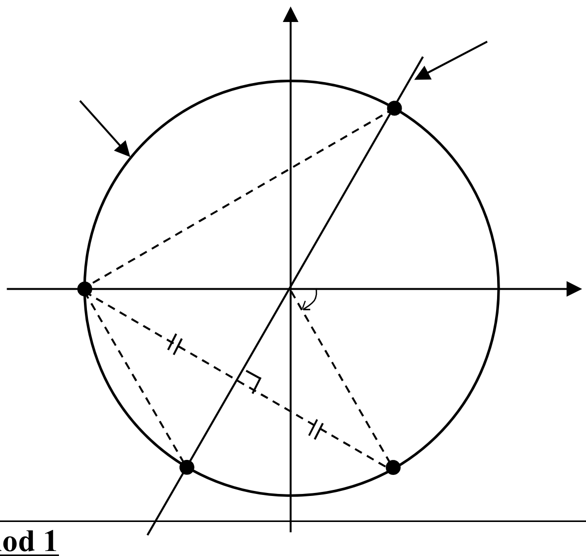
2015 Year 6 H2 Math Prelim Exam Paper 2 Mark Scheme

Qn	Suggested Solution
1(i)	Given $u_1 = \frac{1}{4}$ and $u_{n+1} = u_n - \frac{4}{n^2(n+1)(n+2)^2}, n \geq 1.$ Let P(n) be the proposition that $u_n = \frac{1}{n^2(n+1)^2}$, $n \geq 1$. Consider P(1): LHS = $u_1 = \frac{1}{4}$ (given) RHS $= \frac{1}{1^2(1+1)^2} = \frac{1}{4} = \text{LHS}$ Hence P(1) is true. Assume P(k) is true for some $k \in \mathbb{Z}^+$, i.e. $u_k = \frac{1}{k^2(k+1)^2}$ Consider P(k+1). LHS $= u_{k+1}$ $= u_k - \frac{4}{k^2(k+1)(k+2)^2}$ $= \frac{1}{k^2(k+1)^2} - \frac{4}{k^2(k+1)(k+2)^2}$ $= \frac{(k+2)^2 - 4(k+1)}{k^2(k+1)^2(k+2)^2}$ $= \frac{k^2}{k^2(k+1)^2(k+2)^2}$ $= \frac{1}{(k+1)^2(k+2)^2} = \text{RHS}$ Hence P(k) is true $\Rightarrow$ P(k+1) is true. Since P(1) is true & P(k) is true $\Rightarrow$ P(k+1) is true, by mathematical induction, $u_n = \frac{1}{n^2(n+1)^2}$ for all $n \geq 1$.

(ii)	$\sum_{n=1}^{N}\frac{1}{n^2(n+1)(n+2)^2}=\sum_{n=1}^{N}\left[\frac{1}{4}(u_n-u_{n+1})\right]$ $=\frac{1}{4}\sum_{n=1}^{N}(u_n-u_{n+1})$ $=\left\{\begin{array}{l}\frac{1}{4}[(u_1-u_2)\\+(u_2-u_3)\\ \vdots\\+(u_{N-1}-u_N)\\+(u_N-u_{N+1})]\end{array}\right.$ $=\frac{1}{4}(u_1-u_{N+1})$ $=\frac{1}{4}\left(\frac{1}{4}-\frac{1}{(N+1)^2(N+2)^2}\right)$ $=\frac{1}{16}-\frac{1}{4(N+1)^2(N+2)^2}$
(iii)	As $N\to\infty,\frac{1}{4(N+1)^2(N+2)^2}\to 0$ So $\sum_{n=1}^{N}\frac{1}{n^2(n+1)(n+2)^2}\to\frac{1}{16}-0=\frac{1}{16}$ Hence, the series is convergent and $\sum_{n=1}^{\infty}\frac{1}{n^2(n+1)(n+2)^2}=\frac{1}{16}$
(iv)	$\sum_{n=2}^{N}\frac{1}{n(n^2-1)^2}=\sum_{n=2}^{N}\frac{1}{(n-1)^2n(n+1)^2}$ $=\sum_{k=1}^{N-1}\frac{1}{k^2(k+1)(k+2)^2}$ (Let $k=n-1$) $=\frac{1}{16}-\frac{1}{4(N-1+1)^2(N-1+2)^2}$ $=\frac{1}{16}-\frac{1}{4N^2(N+1)^2}$

