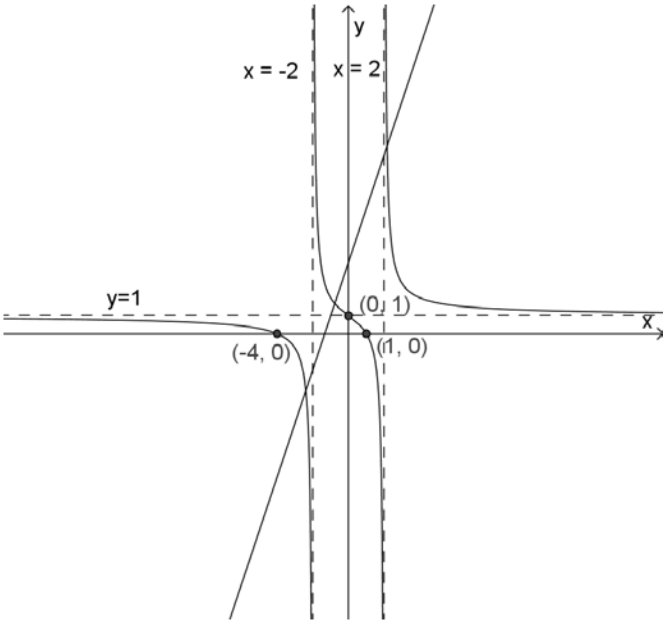
Suggested Solution: 2016 SH2 H1 Mathematics Preliminary Examination

1 (i)

$$y=\frac{3x}{x^2-4}+1$$



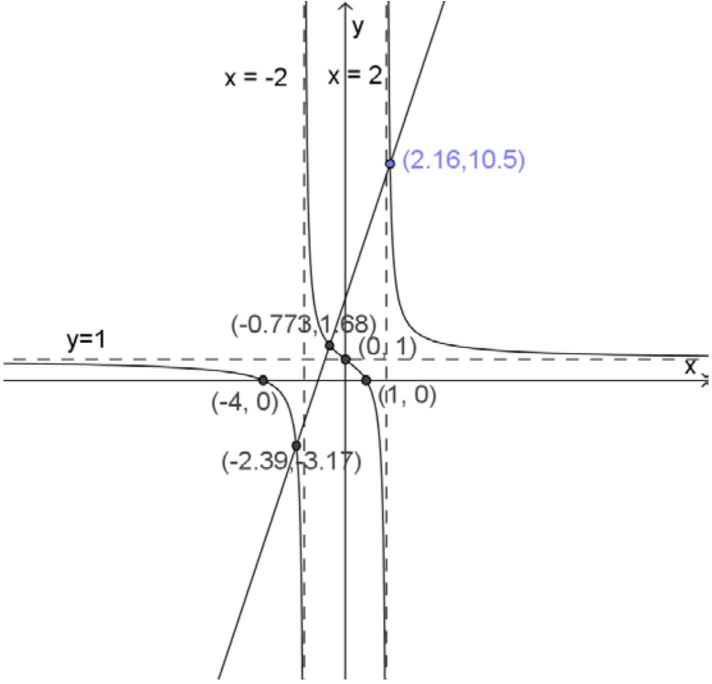
1 (ii)

$$\frac{x}{x^2-4}>x+1$$

$$\frac{3x}{x^2-4}>3x+3$$

$$\frac{3x}{x^2-4}+1>3x+4$$

Sketch the graph if $y = 3x + 4$ on the same diagram in (i):



From graph,

$x<-2.39$ or $-2<x<-0.773$ or $2<x<2.16$

2(i)	$y=\ln\left(\dfrac{x}{3-x}\right)$ $\quad=\ln x-\ln\left(3-x\right)$ $\dfrac{\mathrm{d}y}{\mathrm{d}x}=\dfrac{1}{x}-\dfrac{-1}{3-x}=\dfrac{1}{x}+\dfrac{1}{3-x}$ $\dfrac{\mathrm{d}y}{\mathrm{d}x}=k$ $\dfrac{1}{x}+\dfrac{1}{3-x}=k$ $\dfrac{3-x+x}{x\left(3-x\right)}=k$ $3=kx\left(3-x\right)$ $3=3kx-kx^2$ $kx^2-3kx+3=0$ (shown)
2(ii)	No real roots implies Discriminant < 0 $\left(-3k\right)^2-4k\left(3\right)<0$ $9k^2-12k<0$ $3k\left(3k-4\right)<0$ Therefore, $0<k<\dfrac{4}{3}$.
2(iii)	Since 2 is not in the range $0<k<\dfrac{4}{3}$, there is a point on the curve C with gradient 2.

