ANDERSON SERANGOON JUNIOR COLLEGE

MATHEMATICS **9758**

H2 Math Prelim Paper 2 (100 marks)

19 Sept 2022

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

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.

$$\left(z-(1+2\mathrm{i})\right)\left(z-(1-2\mathrm{i})\right)(3z-k)=3z^3-7z^2+17z+m$$
$$\left(z^2-2z+5\right)(3z-k)=3z^3-7z^2+17z+m$$

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

$$k=1$$
$$\left(z-(1+2\mathrm{i})\right)\left(z-(1-2\mathrm{i})\right)(3z-1)=0$$
$$\therefore z=1+2\mathrm{i},\ 1-2\mathrm{i},\ \frac{1}{3}$$

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

$$\Rightarrow\left(z-(1+2\mathrm{i})\right)\left(z-(1-2\mathrm{i})\right)(3z-k)=3z^3-7z^2+17z+m=0$$
$$\left(z^2-2z+5\right)(3z-k)=3z^3-7z^2+17z+m$$

Comparing coefficients of z

$$-2(-k)+15=17$$
$$k=1$$
$$\left(z^2-2z+5\right)(3z-1)=0$$

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

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

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

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

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

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

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

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

2 Relative to the origin O, the points A, B and C have position vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ 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 $\mathbf{a}$ and $\mathbf{b}$ 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 $\mathbf{d}$.

(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 $\mathbf{d}$ in terms of $\mathbf{a}$ and $\mathbf{b}$. [4]

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.

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

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

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

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

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

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

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

(iii) Since D lies on line AB,

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

OD is perpendicular to OA

$\Rightarrow \left[\mathbf{a} + \lambda(\mathbf{b} - \mathbf{a})\right]\cdot\mathbf{a} = 0$ 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 $(\mathbf{a}\cdot\mathbf{b})^2 = 4\left(|\mathbf{b}|^2 - k^2\right)$

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

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

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

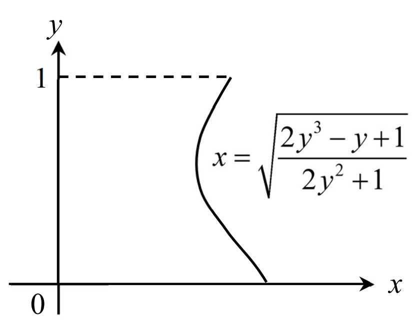
$\lambda = \dfrac{1}{5}$

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

3 **(a)(i)** Use the substitution $u = 1 + x^2$ to find $\displaystyle\int \frac{x\mathrm{e}^{1+x^2}}{\sqrt{1+\mathrm{e}^{1+x^2}}}\,\mathrm{d}x$. [4]

(ii) Curves C_1 and C_2 have equations $y = x\mathrm{e}^{x^2-2} - \dfrac{1}{2\mathrm{e}}$ and $y = \dfrac{x\mathrm{e}^{1+x^2}}{\sqrt{1+\mathrm{e}^{1+x^2}}} - \dfrac{1}{2\mathrm{e}^2}$ 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 $x = \sqrt{\dfrac{2y^3 - y + 1}{2y^2 + 1}}$ 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) $\dfrac{\mathrm{d}u}{\mathrm{d}x} = 2x$

$$\int \frac{xe^{1+x^2}}{\sqrt{1+e^{1+x^2}}}\,dx = \frac{1}{2}\int \frac{e^{1+x^2}}{\sqrt{1+e^{1+x^2}}}(2x)\,dx$$

$$= \frac{1}{2}\int \frac{e^u}{\sqrt{1+e^u}}\,du$$

$$= \frac{1}{2}\frac{(1+e^u)^{\frac{1}{2}}}{\frac{1}{2}} + c$$

$$= \left(1+e^{1+x^2}\right)^{\frac{1}{2}} + c$$

(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 = $\pi \int_0^1 \left(\frac{2y^3 - y + 1}{2y^2+1}\right)dy$