Qn	Suggested Solution

2	Surface area of cone $= \pi(12x)(13x) = 156\pi x^2$ Surface area of side and bottom of model $= \pi(12x)^2 + 2\pi(12x)y = 144\pi x^2 + 24\pi xy$ Cost of making the cone in dollars $= 156\pi x^2\left(0.05\right) = 7.8\pi x^2$ Cost of making the side and bottom of model in dollars $= (144\pi x^2 + 24\pi xy)\left(0.02\right) = 2.88\pi x^2 + 0.48\pi xy$ Total cost of making model in dollars $= 7.8\pi x^2 + 2.88\pi x^2 + 0.48\pi xy = 10.68\pi x^2 + 0.48\pi xy$ Given that the model costs \$100 to make, $10.68\pi x^2 + 0.48\pi xy = 100$ $y = \dfrac{100 - 10.68\pi x^2}{0.48\pi x}$ ...... (1) V, Volume of the model $= \pi\left(12x\right)^2 y + \dfrac{1}{3}\pi\left(12x\right)^2\left(5x\right) = 48\pi x^2(5x + 3y)$ Substituting equation (1) into V, $V = 48\pi x^2\left(5x + 3\left(\dfrac{100 - 10.68\pi x^2}{0.48\pi x}\right)\right)$ $= 240\pi x^3 + 30000x - 3204\pi x^3$ $= 30000x - 2964\pi x^3$ (shown) $\dfrac{dV}{dx} = 30000 - 8892\pi x^2 = 0$ $x = \sqrt{\dfrac{30000}{8892\pi}} = 1.0363$ (5 s.f.) = 1.04 (3 s.f.) $y = \dfrac{100 - 10.68\pi\left(\dfrac{30000}{8892\pi}\right)}{0.48\pi\sqrt{\dfrac{30000}{8892\pi}}} = \dfrac{63.968}{0.48\pi\left(1.0363\right)} = 40.9$ (3 s.f.) $\dfrac{d^2V}{dx^2} = -17784\pi x < 0 \quad \because x > 0$ $\therefore$ maximum volume occurs when $x = 1.04$ and $y = 40.9$.

Qn	Suggested Solution
3(i)	$z = \dfrac{4}{z^*}$ $zz^* = 4$ $\lvert z\rvert^2 = 4$ $\lvert z\rvert = 2 \ \because \lvert z\rvert > 0$ (shown)
(ii)	$1-\sqrt{3}\mathrm{i} = 2\mathrm{e}^{\mathrm{i}\left(-\frac{\pi}{3}\right)}$
(iii)	Let A, B, P, Q be points that represent the complex numbers -2, $1-\sqrt{3}\mathrm{i}$, z_1 and z_2 respectively. 
(iv)	**Method 1** Since $P(z_1)$ and $Q(z_2)$ lie on circle (centred at origin) of radius 2, $\lvert z_1\rvert = \lvert z_2\rvert = 2$ $\angle AOB = \dfrac{2\pi}{3}$ $\angle AOQ = \angle QOB = \dfrac{\pi}{3}$ (angle bisector) $\therefore \arg(z_2) = -\dfrac{2\pi}{3}$ $\Rightarrow z_2 = 2\mathrm{e}^{\mathrm{i}\left(-\frac{2\pi}{3}\right)} = -1-\sqrt{3}\mathrm{i}$ $\arg(z_1) = \dfrac{\pi}{3}$ (angle on straight line) $\Rightarrow z_1 = 2\mathrm{e}^{\mathrm{i}\left(\frac{\pi}{3}\right)} = 1+\sqrt{3}\mathrm{i}$ **Method 2**

Cartesian equation of circle:

$x^2 + y^2 = 2^2$ ---(1)