3 (i)

$$
\begin{aligned}
D &= \sqrt{(x-50)^2+y^2} \\
&= \sqrt{(x-50)^2+\left(\sqrt{x}\right)^2} \\
&= \sqrt{x^2-100x+2500+x} \\
&= \sqrt{x^2-99x+2500}
\end{aligned}
$$

3 (ii)

$$D=\sqrt{x^2-99x+2500}$$

Applying Chain Rule:

$$
\begin{aligned}
\frac{\mathrm{d}D}{\mathrm{d}x} &= \frac{1}{2}\left(x^2-99x+2500\right)^{-\frac{1}{2}}(2x-99) \\
&= \frac{2x-99}{2\sqrt{x^2-99x+2500}}
\end{aligned}
$$

For stationary point(s), $\frac{\mathrm{d}D}{\mathrm{d}x}=0$, so we have

$$
\begin{aligned}
2x-99&=0 \\
x&=49.5
\end{aligned}
$$

<u>Method 1 (1st Derivative Test)</u>

x	49.4	49.5	49.6
$\frac{\mathrm{d}D}{\mathrm{d}x}$	−0.0141762	0	0.0141762
	< 0	$= 0$	> 0

Thus, D is a minimum at $x = 49.5$.

<u>Method 2 (2nd Derivative Test)</u>

$$\frac{\mathrm{d}^2D}{\mathrm{d}x^2}=-\frac{(2x-99)^2}{4\left(x^2-99x+2500\right)^{\frac{3}{2}}}+\frac{2}{2\sqrt{x^2-99x+2500}}$$

When $x=49.5$, $\frac{\mathrm{d}^2D}{\mathrm{d}x^2}=0.14177>0$.

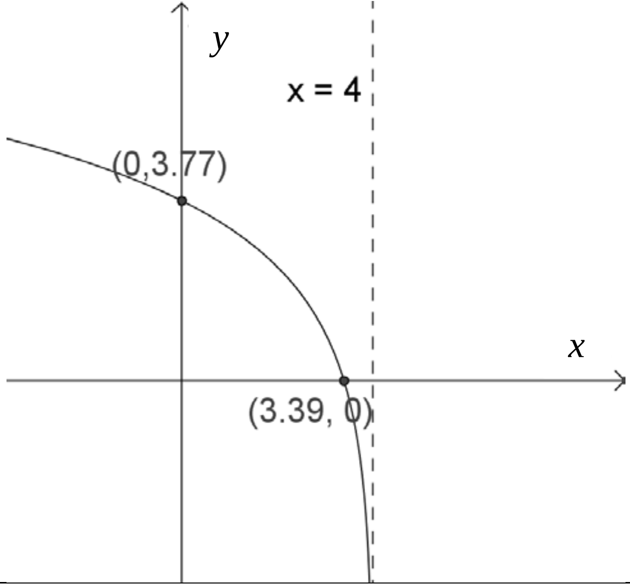
Thus, D is a minimum at $x = 49.5$.

