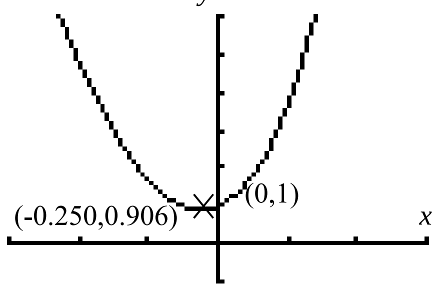
2014 Year 6 Preliminary Examination Solution & Comments

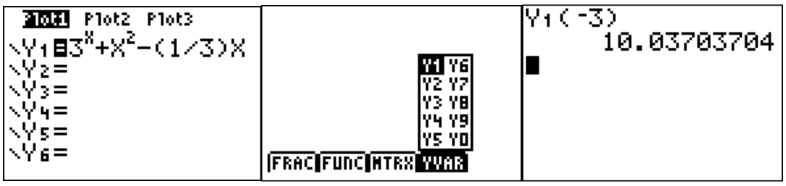
Qn	Suggested Solution
1(i)	$\ln\left(\dfrac{4x^3}{3x+2}\right)-2\ln x=\ln(x-4)$ $\ln\left(\dfrac{4x^3}{3x+2}\right)-\ln x^2=\ln(x-4)$ $\ln\left(\dfrac{4x^3}{(3x+2)x^2}\right)=\ln(x-4)$ $\dfrac{4x}{3x+2}=x-4$ $3x^2-14x-8=0$ $x=\dfrac{7+\sqrt{73}}{3}$ or $x=\dfrac{7-\sqrt{73}}{3}=-0.51467(5\text{ s.f.})$ (rej. for $\ln(x-4)$ to be defined)
(ii)	From **(i)**, replace x by e^{2x}: $\ln\left(\dfrac{4(e^{2x})^3}{3e^{2x}+2}\right)-2\ln e^{2x}=\ln(e^{2x}-4)$ $\ln\left(\dfrac{4e^{6x}}{3e^{2x}+2}\right)-4x=\ln(e^{2x}-4)$ Hence, $e^{2x}=\dfrac{7+\sqrt{73}}{3}$ $2x=\ln\left(\dfrac{7+\sqrt{73}}{3}\right)$ $x=\dfrac{1}{2}\ln\left(\dfrac{7+\sqrt{73}}{3}\right)$

2	$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{x}{2x+3}\right)=\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\frac{1}{2}(2x+3)-\frac{3}{2}}{2x+3}\right)$ $=\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{2}-\frac{3}{2(2x+3)}\right)$ $=-\frac{3}{2}(-1)(2x+3)^{-2}(2)$ $=\frac{3}{(2x+3)^2}$
(a)	$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\ln x}{\ln x^2+3}\right)=\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\ln x}{2\ln x+3}\right)$ $=\frac{3}{x(2\ln x+3)^2}$ For f’ to exist, $x\neq 0$ and $x>0$ and $2\ln x+3\neq 0$ Solving $2\ln x+3=0$ $\ln x=-\frac{3}{2}$ $x=\mathrm{e}^{-\frac{3}{2}}$ The set of values of x where the derivative exists is $\left\{x\in\mathbb{R}: x>0 \text{ and } x\neq \mathrm{e}^{-\frac{3}{2}}\right\}$. **<u>Alternative way of presentation</u>** The set of values of x where the derivative exists is $\left\{x\in\mathbb{R}: 0<x<\mathrm{e}^{-\frac{3}{2}} \text{ or } x>\mathrm{e}^{-\frac{3}{2}}\right\}$.
(b)	$\int_{-1}^{0}\frac{1}{(4x+6)^2}\,\mathrm{d}x=\int_{-1}^{0}\frac{1}{4(2x+3)^2}\,\mathrm{d}x$ $=\frac{1}{3\times 4}\int_{-1}^{0}\frac{3}{(2x+3)^2}\,\mathrm{d}x$ $=\frac{1}{12}\left[\frac{x}{2x+3}\right]_{-1}^{0}$ $=\frac{1}{12}\left(0-\frac{-1}{-2+3}\right)$ $=\frac{1}{12}$ **Alternative (otherwise method):**

	$\int_{-1}^{0} \frac{1}{(4x+6)^2}\,\mathrm{d}x = \frac{1}{4}\int_{-1}^{0} \frac{4}{(4x+6)^2}\,\mathrm{d}x$ $= \frac{1}{4}\left[\frac{(4x+6)^{-1}}{-1}\right]_{-1}^{0}$ $= -\frac{1}{4}\left(\frac{1}{6} - \frac{1}{2}\right)$ $= \frac{1}{12}$

<table>
<tr><td>3(i)</td><td>$y = 3^x + x^2 - \frac{1}{3}x.$

</td></tr>
<tr><td>(ii)</td><td>At $x = -3$, gradient of $C \approx -6.292644 = -6.2926$ (4 d.p.) (by GC)</td></tr>
<tr><td>(iii)</td><td>Gradient of normal to C at $x = -3$ is $\dfrac{-1}{-6.292644} \approx 0.158916$

when $x = -3$, $y_1 = 3^{-3} + (-3)^2 - \frac{1}{3}(-3) = 10.037037$



Equation of normal is $y - 10.037037 = 0.158916(x+3)$
$y = 0.158916x + 10.513785$
$y = 0.1589x + 10.5138$ (4 d.p.)</td></tr>
<tr><td>(iv)</td><td>The normal meets the curve again at
$x \approx 1.8814 = 1.88$ (3 s.f.) (by G.C.)