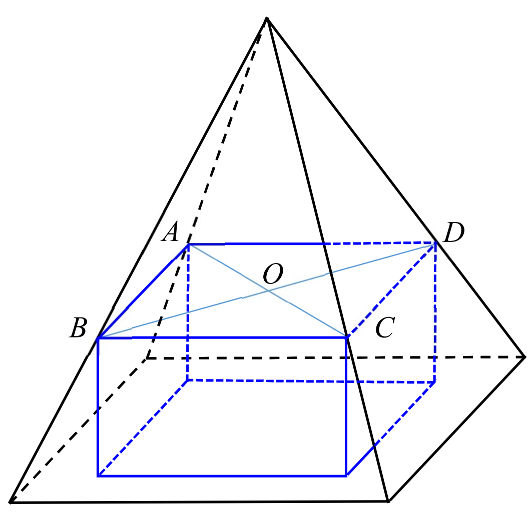
$$= \pi \int_0^1 y + \frac{1-2y}{2y^2+1}\,dy$$

$$= \pi \int_0^1 \left(y - \frac{2y}{2y^2+1} + \frac{1}{2y^2+1}\right)dy$$

$$= \pi \left[\frac{y^2}{2} - \frac{1}{2}\ln\left(1+2y^2\right) + \frac{1}{\sqrt{2}}\tan^{-1}\left(\sqrt{2}y\right)\right]_0^1$$

$$= \pi \left(\frac{1}{2} - \frac{1}{2}\ln 3 + \frac{1}{\sqrt{2}}\tan^{-1}\sqrt{2}\right)$$

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.

(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]

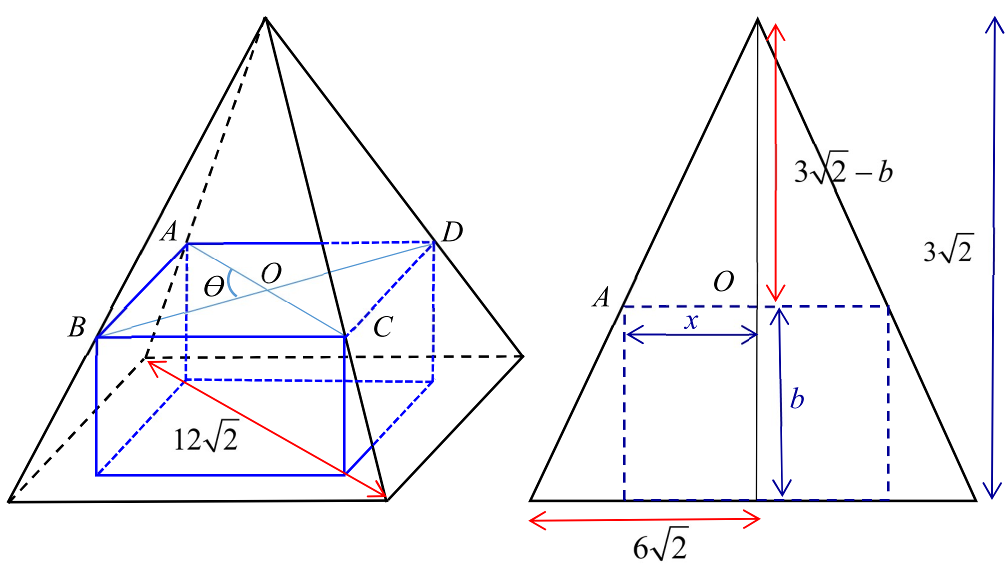
For the rest of the question, it is given that $\theta = \dfrac{\pi}{3}$.

(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]

Solution



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

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

$V = AB \times BC \times b$

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

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

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

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

$$\frac{\mathrm{d}V}{\mathrm{d}b} = 4\sqrt{3}\left[-2b\left(3\sqrt{2}-b\right)+\left(3\sqrt{2}-b\right)^2\right]$$

$$\frac{\mathrm{d}V}{\mathrm{d}b} = 4\sqrt{3}\left(3\sqrt{2}-b\right)\left(3\sqrt{2}-3b\right)$$

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 $b<h$)

$$\frac{\mathrm{d}^2V}{\mathrm{d}b^2} = 4\sqrt{3}\left[-2b(-1)+\left(3\sqrt{2}-b\right)(-2)+2\left(3\sqrt{2}-b\right)(-1)\right]$$

$$= 4\sqrt{3}\left[6b-12\sqrt{2}\right]$$

$$= 24\sqrt{3}\left(b-2\sqrt{2}\right)$$

$$\left.\frac{\mathrm{d}^2V}{\mathrm{d}b^2}\right|_{b=\sqrt{2}} = -24\sqrt{6} < 0$$