The minimum D is

$$\sqrt{(49.5)^2-99(49.5)+2500}=7.05 \text{ metres.}$$

4 (a)(i)	Let $y=\sqrt{3}x-\frac{\pi}{\sqrt{x}}$ $=\sqrt{3}x-\pi x^{0.5}$ $\frac{\mathrm{d}y}{\mathrm{d}x}=\sqrt{3}-\frac{\pi}{2\sqrt{x}}$
4 (a)(ii)	$\mathrm{f}\left(x\right)=\mathrm{e}^{\left(k^2\sqrt{x}+k\right)}$ $\mathrm{f}\,'\left(x\right)=\frac{k^2}{2\sqrt{x}}\mathrm{e}^{\left(k^2\sqrt{x}+k\right)}$
4 (b)	By GC, $\frac{\mathrm{d}y}{\mathrm{d}x}\approx 0.80685=0.81$ (2 d.p)
4 (c)	$\int_0^1\left(\mathrm{e}^{\frac{x}{2}}+3\right)^2\,\mathrm{d}x=\int_0^1\left(\mathrm{e}^{x}+6\mathrm{e}^{\frac{x}{2}}+9\right)\,\mathrm{d}x$ $=\left[\mathrm{e}^{x}+12\mathrm{e}^{\frac{x}{2}}+9x\right]_0^1$ $=\left(\mathrm{e}+12\mathrm{e}^{\frac{1}{2}}+9\right)-\left(1+12+0\right)$ $=\mathrm{e}+12\mathrm{e}^{\frac{1}{2}}-4$

5 (i)



5 (ii)

$$\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{-2}{4-x}$$

When $x=0.5$, $\frac{\mathrm{d}y}{\mathrm{d}x}=\frac{-2}{{}^{7}\!/_{2}}=-\frac{4}{7}$

Gradient of normal = $-\frac{1}{-4/7}=\frac{7}{4}$

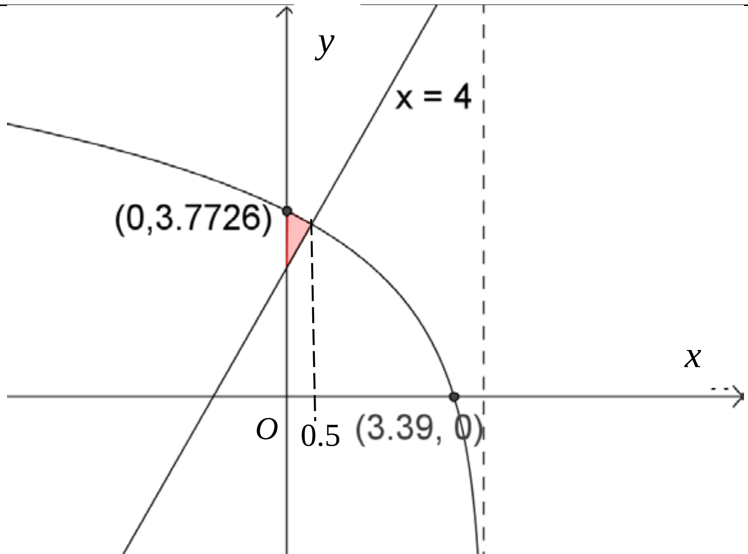
When $x=0.5$, $y=1+2\ln\left(3.5\right)$:

Equation of normal is

$$y-\left(1+2\ln\left(3.5\right)\right)=\frac{7}{4}\left(x-0.5\right)$$

$$y=\frac{7}{4}x+\frac{1}{8}+2\ln\left(\frac{7}{2}\right)$$

5 (iii)



Integrating with respect to x-axis:

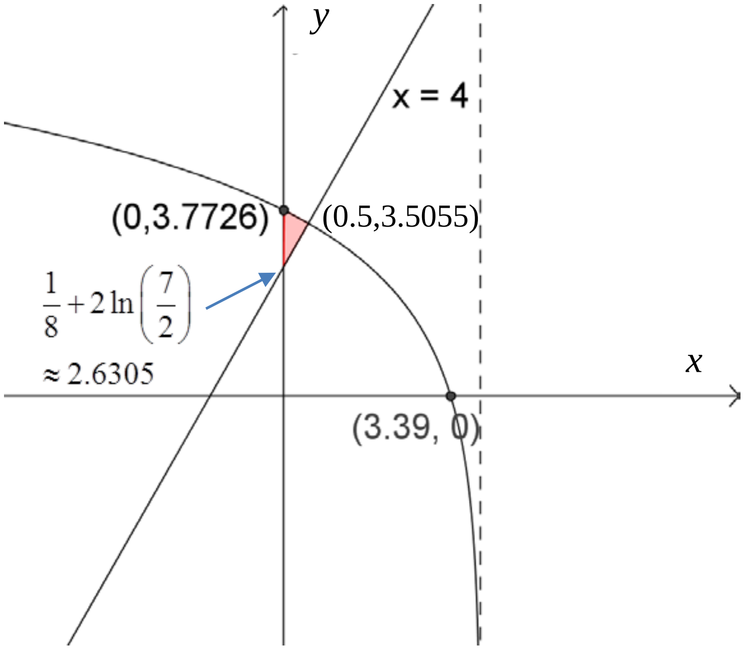
Area of required region (shaded)