Area required =
$\int_{-3}^{1.8814} (0.158916x + 10.513785) - (3^x + x^2 - \frac{1}{3}x)\,\mathrm{d}x$
= 31.6 (3 s.f.) (by GC)</td></tr>
<tr><td></td><td></td></tr>
</table>

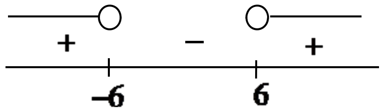
4(i)

$y = 4x^3 + px^2 + 3x + q$

$\frac{dy}{dx} = 12x^2 + 2px + 3$

For C to have two stationary points, $12x^2 + 2px + 3 = 0$ must have two real & distinct solutions. That is,

$(2p)^2 - 4(12)(3) > 0$



$p^2 - 36 > 0$

$(p-6)(p+6) > 0$

$p < -6 \quad \text{or} \quad p > 6$

Set of values of $p = \{p \in \mathbb{R} : p < -6 \text{ or } p > 6\}$

(ii) (a)

For the case in which C has a stationary point at $\left(-\frac{1}{2}, -3\right)$,

at $x = -\frac{1}{2}$, $\frac{dy}{dx} = 0$

$12\left(-\frac{1}{2}\right)^2 + 2p\left(-\frac{1}{2}\right) + 3 = 0$

$3 - p + 3 = 0$

$\Rightarrow p = 6$ (shown)

Since $\left(-\frac{1}{2}, -3\right)$ is a point on C,

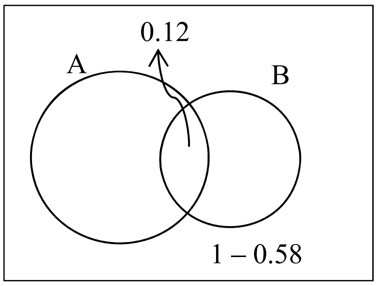
$-3 = 4\left(-\frac{1}{2}\right)^3 + 6\left(-\frac{1}{2}\right)^2 + 3\left(-\frac{1}{2}\right) + q$

$\Rightarrow q = -\frac{5}{2}$

For $p = 6$, $\frac{dy}{dx} = 12x^2 + 12x + 3$

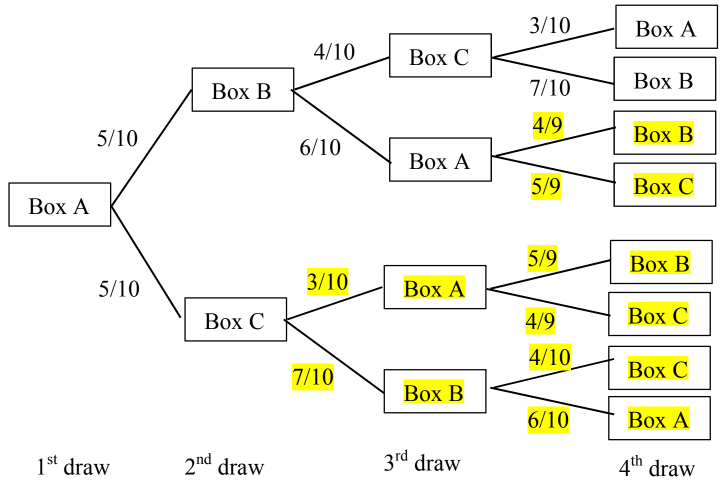
x	$\left(-\frac{1}{2}\right)^-$	$-\frac{1}{2}$	$\left(-\frac{1}{2}\right)^+$
$\frac{dy}{dx}$	+ve	0	+ve
tangent	/	–	/

Therefore, $\left(-\frac{1}{2}, -3\right)$ is a stationary point of inflexion.

(ii) (b)	From (i), the cubic curve C has two stationary points for $p < -6$ or $p > 6$. Thus we can deduce that for <u>$p = 6$</u> in part (ii)(a), <u>discriminant = 0</u> $\Rightarrow \dfrac{\mathrm{d}y}{\mathrm{d}x} = 12x^2 + 2px + 3 = 0$ has only one real solution <u>$\Rightarrow$ Since C is cubic and C has only one stationary point, which also has to be an inflexion point as well.</u> <u>Note</u>: C has no stationary point for $-6 < p < 6$. That is, C has a non-stationary point of inflexion.
5	$\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$ $0.58 = p + q - 0.12$ $p + q = 0.7 \quad \cdots(1)$ Since A and B are independent, $\mathrm{P}(A)\mathrm{P}(B) = \mathrm{P}(A \cap B)$ $pq = 0.12$ $q = \dfrac{0.12}{p} \quad \ldots(2)$ Substitute (2) into (1): $p + \dfrac{0.12}{p} = 0.7$ $p^2 + 0.12 - 0.7p = 0$ $p = 0.4 \quad$ or $\quad p = 0.3$ $q = 0.3 \qquad\quad q = 0.4$ (reject $\because p > q$)  $P(A \cup B') = 1 - (0.3 - 0.12)$ $= 0.82$

