2020 JC2 H2 PRELIMS Paper 2's Suggested Solutions

<u>1:</u>

Let $y = ax^3 + bx^2 + cx + d$, where a, b, c and d are real constants.

When $x = 0$, $y = 0$, $\therefore\ d = 0$

Given $x = 2$, $y = 0$, $a(2^3) + b(2^2) + c(2) = 0$
$8a + 4b + 2c = 0$ ------ (1)

Given $x = 2.55$, $y = -0.0631$,
$a(2.55^3) + b(2.55^2) + c(2.55) = -0.0631$ ------ (2)

Differentiating y with respect to x, $\frac{\mathrm{d}y}{\mathrm{d}x} = 3ax^2 + 2bx + c$

Given $x = 0.785$, $\frac{\mathrm{d}y}{\mathrm{d}x} = 0$,
$3(0.785^2)a + 2(0.785)b + c = 0$ ------ (3)

Solving (1), (2) and (3) using GC,
$a = 0.0993$, $b = -0.497$, $c = 0.596$ (3 s.f.)

Hence, $y = 0.1x^3 - 0.5x^2 + 0.6x$.

<u>**2:**</u>

$$x=\operatorname{cosec}\theta \Leftrightarrow \frac{\mathrm{d}x}{\mathrm{d}\theta}=-\operatorname{cosec}\theta\cot\theta$$

$$\int_{\frac{2}{\sqrt{3}}}^{2}\frac{1}{x^3\sqrt{x^2-1}}\,\mathrm{d}x=\int_{\frac{\pi}{3}}^{\frac{\pi}{6}}\frac{1}{\operatorname{cosec}^3\theta\sqrt{\operatorname{cosec}^2\theta-1}}\left(-\operatorname{cosec}\theta\cot\theta\right)\mathrm{d}\theta$$

$$=-\int_{\frac{\pi}{3}}^{\frac{\pi}{6}}\frac{1}{\operatorname{cosec}^2\theta\sqrt{\operatorname{cosec}^2\theta-1}}\cot\theta\,\mathrm{d}\theta$$

$$=-\int_{\frac{\pi}{3}}^{\frac{\pi}{6}}\frac{1}{\operatorname{cosec}^2\theta\cot\theta}\cot\theta\,\mathrm{d}\theta$$

$$=-\int_{\frac{\pi}{3}}^{\frac{\pi}{6}}\frac{1}{\operatorname{cosec}^2\theta}\,\mathrm{d}\theta$$

$$=-\int_{\frac{\pi}{3}}^{\frac{\pi}{6}}\sin^2\theta\,\mathrm{d}\theta$$

$$=\int_{\frac{\pi}{6}}^{\frac{\pi}{3}}\sin^2\theta\,\mathrm{d}\theta \qquad \text{(shown)}$$

$$=\int_{\frac{\pi}{3}}^{\frac{\pi}{6}}\frac{\cos 2\theta-1}{2}\,\mathrm{d}\theta$$

$$=\frac{1}{2}\left[\frac{\sin 2\theta}{2}-\theta\right]_{\frac{\pi}{3}}^{\frac{\pi}{6}}$$

$$=\frac{1}{2}\left\{\left[\frac{\sin\frac{2\pi}{6}}{2}-\frac{\pi}{6}\right]-\left[\frac{\sin\frac{2\pi}{3}}{2}-\frac{\pi}{3}\right]\right\}$$

$$=\frac{\pi}{12}$$

3:

(i) $$\frac{3}{(3r+1)(3r+4)} \equiv \frac{A}{3r+1} + \frac{B}{3r+4}$$

When $r = -\frac{1}{3}, A = \frac{3}{-1+4} = 1$

When $r = -\frac{4}{3}, B = \frac{3}{-4+1} = -1$

$$\therefore \frac{3}{(3r+1)(3r+4)} \equiv \frac{1}{3r+1} - \frac{1}{3r+4}$$

$$\begin{aligned}
\sum_{r=1}^{n} \frac{3}{(3r+1)(3r+4)} &= \sum_{r=1}^{n} \left(\frac{1}{3r+1} - \frac{1}{3r+4} \right) \\
&= \left(\frac{1}{4} - \frac{1}{7} \right) + \\
&\quad \left(\frac{1}{7} - \frac{1}{10} \right) + \\
&\quad \vdots \\
&\quad \left(\frac{1}{3n-2} - \frac{1}{3n+1} \right) + \\
&\quad \left(\frac{1}{3n+1} - \frac{1}{3n+4} \right) \\
&= \frac{1}{4} - \frac{1}{3n+4}
\end{aligned}$$