Gradient of line AB: $\dfrac{-\sqrt{3}-0}{1-(-2)} = -\dfrac{\sqrt{3}}{3} = -\dfrac{1}{\sqrt{3}}$

Gradient of perpendicular bisector $= -\dfrac{1}{-\dfrac{1}{\sqrt{3}}} = \sqrt{3}$

Cartesian equation of perpendicular bisector:

$y - 0 = \sqrt{3}(x - 0)$

$y = \sqrt{3}x$ ---(2)

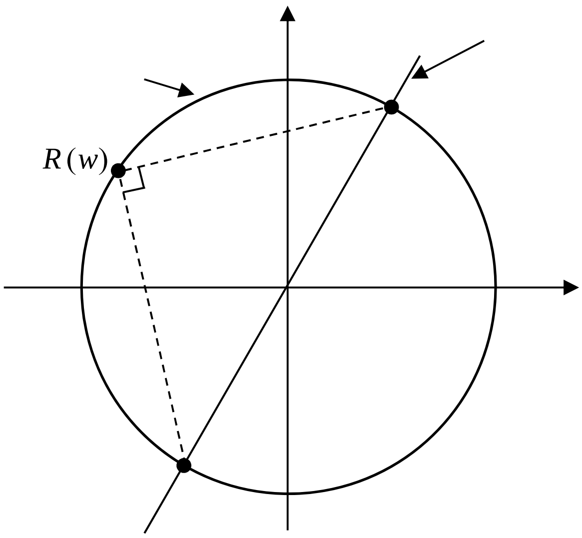
Solving (1) and (2),

$x = 1$ or $x = -1$

$y = \sqrt{3}$ $\quad y = -\sqrt{3}$

$\therefore z_1 = 1 + \sqrt{3}\mathrm{i}$ and $z_2 = -1 - \sqrt{3}\mathrm{i}$

(v) Let R be the point that represent the complex numbers w.



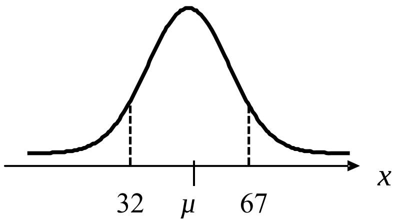
Note that PQ forms the diameter of the circle centred at origin with radius 2 units.

$\arg(z_1 - w) - \arg(z_2 - w)$

$= \angle PRQ = \dfrac{\pi}{2}$ (right angle in semicircle)

Qn	Suggested Solution
4(i)(a)	Area covered by weed at the end of the first week $=0.9(500+80)$ $=0.9(500)+80(0.9)\ \text{m}^2$ Area covered by weed at the end of the second week $=0.9(0.9(500)+80(0.9)+80)$ $=0.9^2(500)+80(0.9+0.9^2)$ $=541.8\quad \text{m}^2$
(b)	Area covered by weed at the end of the *n*th week $=0.9^n\left(500\right)+80(0.9+0.9^2+...+0.9^n)$ $=0.9^n\left(500\right)+0.9(80)\left(\dfrac{1-0.9^n}{0.1}\right)$ $=0.9^n\left(500\right)+720\left(1-0.9^n\right)\text{m}^2$ Therefore, $k=720$.
(c)	As $n\to\infty, 0.9^n\to 0$ So $0.9^n\left(500\right)+720(1-0.9^n)\to 0(500)+720(1-0)=720$ Hence the area covered with weed at the end of the week in the long run is 720 m^2.
(ii)(a)	Change in area covered with weed in the *n*th week $=80-(50+10(n-1))$ $=40-10n\ \text{m}^2$
(b)	Area covered with weed at the end of the *n*th week $=500+\sum_{r=1}^{n}(40-10r)$ $=500+\sum_{r=1}^{n}40-10\sum_{r=1}^{n}r$ $=500+40n-10\left[\dfrac{n}{2}(1+n)\right]$ $=500+35n-5n^2\ \text{m}^2$