= "Area bounded between curve C and the normal between x = 0 and x = 0.5"

$$= \int_0^{0.5} \left(1+2\ln(4-x)-\left(\frac{7}{4}x+\frac{1}{8}+2\ln\left(\frac{7}{2}\right)\right)\right) \mathrm{d}x$$

$= 0.2870$ (4d.p) by GC

<u>Alternative Method</u> (Ir ting with respect to y-axis)



$y = 1+2\ln(4-x)$

$\Rightarrow y-1 = 2\ln(4-x)$

$\Rightarrow \dfrac{y-1}{2} = \ln(4-x)$

$\Rightarrow \mathrm{e}^{\frac{y-1}{2}} = 4-x$

$\Rightarrow x = 4-\mathrm{e}^{\frac{1}{2}y-\frac{1}{2}}$

Area of required region (shaded)

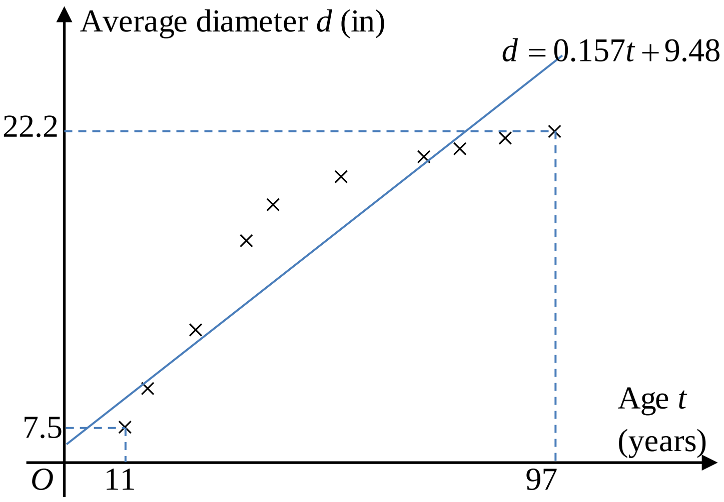
$$= \int_{3.5055}^{3.7726} \left(4-\mathrm{e}^{\frac{1}{2}y-\frac{1}{2}}\right) \mathrm{d}y + \frac{1}{2}\left(\frac{1}{2}\right)(3.5055-2.6305)$$

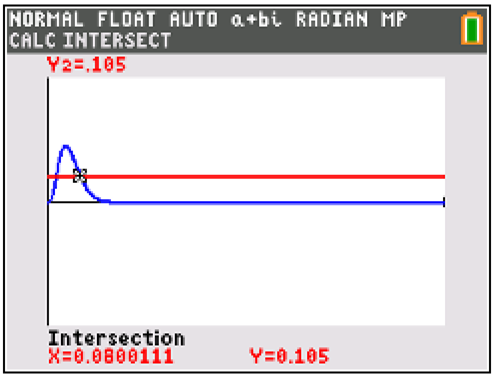
$\approx 0.0682611 + 0.21875$

$= 0.2870$ (4d.p) by GC

6 (a)	Assign a number from 1 to N to each of the students, where N represents the student population size OR obtain a list of the students from the administration office in order of their identification numbers or registration numbers. Next, determine the sampling interval size $k = \frac{1}{0.02} = 50.$ Randomly select a student from the 1st 50 students. Select every 50th student thereafter until the required sample is obtained.
6 (b)	**Advantages**: • **Representativeness of Sample** Quota sampling allows the survey to capture the responses that represent various groups of students (e.g. different PM classes, or CCAs); this may be preferred as certain homeroom or sports facilities may not be in as good a condition as others, and the representation of each group will ensure that the results will not be biased towards those who are often using these less functional facilities or towards those who are often using the more functional facilities. • **Efficiency of Collecting the Sample** Quota sampling may be more efficient as systematic sampling in this case requires the surveyor to identify the selected respondents and to contact them, which can be time consuming (e.g. student selected may be on MC on day of survey, selected students do not respond to online survey etc) **Disadvantages**: • **Non-randomness/Selection Bias** Quota sampling is non-random and may contain selection bias, where the surveyor chooses people who may appear more friendly or choose students in the canteen only at a selected time period. This results in certain students having no chance of being selected at all, which may affect the validity of the survey results. • **Non-representativeness of Sample** Quota sampling may result in a group (e.g. one entire cohort, or people coming later to the canteen etc) being excluded entirely from the selection, which may result in the data collected being an inaccurate representation of the entire school population.