(ii) $$\frac{3}{22}\times\frac{1}{25}+\frac{3}{25}\times\frac{1}{28}+\frac{3}{28}\times\frac{1}{31}+\cdots=\sum_{r=7}^{\infty}\frac{3}{(3r+1)(3r+4)}$$

$$=\sum_{r=1}^{\infty}\frac{3}{(3r+1)(3r+4)}-\sum_{r=1}^{6}\frac{3}{(3r+1)(3r+4)}$$

$$=\lim_{n\to\infty}\left(\frac{1}{4}-\frac{1}{3n+4}\right)-\left(\frac{1}{4}-\frac{1}{3(6)+4}\right)$$

$$=\frac{1}{4}-\left(\frac{1}{4}-\frac{1}{22}\right)$$

$$=\frac{1}{22}$$

4:

(i)

$$\frac{dx}{dt} = kx - 3$$

Since the population remains constant when $x = p$,

$0 = kp - 3$

$k = \frac{3}{p}$

Thus $\frac{dx}{dt} = \frac{3}{p}x - 3$

(ii)

$$\frac{dx}{dt} = \frac{3}{p}x - 3$$

$$\frac{dx}{dt} = \frac{3x - 3p}{p} = \frac{3}{p}(x - p)$$

$$\int \frac{1}{x - p}\, dx = \int \frac{3}{p}\, dt$$

$$\ln|x - p| = \frac{3}{p}t + c$$

$$|x - p| = e^{\frac{3}{p}t + c}$$

$$x - p = Ae^{\frac{3}{p}t} \quad \text{where } A = \pm e^{c}$$

$$x = Ae^{\frac{3}{p}t} + p$$

$$5 = Ae^{\frac{3}{p}(0)} + p \Rightarrow A = 5 - p$$

Thus $x = (5 - p)e^{\frac{3}{p}t} + p$ (shown)

(iii)

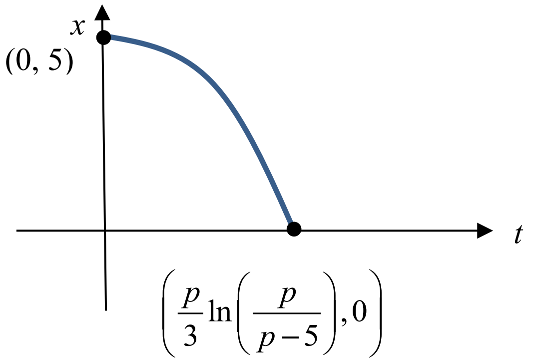
Since the population of trees decreases over time, $5 - p < 0 \Rightarrow p > 5$

Let $x = 0$, $0 = (5 - p)e^{\frac{3}{p}t} + p$

$$-\frac{p}{5 - p} = e^{\frac{3}{p}t}$$

$\frac{p}{3}\ln\left|-\frac{p}{5 - p}\right| = t$ or $t = \frac{p}{3}\ln\left|\frac{p}{p - 5}\right|$ or $t = \frac{p}{3}\ln\left(\frac{p}{p - 5}\right)$ since $p > 5$

Thus,



Starting from the initial year $t = 0$, it takes $\frac{p}{3}\ln\left(\frac{p}{p-5}\right)$ years for the tree population to be fully depleted.

