keep answers to establishing if it is a linear model then it should be decreasing indefinitely and will generate negative values for concentration. Which is not possible based on the context of the question.

	(iii) Using GC, $r_A = -0.92958$ while $r_B = -0.97521$.	
	Since the r value for model B is closer to -1 than model A, model B is more appropriate for modelling this set of data.	
	(iv) $c = ae^{bd}$	
	$\ln c = \ln a + bd$	
	From GC, $\ln c = 4.1696 - 0.0066478d$	
	$\ln c = 4.16 - 0.00665d$	
	When $d = 140$, $\ln c = 4.1696 - 0.0066478(140)$	
	$c = 25.5059 \approx 25.5$	
	(v) The estimate is unreliable because the data substituted is outside the data range [20,120] and so the linear relationship between d and $\ln c$ may not hold.	
	(vi) Initially, the concentration of herbicides in the soil is 64.7%.	

Commented [SH45]: Question reading
For calculation of (iii)
1. Outlier/ incorrect observation must be removed before calculating the r value for each of the model.
Many students did not omit this (40,57) from their calculation thus providing incorrect r-values for their models. Show your GC answers up to 5sf then give final answer to 3sf.

Majority of the students did well in choosing the correct model with $|r|$ is closet to 1. Well done!

Commented [SH46]: Things to note
Many students did not know how to linearise ($c = ae^{bd}$) to $\ln c = \ln a + bd$. And calculating the value of c posed a problem for many.

Commented [SH47]: Things to note
1.State the date range clearly for the examiner and add on to say that the trend may not hold and thus the estimate is not reliable.
2.Extrapolation is a process – of using a data point(out of the data range) to calculate an estimate. It doesn't warrant as an answer for marks to be awarded.

Commented [SH48]: Things to note
1.Initial concentration of the herbicide in percentage
2. Finding the $y-$ intercept when d=0.
3. Give your interpretation is required to show understanding. Giving answer alone is not sufficient.