7 (i)	$\mathrm{P}(A \mid B)=p$ $\dfrac{\mathrm{P}(A\cap B)}{\mathrm{P}(B)}=p$ $\mathrm{P}(A\cap B)=p^2$ $\mathrm{P}(A\cup B)=\mathrm{P}(A)+\mathrm{P}(B)-\mathrm{P}(A\cap B)$ $0.8=3p-1+p-p^2$ $-p^2+4p-1=0.8$ Using GC, *p* = 0.5167603 or *p* = 3.4832397 (reject since *p*>1) Therefore, *p* = 0.517 (2 d.p)
7 (ii)	$\mathrm{P}(A'\cup B')=1-\mathrm{P}(A\cap B)$ $=1-p^2$ $=1-(0.51676)^2$ $=0.733$ (3s.f)
7 last part	$\mathrm{P}(A)=(3(0.52)-1)=0.56$ $\mathrm{P}(A \mid B)=0.52$ *A* and *B* are not independent since $\mathrm{P}(A)\neq \mathrm{P}(A \mid B)$. **Alternative Method** $\mathrm{P}(A)\mathrm{P}(B)=(3p-1)p$ $=(3(0.52)-1)0.52$ $=0.2912$ $\mathrm{P}(A\cap B)=p^2=(0.52)^2=0.2704$ *A* and *B* are not independent since $\mathrm{P}(A)\mathrm{P}(B)\neq \mathrm{P}(A\cap B)$

8 (i)	
(ii)	r = 0.918 (to 3 s.f)
(iii)	The regression line of d on t is $d = 0.15657t + 9.4765$ $d = 0.157t + 9.48$ (3s.f) For every increase in 1 year of the age of a teak tree, the trunk diameter increases **approximately** by 0.157 inch.
(iv)	When t = 40, $d = 0.15657(40) + 9.4765 = 15.7393$ d = 15.7 (3 s.f) Hence the diameter is 15.7 inches (to 3 s.f.) Since $t = 40$ is within the range of values of t, [11, 97] and r = 0.918 is close to 1, suggesting a strong positive linear correlation between d and t, therefore the estimate is reliable.
(v)	A linear model predicts the diameter will keep increasing indefinitely. Therefore a linear model is not appropriate. **OR** The diameter of any teak tree should approach a constant value over time, hence a linear model is not suitable.
(vi)	The regression line of d on t is $d = 0.15657t + 9.4765 + k$ $d = 0.157t + 9.48 + k$ (3s.f)

9 (i)	The eye colour of a person is independent of that of another person. **OR** The probability that the eye colour of a person is blue is constant.
9 (ii)	$Y \sim \mathrm{B}(70, p)$ $\mathrm{P}(Y = 3) = 0.105$ $\binom{70}{3} p^3 (1-p)^{67} = 0.105$ from MF15 $54740 p^3 (1-p)^{67} = 0.105$  (*Tip: Set the window for* $x_{min} = 0$ and $x_{max} = 1$ since x (which is p in the above equation) represents probability) From graph, since $p > 0.05$, $p = 0.080011$ $= 0.0800$ (4 d.p)
9 (iii)	$Y \sim \mathrm{B}(70, 0.08)$ Then $\mathrm{P}(5 \le Y < 21)$ $= \mathrm{P}(Y \le 20) - \mathrm{P}(Y \le 4)$ $= 0.668$ (3 s.f)
9 (iv)	Since $n = 70$ is sufficiently large, $np = 5.6 > 5$ and $nq = 64.4 > 5$, hence $Y \sim \mathrm{N}(5.6,\ 5.152)$ approximately. $\mathrm{P}(Y > 9) = \mathrm{P}(Y > 9.5)$ $= 0.0429$ (to 3 s.f.)