5:

(ai)

$\overrightarrow{AD} = \overrightarrow{AO} + \overrightarrow{OD} = -\underset{\sim}{a} + \frac{1}{2}\underset{\sim}{b}$

$\overrightarrow{BE} = \overrightarrow{BO} + \overrightarrow{OE} = -\underset{\sim}{b} + \frac{1}{2}\underset{\sim}{a}$

Method 1 (Using Triangular Law of Vectors)

$\overrightarrow{OF} = \overrightarrow{OA} + \overrightarrow{AF} = \underset{\sim}{a} + \lambda(-\underset{\sim}{a} + \frac{1}{2}\underset{\sim}{b})$ for some $\lambda \in \mathbb{R}$

Also,

$\overrightarrow{OF} = \overrightarrow{OB} + \overrightarrow{BF} = \underset{\sim}{b} + \mu(-\underset{\sim}{b} + \frac{1}{2}\underset{\sim}{a})$ for some $\mu \in \mathbb{R}$

Equating we get,

$(1-\lambda)\underset{\sim}{a} + \frac{1}{2}\lambda\underset{\sim}{b} = \frac{1}{2}\mu\underset{\sim}{a} + (1-\mu)\underset{\sim}{b}$

Since $\underset{\sim}{a}$ and $\underset{\sim}{b}$ are non-parallel and non-zero, we can equate coefficients.

$$\therefore \begin{cases} 1-\lambda = \frac{\mu}{2} \\ \frac{\lambda}{2} = 1-\mu \end{cases}$$

Solving, $\lambda = \mu = \frac{2}{3}$

$\therefore \overrightarrow{OF} = \underset{\sim}{a} + \frac{2}{3}(-\underset{\sim}{a} + \frac{1}{2}\underset{\sim}{b}) = \frac{1}{3}\underset{\sim}{a} + \frac{1}{3}\underset{\sim}{b}$ (shown)

Method 2 (Using vector equation of lines)

$l_{AD} : \underset{\sim}{r} = \underset{\sim}{a} + \lambda\left(\frac{1}{2}\underset{\sim}{b} - \underset{\sim}{a}\right), \lambda \in \mathbb{R}$

$l_{BE} : \underset{\sim}{r} = \underset{\sim}{b} + \mu\left(\frac{1}{2}\underset{\sim}{a} - \underset{\sim}{b}\right), \mu \in \mathbb{R}$

At the point of intersection of the two lines,

$$\underset{\sim}{a} + \lambda\left(\frac{1}{2}\underset{\sim}{b} - \underset{\sim}{a}\right) = \underset{\sim}{b} + \mu\left(\frac{1}{2}\underset{\sim}{a} - \underset{\sim}{b}\right)$$

$$(1-\lambda)\underset{\sim}{a} + \frac{1}{2}\lambda\underset{\sim}{b} = (1-\mu)\underset{\sim}{b} + \frac{1}{2}\mu\underset{\sim}{a}$$

Since $\underset{\sim}{a}$ and $\underset{\sim}{b}$ are non-zero and not equal, we can equate coefficients,

$$\therefore \begin{cases} 1-\lambda = \frac{\mu}{2} \\ \frac{\lambda}{2} = 1-\mu \end{cases}$$

Solving, $\lambda = \mu = \frac{2}{3}$

$\therefore \overrightarrow{OF} = \underset{\sim}{a} + \frac{2}{3}(-\underset{\sim}{a} + \frac{1}{2}\underset{\sim}{b}) = \frac{1}{3}\underset{\sim}{a} + \frac{1}{3}\underset{\sim}{b}$ (shown)

Method 3 (Using property of medians)

By property of medians,

$$\overrightarrow{BF} = \frac{2}{3}\overrightarrow{BE} = \frac{2}{3}\left(-\underset{\sim}{b} + \frac{1}{2}\underset{\sim}{a}\right) = \frac{1}{3}\underset{\sim}{a} - \frac{2}{3}\underset{\sim}{b}$$

So, $\overrightarrow{OF}=\overrightarrow{OB}+\overrightarrow{BF}=\underset{\sim}{b}+\frac{1}{3}\underset{\sim}{a}-\frac{2}{3}\underset{\sim}{b}=\frac{1}{3}\left(\underset{\sim}{a}+\underset{\sim}{b}\right)$ (Shown)