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 $= 0.03\times 2\left[4\sqrt{6}\left(4\sqrt{2}\right)+\sqrt{2}\left(4\sqrt{6}\right)+\sqrt{2}\left(4\sqrt{2}\right)\right]$

– to find surface area

$= \$4.64$

(iii) $\frac{\mathrm{d}V}{\mathrm{d}t} = \frac{\mathrm{d}V}{\mathrm{d}b}\times\frac{\mathrm{d}b}{\mathrm{d}t}$

When $b=\frac{3}{2}\sqrt{2}$,

$$\left.\frac{\mathrm{d}V}{\mathrm{d}t}\right|_{b=\frac{3}{2}\sqrt{2}} = 4\sqrt{3}\left(3\sqrt{2}-\frac{3}{2}\sqrt{2}\right)\left(3\sqrt{2}-\frac{9}{2}\sqrt{2}\right)\times\left(-2 \text{ units/s}\right)$$

$$= 36\sqrt{3} \text{ units}^3\text{/s}$$

Since $\left.\frac{\mathrm{d}V}{\mathrm{d}t}\right|_{b=\frac{3}{2}\sqrt{2}} > 0$, 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

(i) Required probability $= \dfrac{2\times 5!\times 5!}{10!}$

$= \dfrac{1}{126}$

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

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

$= 72000$

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 $\dfrac{2}{5}$. For all the subsequent kicks, the probability of scoring on that kick will be $\dfrac{4}{5}$ if the player scores in the preceding kick, and probability of scoring on that kick will be $\dfrac{1}{6}$ 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]

(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)

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

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

$$=\frac{\left(\frac{2}{5}\times\frac{1}{5}\times\frac{1}{6}\right)+\left(\frac{3}{5}\times\frac{1}{6}\times\frac{4}{5}\right)}{\left(\frac{2}{5}\times\frac{1}{5}\times\frac{1}{6}\right)+\left(\frac{3}{5}\times\frac{1}{6}\times\frac{4}{5}\right)+\left(\frac{2}{5}\times\frac{4}{5}\times\frac{1}{5}\right)}$$

≈ 0.593 (3 s.f.)

(ii) Required probability = P(SSSS) $\times$ P(SS'S'S') $\times$ P(S'SSS') $\times 3!$

$$=\left(\frac{2}{5}\times\left(\frac{4}{5}\right)^3\right)\left(\frac{2}{5}\times\frac{1}{5}\times\left(\frac{5}{6}\right)^2\right)\left(\frac{3}{5}\times\frac{1}{6}\times\frac{4}{5}\times\frac{1}{5}\right)\times 3!$$

$= 0.00109$ (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

$Y \sim \mathrm{N}(2.05, \sigma^2)$

$\mathrm{P}(Y < 2) = 0.01$

$\Rightarrow \mathrm{P}(Z < \frac{2-2.05}{\sigma}) = 0.01$

$\Rightarrow \frac{2-2.05}{\sigma} = -2.32635$

$\Rightarrow \sigma = 0.021493$

(ii) $C = (50)(X_1 + X_2 + X_3) - 3(40)Y$

$\mathrm{E}(C) = 3(50)(2.05) - 3(40)(2.05) = 61.5$

$\mathrm{Var}(C) = 3(50)^2(0.025^2) + (3\times 40)^2(0.021493^2) = 11.33957$

Thus $C \sim N(61.5, 11.33957)$

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

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

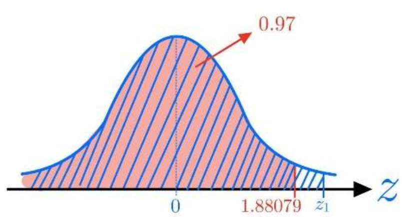
$$T=\frac{X_1+X_2+Y_1+Y_2+\ldots+Y_n}{n+2}$$

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

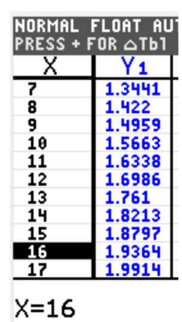
$$\mathrm{P}\left(T<2.06\right)\ge 0.97$$

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

$$\Rightarrow \mathrm{P}\left(Z<z_1\right)\ge 0.97$$
$$\Rightarrow z_1>1.88079 \text{ (from graph)}$$



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


NORMAL FLOAT AU
PRESS + FOR △Tbl

X	Y1
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	$\mathrm{P}\left(T<2.06\right)$
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

$$s^2 = \frac{60}{59}\left(16.4^2\right) = 273.5186441 \approx 273.52 \text{ (2 decimal places)}$$