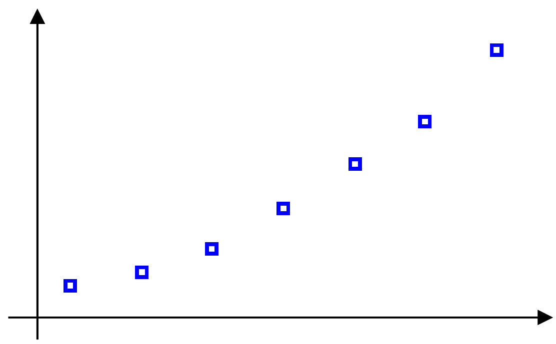
Qn	Suggested Solution
5(i)	Each level/strata is not proportionally represented and hence not stratified sampling.
(ii)	Divide the students of the school into strata in terms of age or gender. Select the number of students in each stratum e.g. 20 male and 20 female students for the survey in order to meet the quota. Then stand at the entrance of the school at the start of a school day to survey the first 20 male and first 20 female students that enter the school. One disadvantage of quota sampling is that the sample obtained is likely to be biased as students who come to school later will not have a chance to be considered (or selected).

Qn	Suggested Solution
6(i)	Given $P(M < 32) = P(M > 67)$, By symmetry, $\mu = \frac{32+67}{2} = 49.5$ 
(ii)	Given that M and F are times spent in minutes by a randomly chosen male and female customer, respectively. i.e. $M \sim \mathrm{N}(49.5,\ 18^2)$, $F \sim \mathrm{N}(71,\ 35^2)$ Then $\overline{F} = \frac{F_1 + F_2 + F_3 + F_4 + F_5}{5} \sim \mathrm{N}(71,\ \frac{35^2}{5})$ $\overline{F} - 3M \sim \mathrm{N}(71 - 3(49.5),\ \frac{35^2}{5} + 3^2(18^2))$ i.e. $\overline{F} - 3M \sim \mathrm{N}(-77.5,\ 3161)$ $P(\overline{F} > 3M) = P(\overline{F} - 3M > 0) \approx 0.084033 = 0.0840$ (3 s.f.)
(iii)	For $F \sim \mathrm{N}(71,\ 35^2)$, $P(F < 0) = 0.0213$ The probability that the time spent by a female customer being less than zero is not negligible. <u>OR</u> The probability 0.0213 suggests that for every 100 customers, more than 2 spend less than zero minutes at the salon. Since time is a non-negative quantity, the normal distribution with the given mean and standard deviation is not an appropriate model.

Qn	Suggested Solution
7(i)	Let X denote the number of bits that are corrupted during the transmission process, out of 8 bits. $X \sim \text{B}(8, 0.03)$. $\text{P}(X \geq 2)$ $= 1 - \text{P}(X \leq 1)$ $= 0.022341$ (to 5 s.f.) $= 0.0223$ (3 s.f.) (shown)
(ii)	Required probability $= \text{P}(X = 0 \mid X < 2)$ $= \dfrac{\text{P}(\{X = 0\} \cap \{X < 2\})}{\text{P}(X < 2)}$ $= \dfrac{\text{P}(X = 0)}{\text{P}(X \leq 1)}$ $= \dfrac{0.78374}{1 - 0.0233}$ $= 0.802$ (3 s.f.)
(iii)	Let Y denote the number of bytes that are corrupted during the transmission process, out of 100 bytes. $Y \sim \text{B}(100, 0.0223)$. Since $n = 100 \geq 50$ is large and p is small enough such that $np = 0.0223(100) = 2.23 < 5$, $Y \sim \text{P}_\text{o}(2.23)$ approximately. $\text{P}(5 < Y < 10)$ $= \text{P}(Y \leq 9) - \text{P}(Y \leq 5)$ $= 0.0263$ (3 s.f.)