<u>Method 3 (Using Ratio Theorem and property of medians)</u>
By property of medians, $DF:FA=1:2$
By ratio theorem,
$\overrightarrow{OF}=\frac{\overrightarrow{OA}+2\overrightarrow{OD}}{3}=\frac{\underset{\sim}{a}+2\left(\frac{1}{2}\underset{\sim}{b}\right)}{3}=\frac{1}{3}\left(\underset{\sim}{a}+\underset{\sim}{b}\right)$ (Shown)

(aii)
$\overrightarrow{OC}=\frac{1}{2}(\underset{\sim}{a}+\underset{\sim}{b})$ using the midpoint theorem.
$\overrightarrow{OF}=\frac{1}{3}(\underset{\sim}{a}+\underset{\sim}{b})$
$\quad=\frac{2}{3}\left[\frac{1}{2}(\underset{\sim}{a}+\underset{\sim}{b})\right]=\frac{2}{3}\overrightarrow{OC}$

Since $\overrightarrow{OF}=k\overrightarrow{OC}$ for some scalar k where $0<k<1$, point F lies on OC.

(bi)
$\overrightarrow{PQ}=\underset{\sim}{q}-\underset{\sim}{p}$.
Since $\overrightarrow{PQ}\cdot\underset{\sim}{n}=(\underset{\sim}{q}-\underset{\sim}{p})\cdot\underset{\sim}{n}$
$\quad=\underset{\sim}{q}\cdot\underset{\sim}{n}-\underset{\sim}{p}\cdot\underset{\sim}{n}$
$\quad=\underset{\sim}{p}\cdot\underset{\sim}{n}-\underset{\sim}{p}\cdot\underset{\sim}{n}$ (since $\underset{\sim}{q}.\underset{\sim}{n}=\underset{\sim}{p}.\underset{\sim}{n}$)
$\quad=0$
$\therefore\ \overrightarrow{PQ}\perp\underset{\sim}{n}$ (shown).

So, line PQ // plane Π

Since P is a point on line PQ, but P is not on plane Π ($\because\ \overrightarrow{OP}\cdot\underset{\sim}{n}=\underset{\sim}{p}\cdot\underset{\sim}{n}\neq 0$),
line PQ is parallel to Π but non-intersecting.

(b)(ii)
Point P with position vector $\underset{\sim}{p}$ lies on the plane.
$\overrightarrow{PQ}=\underset{\sim}{q}-\underset{\sim}{p}$ is parallel to the plane.
Since $\underset{\sim}{n}$ is perpendicular to plane Π, and plane Π is perpendicular to the new plane, another vector parallel to this plane is $\underset{\sim}{n}$.

$\therefore$ New plane : $\underset{\sim}{r}=\underset{\sim}{p}+\lambda(\underset{\sim}{q}-\underset{\sim}{p})+\mu\underset{\sim}{n},\ \lambda,\mu\in\mathbb{R}$.
Equivalent answers:
New plane : $\underset{\sim}{r}=\underset{\sim}{q}+\lambda(\underset{\sim}{q}-\underset{\sim}{p})+\mu\underset{\sim}{n},\ \lambda,\mu\in\mathbb{R}$ or $\underset{\sim}{r}=\underset{\sim}{q}+\lambda(\underset{\sim}{p}-\underset{\sim}{q})+\mu\underset{\sim}{n},\ \lambda,\mu\in\mathbb{R}$
New plane : $\underset{\sim}{r}\bullet\left[(\underset{\sim}{q}-\underset{\sim}{p})\times\underset{\sim}{n}\right]=\underset{\sim}{p}\bullet\left[(\underset{\sim}{q}-\underset{\sim}{p})\times\underset{\sim}{n}\right]$ or $\underset{\sim}{r}\bullet\left[(\underset{\sim}{p}-\underset{\sim}{q})\times\underset{\sim}{n}\right]=\underset{\sim}{p}\bullet\left[(\underset{\sim}{p}-\underset{\sim}{q})\times\underset{\sim}{n}\right]$
New plane : $\underset{\sim}{r}\bullet\left[(\underset{\sim}{q}-\underset{\sim}{p})\times\underset{\sim}{n}\right]=\underset{\sim}{q}\bullet\left[(\underset{\sim}{q}-\underset{\sim}{p})\times\underset{\sim}{n}\right]$ or $\underset{\sim}{r}\bullet\left[(\underset{\sim}{p}-\underset{\sim}{q})\times\underset{\sim}{n}\right]=\underset{\sim}{q}\bullet\left[(\underset{\sim}{p}-\underset{\sim}{q})\times\underset{\sim}{n}\right]$