10 (i)(a)

Let E and F denote the event that the sum of scores is at least 8 and the event that the difference between the scores is at least 4 respectively.

Sum of Scores

1st die \ 2nd die	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$\mathrm{P}(E)=\frac{1}{6}\times\frac{1}{6}\times 15=\frac{5}{12}$$

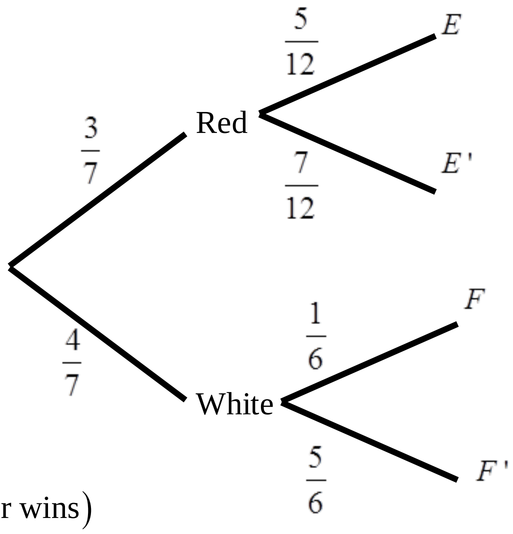
10 (i)(b)

Difference between Scores

1st die \ 2nd die	1	2	3	4	5	6
1	0	1	2	3	4	5
2	1	0	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	0	1	2
5	4	3	2	1	0	1
6	5	4	3	2	1	0

$$\mathrm{P}(F)=\frac{1}{6}\times\frac{1}{6}\times 6=\frac{1}{6}$$

10 (ii)



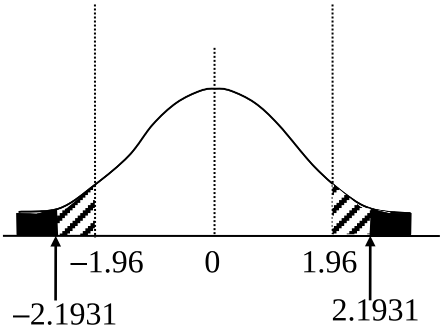
$$\mathrm{P}(\text{player wins})$$

$$=\mathrm{P}(\text{red and sum at least 8})+\mathrm{P}(\text{white and diff at least 4})$$

$$=\left(\frac{3}{7}\right)\left(\frac{5}{12}\right)+\left(\frac{4}{7}\right)\left(\frac{1}{6}\right)$$

$$=\frac{23}{84}$$

10 **(iii)**	P(red ball selected | player wins) $= \dfrac{\text{P(red ball selected } \cap \text{player wins)}}{\text{P(player wins)}}$ $= \dfrac{\left(\dfrac{3}{7}\right)\left(\dfrac{5}{12}\right)}{\dfrac{23}{84}} = \dfrac{15}{23}$

11 (i)	Unbiased estimate of μ, $\bar{x} = \frac{\sum x}{n} = \frac{38100}{40} = 952.50$ Unbiased estimate of σ^2, $s^2 = \frac{1}{39}(731800)$ $= 18764.10$ (2 d.p.)
11 (ii)	$\mathrm{H_0}$: $\mu = 1000$ $\mathrm{H_1}$: $\mu \neq 1000$ Level of Significance: 5% Under $\mathrm{H_0}$, $Z = \frac{\overline{X} - 1000}{S/\sqrt{40}} \sim \mathrm{N}(0,1)$ approximately by Central Limit Theorem Method 1: *Compare critical region and observed test statistic* Critical region: $\lvert z \rvert > 1.960$ $z = \frac{952.5 - 1000}{s/\sqrt{40}} \approx -2.1931$ $\left(\text{where } s = \sqrt{\frac{731800}{39}} \right)$ Since $\lvert z \rvert = 2.1931 > 1.960$, we reject $\mathrm{H_0}$.  Method 2: *Using p-value* *p*-value = 0.028299 Since *p*-value = 0.0283 < 0.05, we reject $\mathrm{H_0}$. We conclude that there is sufficient evidence at 5% level of significance that the mean amount of loans borrowed by its clients differs from \$1000.
11 (iii)	The meaning of 'at the 5% significance level' is that there is a probability of 0.05 of rejecting the claim that the **mean** amount of loans borrowed by its clients is \$1000 given that it is true. **OR**