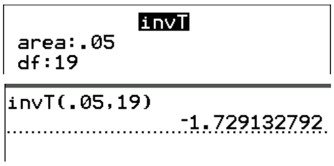
Qn	Suggested Solution
8(i)	Required number of ways $= 5!\times 2!$ $= 240$
(ii)	Required number of ways $= {}^{3}C_{1} \times 2!\times 5!$ $= 720$ <u>Explanation:</u> Number of ways to choose a brother = ${}^{3}C_{1}$ Number of ways to arrange the 2 parents in the group of chosen brother and parents = 2! Number of ways to arrange the group of 3 with the remaining 4 people =5!
(iii)	Required number of ways $= (4-1)!\times {}^{4}P_{3}$ $= 144$ <u>Explanation:</u> Number of ways to arrange the remaining 4 people at a round table = $(4-1)!$ Number of ways to slot in the brothers = ${}^{4}P_{3}$

Qn	Suggested Solution
9(i)	Let $\mathrm{P}(A)=\alpha$ and $\mathrm{P}(B)=\beta$. $\mathrm{P}(A)+\mathrm{P}(B)=\mathrm{P}(A\cup B)+\mathrm{P}(A\cap B)$ $=0.55+0.05=0.6$ $\Rightarrow \alpha+\beta=0.6\cdots\cdots(1)$ Since A and B are independent, $\mathrm{P}(A\cap B)=\mathrm{P}(A)\mathrm{P}(B)$ $\Rightarrow \alpha\beta=0.05$ i.e. $\alpha=\dfrac{0.05}{\beta}\cdots\cdots(2)$ Sub (2) into (1): $\dfrac{0.05}{\beta}+\beta=0.6$ $\beta^2-0.6\beta+0.05=0$ $\beta=0.5$ or $\beta=0.1\cdots\cdots(3)$ Sub (3) into (1): $\alpha=0.1$ or $\alpha=0.5$ Since $\alpha<\beta$, $\mathrm{P}(A)=0.1$ and $\mathrm{P}(B)=0.5$
(ii)	$\mathrm{P}(A'\cup B'\mid C)$ $=1-\mathrm{P}((A'\cup B')'\mid C)$ $=1-\mathrm{P}(A\cap B\mid C)$ $=1-\dfrac{\mathrm{P}(A\cap B\cap C)}{\mathrm{P}(C)}$ $=1-\dfrac{\mathrm{P}(A\cap B\cap C)}{0.5}=0.95$ $\therefore \mathrm{P}(A\cap B\cap C)=0.05\times0.5=0.025.$
(iii)	Since $\mathrm{P}(A\cap C)\geq\mathrm{P}(A\cap B\cap C)=0.025>0$, A and C are not mutually exclusive.

Qn	Suggested Solution
10(i)	
(ii)	**(a)** $r = 0.9781$ (4 d.p) **(b)** $r = 0.9985$ (4 d.p) (Insert GC screenshot)
(iii)	As the r-value between x^2 and y is closer to 1 than the r-value between x and y indicating a stronger positive linear relationship between x^2 and y, $y = c + dx^2$ is the better model. (Insert reason from part (i) using scatter diagram)
(iv)	As y is the dependent variable and x is the independent (controlled) variable, neither the regression line of x on y nor the regression line of x^2 on y should be used.
(v)	Equation of regression line of y on x^2, $y = 399.60 + 0.055326x^2$ (5 s.f) $y = 400 + 0.0553x^2$ (3 s.f) When $y = 440$, $440 = 399.60 + 0.055326x^2$ $x = 27.0$ $(x > 0)$ Alternatively, (ER 2010) Using regression line y on x: $y = 370.96 + 2.75x$ (5 s.f) $x = 25.1$ The estimate is reliable as there is a strong positive linear relationship between x^2 and y and interpolation is used.