<u>6:</u>

(ai) P (all counters blue) $= \frac{{}^{6}C_{4}}{{}^{10}C_{4}} = \frac{1}{14}$

(aii) <u>Case 1</u> : P(3 blue,1 red) $= \frac{{}^{6}C_{3} \times {}^{4}C_{1}}{{}^{10}C_{4}} = \frac{8}{21}$

<u>Case 2</u> : P (all blue) $= \frac{1}{14}$

P(at least 3 blue)
= P(3 blue, 1 red) + P(all blue)
$= \frac{8}{21} + \frac{1}{14} = \frac{19}{42}$

(aiii) P (all counters red) $= \frac{{}^{4}C_{4}}{{}^{10}C_{4}} = \frac{1}{210}$

P(at least 1 counter of each colour is drawn)
= 1 – P(all blue) – P(all red)
$= 1 - \frac{1}{14} - \frac{1}{210} = \frac{97}{105}$

(aiv) P(at least 3 blue | at least 1 of each colour)

$$= \frac{\text{P(at least 3 blue } \cap \text{ at least 1 of each colour)}}{\text{P(at least 1 of each colour)}}$$

$$= \frac{\text{P(3 blue,1 red)}}{\text{P(at least 1 of each colour)}}$$

$$= \frac{\frac{8}{21}}{\frac{97}{105}} = \frac{40}{97}$$

(b) Since P(at least 3 blue | at least 1 of each colour) = $\frac{40}{97} \neq \frac{19}{42}$ = P (at least 3 blue),

the two events are **not** independent.

<u>7:</u>

(i)
Any two of the following:

- the event whether a JC2 student sign up for the GP tuition is independent of one another.
- The probability of a JC2 student sign up for the GP tuition is a fixed constant 0.08.
- There are a fixed number of JC2 students received the flyers.

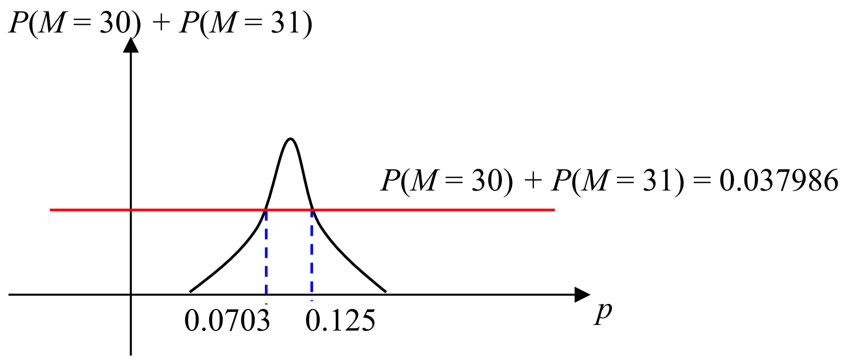
(ii) Let G be the r.v. denoting the number of students sign up for GP tuition out of 120 students.
$G \sim B(120, 0.08)$

(a) $P(G > 15) = 1 - P(G \leq 15)$
$= 0.0301$

(b) $Var(G) = np(1-p)$
$= 120(0.08)(0.98)$
$= 8.83 \text{ or } \dfrac{1104}{125}$

(iii) $P(M = 30 \text{ or } 31) = 0.037986$
$P(M = 30) + P(M = 31) = 0.037986$

From GC,



Since $p > 0.08$, $p = 0.125$

8:

(i)

y	0	1	2
$\mathrm{P}(Y=y)$	$\left(\dfrac{2}{n+2}\right)\left(\dfrac{1}{n+1}\right)$ $=\dfrac{2}{(n+2)(n+1)}$	$2\left(\dfrac{2}{n+2}\right)\left(\dfrac{n}{n+1}\right)$ $=\dfrac{4n}{(n+2)(n+1)}$	$\left(\dfrac{n}{n+2}\right)\left(\dfrac{n-1}{n+1}\right)$ $=\dfrac{n(n-1)}{(n+2)(n+1)}$

(ii)
$$\begin{aligned}\mathrm{E}(Y)&=1\left(\frac{4n}{(n+2)(n+1)}\right)+2\left(\frac{n(n-1)}{(n+2)(n+1)}\right)\\&=\frac{4n+2n^2-2n}{(n+2)(n+1)}\\&=\frac{2n(n+1)}{(n+2)(n+1)}\\&=\frac{2n}{n+2}\end{aligned}$$
$\therefore A=2$

(iii)
$$\begin{aligned}\mathrm{Var}(Y)&=1^2\left(\frac{4n}{(n+2)(n+1)}\right)+2^2\left(\frac{n(n-1)}{(n+2)(n+1)}\right)-\left(\frac{2n}{n+2}\right)^2\\&=\frac{4n}{(n+2)(n+1)}+\frac{4n^2-4n}{(n+2)(n+1)}-\left(\frac{2n}{n+2}\right)^2\\&=\frac{4n^2}{(n+2)(n+1)}-\frac{4n^2}{(n+2)^2}\\&=\frac{4n^2(n+2)-4n^2(n+1)}{(n+2)^2(n+1)}\\&=\frac{4n^2(n+2-n-1)}{(n+2)^2(n+1)}\\&=\frac{4n^2}{(n+2)^2(n+1)}\end{aligned}$$

9:

(a) **(i)** No of arrangements = $1\times1\times\frac{7!}{2!2!}=1260$

(ii) No of arrangements = $4!\times\frac{4!}{2!2!}\times{}^{5}C_{2}=1440$

(b) Case 1: 2 Green, 1 Blue, 1 Red

No of arrangements = $1\times{}^{3}C_{1}\times{}^{2}C_{1}\times\frac{4!}{2!}=72$

Case 2: 1 Green, 2 Blue, 1 Red

No of arrangements = $1\times{}^{3}C_{2}\times{}^{2}C_{1}\times4!=144$

Case 3(a): 1 Green, 1 Blue, 2 Red (1 small and 1 medium)

No of arrangements = $1\times{}^{3}C_{1}\times1\times1\times4!=72$

Case 3(b): 1 Green, 1 Blue, 2 Red (2 small)

No of arrangements = $1\times{}^{3}C_{1}\times1\times\frac{4!}{2!}=36$

Case 3(c): 1 Green, 1 Blue, 2 Red (2 medium)

No of arrangements = $1\times{}^{3}C_{1}\times1\times\frac{4!}{2!}=36$

Total no of arrangements = $72+144+72+36+36=360$

10:

(i) The distribution might become bimodal when the data for both volumes of Grade X and Grade Y petrol sold in an hour are combined.

(ii) Given $X \sim N(\mu , \sigma^2)$,

$\mathrm{P}(X \le 170) = 0.023$

$\mathrm{P}\left(Z \le \frac{170-\mu}{\sigma}\right) = 0.023$

$\frac{170-\mu}{\sigma} = -1.9954$ (5s.f.)

$\mu - 1.9954\sigma = 170 --- (1)$

$\mathrm{P}(X > 180) = 0.16$

$\mathrm{P}\left(\mathrm{Z} > \frac{180-\mu}{\sigma}\right) = 0.16$

$\frac{180-\mu}{\sigma} = 0.99446$ (5s.f.)

$\mu + 0.99446\sigma = 180 --- (2)$

Solving (1) and (2) using GC,

$\mu = 176.67$ (5 s.f.) and $\sigma = 3.3446$ (5 s.f.)

$\therefore$ $\mu = 177$ (3 s.f.) and $\sigma = 3.34$ (3 s.f.)

(iii) Let Y be the random variable denoting the volume of Grade Y petrol sold in an hour.