	The meaning of ‘at the 5% significance level’ is that there is a probability of 0.05 that it was **wrongly concluded** that the **mean** amount of loans borrowed by its clients differs from $1000.
11(iv)	Test Statistic: $z = \frac{k-1000}{250/\sqrt{40}}$ Do not reject $\text{H}_0 \Rightarrow -1.96 < z < 1.96$ $-1.96 < \frac{k-1000}{250/\sqrt{40}} < 1.96$ $-77.4758 < k-1000 < 77.4758$ $922.524 < k < 1077.4758$ $922.53 \le k \le 1077.47$

12 (a)	$X \sim \mathrm{N}(\mu,\sigma^2)$ By symmetry, $\mu = \frac{18.1+21.9}{2} = 20$ $\mathrm{P}\left(X < 18.1\right) = 0.2$ $\mathrm{P}\left(Z < \frac{18.1-20}{\sigma}\right) = 0.2$ $\frac{18.1-20}{\sigma} = -0.841621$ --- (1) ***OR*** $\mathrm{P}\left(X > 21.9\right) = 0.2$ $\mathrm{P}\left(X < 21.9\right) = 0.8$ $\mathrm{P}\left(Z < \frac{21.9-20}{\sigma}\right) = 0.8$ $\frac{21.9-20}{\sigma} = 0.841621$ --- (2) $\sigma = 2.26 \ \left(3\text{s.f}\right)$
12 (bi)	Let D and E be the battery life of a DuraSell and Energise battery respectively. $D \sim \mathrm{N}\left(m, 1.5^2\right)$ and $E \sim \mathrm{N}\left(19.8, 2.7^2\right)$ $D - m \sim \mathrm{N}\left(0, 1.5^2\right)$ $\mathrm{P}\left(m-1 < D < m+1\right)$ $= \mathrm{P}\left(-1 < D - m < 1\right)$ $= 0.495015066$ $= 0.495$ (3s.f) Alternative Method (Standardisation) $\mathrm{P}\left(m-1 < D < m+1\right)$ $= \mathrm{P}\left(\frac{-1}{1.5} < \frac{D-m}{1.5} < \frac{1}{1.5}\right)$ $= \mathrm{P}\left(\frac{-2}{3} < Z < \frac{2}{3}\right)$ $= 0.495015066$ $= 0.495$ (3s.f)
12 (bii)	Let X be the number of DuraSell batteries, out of 3, with battery life within 1 hour of its mean. $X \sim \mathrm{B}(3, 0.49502)$ $\mathrm{P}\left(X = 2\right) = 0.371$ (3s.f) Alternative Method

	Let A be the event that a DuraSell battery has battery life within 1 hour of its mean. P(exactly 2 DuraSell batteries with battery life within 1 hour of its mean) $=\mathrm{P}(A'AA)+\mathrm{P}(AA'A)+\mathrm{P}(AAA')$ $=\mathrm{P}(A')\mathrm{P}(A)\mathrm{P}(A)+\mathrm{P}(A)\mathrm{P}(A')\mathrm{P}(A)+\mathrm{P}(A)\mathrm{P}(A)\mathrm{P}(A')$ $=3(1-0.49502)(0.49502)^2$ $= 0.371$ (3s.f)
12 (biii)	$D-E \sim \mathrm{N}(2.6, 9.54)$ $\mathrm{P}(D>E)$ $=\mathrm{P}(D-E>0)$ $= 0.800$ (3s.f)
12 (biv)	$\overline{E} \sim \mathrm{N}\left(19.8, \frac{2.7^2}{50}\right)$ $\mathrm{P}(\overline{E}>19) \approx 0.98192$ Expected number of samples that passes quality check = 100(0.98192) = 98.2 (3 s.f)