Qn	Suggested Solution	Mark Scheme
11(i)	1) The number of calls received by each hotline occurs at a constant average rate per unit time. 2) Each call received by each hotline is made independently of other calls received by the same hotline.	
(ii)	Let C denote the number of calls received by the complaint hotline in a randomly chosen period of t minutes. $C \sim \text{Po}(3t)$. Given $\text{P}(C \le 1) = 1\times 10^{-5}$ $\Rightarrow \text{P}(C=0)+\text{P}(C=1) = 1\times 10^{-5}$ $\Rightarrow \text{e}^{-3t}(1+3t) = 1\times 10^{-5}$ From GC, $t = 4.74 = 5$ (to nearest whole number)	
(iii)	Let X and Y denote the number of calls received by the enquiry hotline and the complaint hotline in a randomly chosen period of 10 minutes, respectively. $X \sim \text{Po}(10(2))$ i.e. $X \sim \text{Po}(20)$. $Y \sim \text{Po}(10(3))$ i.e. $Y \sim \text{Po}(30)$. Since X and Y are independent, $X+Y \sim \text{Po}(20+30)$ i.e. $X+Y \sim \text{Po}(50)$. $\text{P}(X+Y > 40)$ $= 1-\text{P}(X+Y \le 40)$ $= 0.914$ (to 3 s.f.)	
(iv)	Since λ_X and λ_Y are sufficiently large such that $\lambda_X = 20 > 10$, and $\lambda_Y = 30 > 10$, $X \sim \text{N}(20,20)$ approximately and $Y \sim \text{N}(30,30)$ approximately. Since X and Y are independent, $Y-X \sim \text{N}(30-20, 30+20)$ approximately i.e. $Y-X \sim \text{N}(10,50)$ approximately. $\text{P}(Y > X)$ $= \text{P}(Y-X > 0)$ $= \text{P}(Y-X > 0.5)$ (by continuity correction) $= 0.910$ (to 3 s.f.)	

Qn	Suggested Solution	
12(i)	An unbiased estimate for the population mean is $\bar{x} = \frac{\sum(x-120)}{n} + 120 = \frac{7380}{90} + 120 = 202$ An unbiased estimate for the population variance is $s^2 = \frac{1}{n-1}\left[\sum(x-120)^2 - \frac{(\sum(x-120))^2}{n}\right]$ $= \frac{1}{89}\left[629982 - \frac{7380^2}{90}\right] \approx 278.90 = 279$ (3 s.f.)	
(ii)	Let X represent the cholesterol level (in mg/dL) of a city-dweller with population mean μ. To test H_0: $\mu = 199$ vs H_1: $\mu > 199$ Perform 1-tail test at 5% significance level Under H_0, $\bar{X} \sim N\left(\mu_0, \frac{s^2}{n}\right)$ approximately, by Central Limit Theorem, since n is large, where $\mu_0 = 199$. From the sample, $\bar{x} = 202$, $s = \sqrt{278.90}$ and $n = 90$. Using z-test, p-value = 0.0442 (3 s.f.) Since p-value < 0.05, we reject H_0 and conclude that there is sufficient evidence at 5% significance level that the **mean** cholesterol level of city-dwellers has increased (or the **mean** cholesterol level of city-dwellers is higher than the average reported in 2012).	
(iii)	Let Y represent the cholesterol level (in mg/dL) of a person living in the countryside with population mean μ. To test H_0 : $\mu = 199$ vs H_1 : $\mu \neq 199$ Perform 2-tail test at 10% significance level. Assume that cholesterol levels of people living in the countryside are normally distributed. Since the sample size of 20 is small and σ^2 is unknown, Under H_0, $T = \frac{\bar{Y} - \mu_0}{S/\sqrt{n}} \sim t(n-1)$, where $\mu_0 = 199$	

	From the sample, $n = 20,\ s^2 = \frac{n}{n-1}(\text{sample variance}) = \frac{20}{20-1}(281.5) = 296.32$ For H_0 to be NOT rejected at 10% significance level, p-value > 0.10.  $\Rightarrow -1.7291 < \frac{\overline{y}-199}{\sqrt{\frac{296.32}{20}}} < 1.7291$ $\Rightarrow -21.047 < \overline{y}-199 < 21.047$ $\Rightarrow 192.34 < \overline{y} < 205.66$ $\therefore$ Set of values of $\overline{y} = \{\overline{y} \in \mathbb{R} : 193 \le \overline{y} \le 205\}$ (3 s.f.)	
	For Betty's *t*-test to be valid, the cholesterol levels of people living in the countryside are assumed to be normally distributed. For Abbey's *z*-test to be valid, the cholesterol levels of people living in the cities need not be assumed to be normally distributed because the distribution of the sample mean is approximately normal by Central Limit Theorem since the **sample size ($n \ge 50$) is large**.	