$X \sim N(176.67 , 3.3446^2)$ and $Y \sim N(200 , 5^2)$

Let $W = X_1 + X_2 + X_3 - 3Y$,

$W \sim N(3(176.67) - 3(200) , 3(3.3446^2) + 3^2(5^2))$

$\sim N(-69.99 , 258.56)$

$\mathrm{P}(|W| < 72) = 0.5497389693$

$= 0.54974$(5 S.F.)

(iv) The volumes of Grade X and Grade Y petrol sold in an hour is independent of each other.

(v) Let J be the random variable denoting the volume of Grade Y petrol sold in an hour during the month of June.

$J = 1.1Y \sim \mathrm{N}(1.1(200)\,,\,1.1^2(5^2))$

$\sim \mathrm{N}(220\,,\,30.25)$

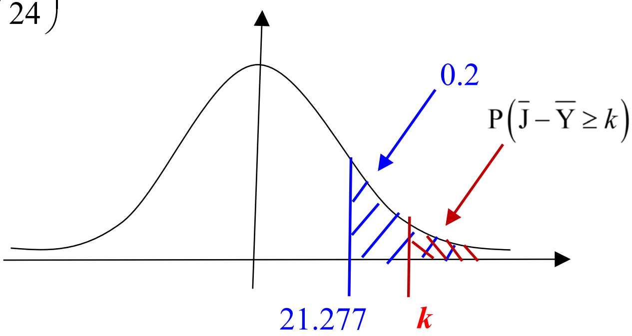
$\overline{J} \sim \mathrm{N}\left(220, \dfrac{30.25}{24}\right)$ and $\overline{Y} \sim \mathrm{N}\left(200, \dfrac{5^2}{24}\right)$

$\overline{J} - \overline{Y} \sim \mathrm{N}(20, 2.3021)$

Given $\mathrm{P}\left(\overline{J} - \overline{Y} \geq k\right) \leq 0.2$,

$k \geq 21.277$ (5 s.f.)

$\therefore$ minimum value of $k = 22$



11:

(a) Given sample size $n = 59$, and $\sum x = 344.5$,

$\bar{x}$, unbiased estimate of population mean incubation period / days

$$= \frac{\sum x}{n} = \frac{344.5}{59} = 5.8389... = 5.84 \text{ (3 s.f.)}$$

s^2, unbiased estimate of population variance incubation period / days^2

$$= \frac{n}{n-1}(\text{Sample variance}) = \frac{59}{58}(8.46136)$$
$$= 8.6072... \quad = 8.61 \text{ (3 s.f.)}$$

(bi) A 2-tail test should be used, as the official intends to test whether the population mean incubation period for SARS-CoV-2 differs from the specific duration of 6.5 days, which could be a change in either direction.

Let μ denote the actual population mean incubation period of the SARS-CoV-2 virus, in days.

Null hypothesis, H_0 : $\mu = 6.5$

Alternative hypothesis, H_1 : $\mu \neq 6.5$

(bii) By the Central Limit Theorem, the sample mean of a sufficiently large number (>20) of 59 independent observations follows an approximate normal distribution, and the hypothesis test could still be carried out without knowledge of the incubation period's population distribution.

(biii) Hypothesis testing : p-value method (using GC)

Let X be a random variable denoting the virus' incubation period of a randomly selected COVID-19 patient, in days.

Under the null hypothesis H_0 ($\mu = 6.5$)

the sample mean virus incubation period of a random sample of 59 COVID-19 patients

$$\bar{X} \sim \mathrm{N}\left(6.5, \frac{\sigma^2}{59}\right) \text{ approximately by CLT, where } \sigma^2 = \mathrm{Var}(X)$$

Against the alternative hypothesis H_1, with σ^2 approximated by its unbiased estimate $s^2 = 8.6072...$ from the sample taken, p-value $= 0.083516... = 0.0835$ (3 s.f.)

Since p-value > 0.05, there is insufficient/no evidence at the 5% level of significance to conclude that the population mean incubation period of SARS-CoV-2 is different from 6.5 days.

Alternative solution (manual calculation of p-value)

Let X be a random variable denoting the virus' incubation period of a randomly selected COVID-19 patient, in days.