(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, $\overline{T} \sim \mathrm{N}\left(50.0, \dfrac{273.5186441}{60}\right)$

$\bar{t} = 47$
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: $z \leq -1.959963986$ or $z \geq 1.959963986$

Test Statistics, $Z = \dfrac{\overline{T} - 50.0}{\sqrt{\dfrac{273.5186441}{n}}} \sim \mathrm{N}(0, 1)$

$$\therefore\ z_{\text{calc}} = \frac{46.0 - 50.0}{\sqrt{\dfrac{273.5186441}{n}}} \leq -1.959963986 \quad \text{or} \quad \frac{46.0 - 50.0}{\sqrt{\dfrac{273.5186441}{n}}} \geq 1.959963986$$

$$\frac{-4\sqrt{n}}{\sqrt{273.5186441}} \le -1.959963986 \quad \text{or} \quad \frac{-4\sqrt{n}}{\sqrt{273.5186441}} \ge 1.959963986$$

$$4\sqrt{n} \ge 32.41466658 \quad \text{or} \quad 4\sqrt{n} \le -32.41466658 \text{ (rejected)}$$

$$\sqrt{n} \ge 8.103666645$$

$$n \ge 65.669$$

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 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 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) $X \sim B(10, \frac{1}{5})$

$\mu = E(X) = 10\left(\frac{1}{5}\right) = 2$, $\sigma^2 = 10\left(\frac{1}{5}\right)\left(\frac{4}{5}\right) = \frac{8}{5}$

$P(\mu - \sigma < X < \mu + \sigma)$

$= P(2 - \sqrt{\frac{8}{5}} < X < 2 + \sqrt{\frac{8}{5}})$

$= P(0.73509 < X < 3.2649)$

$= P(X \le 3) - P(X = 0)$

$= 0.77175$

$= 0.77$

(b)(i) Let X be the random variable denoting "the number of days in which Mr Chua solves the puzzle out of 7 days"

$X \sim B(7, 0.75)$

Required probability = $P(X = 2) \times 0.75$

$= 0.0086517$

$= 0.00865$

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

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

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

$= \frac{295}{9}$

$Var(X) = E(X^2) - (E(X))^2$

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

$= \frac{539}{81}$

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

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

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

$= \dfrac{197}{216}$

(iv) Since $n = 50$ is large, by Central Limit Theorem,

Let $T = X_1 + X_2 + \ldots + X_{50} \sim N(50 \times \frac{46}{9}, 50 \times \frac{539}{81})$ approximately

$T \sim \mathrm{N}\left(\frac{2300}{9}, \frac{26950}{81}\right)$ approximately

$\mathrm{P}(250 < T < 260) \approx 0.216$ (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: $c = a\mathrm{e}^{bd}$

(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]

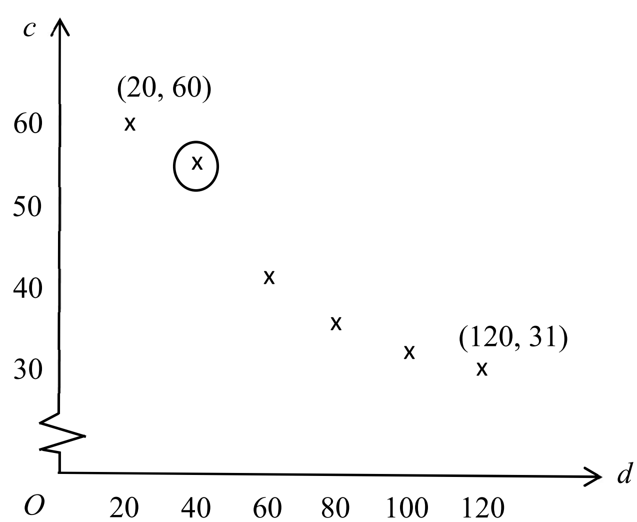
(iv) Use the model you identified in **(iii)** to find the equation of a suitable regression line, and use your equation to estimate the concentration of the herbicide in the soil after 140 days. [2]

(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)



(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.
(iii) Using GC, $r_{\mathrm{A}} = -0.92958$ while $r_{\mathrm{B}} = -0.97521$.
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) $c = a\mathrm{e}^{bd}$

$\ln c = \ln a + bd$

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%.