Under the null hypothesis H_0 ($\mu = 6.5$)
the sample mean virus incubation period of a random sample of 59 COVID-19 patients

$$\bar{X} \sim \mathrm{N}\left(6.5,\ \frac{\sigma^2}{59}\right) \text{ approximately by CLT, where } \sigma^2 = \mathrm{Var}(X)$$

Estimating σ^2 with its unbiased estimator S^2,
Test Statistic $Z = \dfrac{\bar{X} - 6.5}{\sqrt{\dfrac{S^2}{59}}} \sim \mathrm{N}(0,\ 1)$ approximately.

From the taken sample of 59 COVID-19 patients, an observed value of the test statistic

$$z_{\text{test}} = \frac{\bar{x} - 6.5}{\sqrt{\dfrac{s^2}{59}}} = \frac{5.8389... - 6.5}{\sqrt{\dfrac{8.6072...}{59}}}$$

$= -1.7306...$, where $\bar{x}$ denotes the sample mean incubation period from the taken sample of 59 patients.

Against the alternative hypothesis H_1,

$$\begin{aligned} p\text{-value} &= \mathrm{P}\left(Z < -1.7306... \text{ or } Z > 1.7306...\right) \\ &= 2\ \mathrm{P}(Z < -1.7306...) \\ &= 2(0.041758...) \quad = 0.083516... \end{aligned}$$

Level of significance = 5% = 0.05
Since p-value = 0.083516… > 0.05,
there is insufficient/no evidence at this level of significance to conclude that the population mean incubation period of SARS-CoV-2 is different from 6.5 days.

(biv) Simplified description (modelled after phrasing for p-value from lecture notes) :

The p-value is the probability of obtaining a random sample of 59 patients having a sample mean incubation period that is **as extreme or more extreme than** the collected sample mean of 5.8389… days from 6.5 days, assuming that the population mean incubation period is 6.5 days.

Detailed description :

The p-value represents the conditional two-tailed probability of independently obtaining a random sample of 59 patients producing a sample mean incubation period that is as far off from 6.5 days as the collected sample mean of 5.8389 days is, or further, under the null hypothesis that the population mean incubation period is 6.5 days.

(bv) Assume sample size n large enough for CLT to be used.

Given sample mean $\bar{x} = 5.8389...$, $\sigma_x^{\ 2} = 8.46136$.

Then, $s^2 = \dfrac{n}{n-1}\sigma_x^{\ 2} = \dfrac{n}{n-1}(8.46136)$.

In this test, $z_{\text{test}} = \dfrac{\bar{x}-6.5}{\sqrt{\dfrac{s^2}{n}}} = \dfrac{5.8389...-6.5}{\sqrt{\dfrac{\frac{n}{n-1}(8.46136)}{n}}}$

$= \dfrac{5.8389...-6.5}{\sqrt{8.46136}}\sqrt{n-1}$

The level of significance of 0.02 in this test,
is a **2**-tailed probability, which defines a critical value,

$$z_{\text{crit}} = 2.326347... \quad \text{invNorm}\left(\frac{0.02}{\mathbf{2}},\ 0,\ 1,\ \text{RIGHT}\right)$$

that determines a rejection region on both ends.

Since H_0 is rejected,

$$z_{\text{test}} < -z_{\text{crit}} \quad \text{or} \quad z_{\text{test}} > z_{\text{crit}}$$

$$\frac{5.8389...-6.5}{\sqrt{8.46136}}\sqrt{n-1} < -2.326347... \quad \text{or} \quad \frac{5.8389...-6.5}{\sqrt{8.46136}}\sqrt{n-1} > 2.326347... \quad \text{(rej. } \because \text{ LHS} < 0)$$

$$\sqrt{n-1} > (-2.326347...)\frac{\sqrt{8.46136}}{5.8389...-6.5}$$

$$= 10.23722...$$

$$n-1 > (10.23722...)^2 = 104.800...$$

$$n > 105.800...$$

Least $n = 106$

The least sample size for which the null hypothesis is rejected at 2% level of significance , is 106.