6 (i)	Advantages of stratified sampling over quota sampling - More likely to give a representative sample of the people attending the event as this method ensures an adequate sample size for each of the two sub-groups - The stratified sample is a random sample. It is unlike quota sampling which produces a biased sample by using any method of convenience to select the people for each sub-group. A stratified sample is difficult to carry out as it would be difficult to obtain the sampling frame prior to the event.
(ii)	**Remark to students:** Please note that even though sampling frame was not readily available before the event, by standing at a common place that all people will pass through (eg: exit of the venue), the surveyor would still be able to have access to the entire sampling frame for systematic sampling. To obtain a systematic sample of 1% of the population of people attending the event: • Select a random number between 1 to 100 to get a random starting point e.g. 10 • From the start to the end of the one-day event, sample every 100^{th} person at the exit of the venue, starting from the 10^{th} person, i.e. 10, 110, 210, 310,…

7(i)



(ii)

Probability

= P(drawing from box A,B,A,B)

=P(drawing R,W,R)

$$=\left(\frac{5}{10}\right)\left(\frac{6}{10}\right)\left(\frac{4}{9}\right)$$

$$=\frac{2}{15}$$

(ii)

Let event E be the event that he drew from box B on the third draw.

Let event F be the event that he has not drawn from box C from the first to his fourth draw.

$$\mathrm{P}\left(F\cap E\right)$$

$$=\frac{2}{15}$$

$$\mathrm{P}\left(E\right)$$

= P(drawing from A,B,C,B or A,B,A,B or A,C,A,B)

$$=\left(\frac{5}{10}\right)\left(\frac{4}{10}\right)\left(\frac{7}{10}\right)+\frac{2}{15}+\left(\frac{5}{10}\right)\left(\frac{3}{10}\right)\left(\frac{5}{9}\right)$$

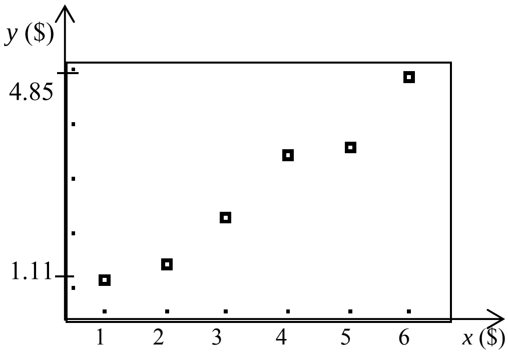
$$=\frac{107}{300}$$

$$\mathrm{P}\left(F|E\right)$$

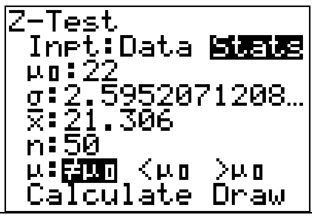
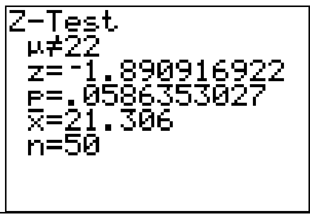
$$=\frac{\mathrm{P}(F\cap E)}{\mathrm{P}(E)}$$

$$=\frac{\frac{2}{15}}{\frac{107}{300}}$$

$$=\frac{40}{107}$$

8(i)	
(ii)	Regression line of y on x: $y = 0.13933 + 0.74971x$ $y = 0.139 + 0.750x$ (3 s.f.) Regression line of x on y: $x = -0.061412 + 1.2888y \quad \Rightarrow x = -0.0614 + 1.29y\ (3 \text{ s.f.})$ Alternatively, may express x on y as $\Rightarrow \quad y = \dfrac{x + 0.06141}{1.2888}$ $\Rightarrow \quad y = 0.77592x + 0.047651$ $\Rightarrow \quad y = 0.776x + 0.0477$
(iii)	Gradient: For every \$1000 increase in monthly household income, the predicted increase in monthly expenditure is \$775.92 (\$776). y-intercept: When there is zero household income, the predicted monthly expenditure is \$476.49 (\$476).
(iv)	$r = 0.98297\ (5 \text{ s.f.}) = 0.983$ (3 s.f.). r is close to 1, hence there is a <u>strong positive linear correlation</u> between average monthly household income and expenditure, as shown in the scatter diagram where the data points are <u>close to the regression line.</u>
(v)	Since value of x is given, we should use regression line of y on x : $x = 9.35 \Rightarrow y = 0.13933 + 0.74971(9.35) = 7.1491$ When income is \$9350, the expenditure is estimated to be <u>\$7, 150</u>.
(vi)	Since $x = 9.35$ is out of the data range of $1 \le x \le 6$ (extrapolation), the linear model may not be valid outside the data range and hence the estimate may not be reliable.

9(i)	Let X be the sleeping hours of a randomly chosen baby who drinks BabyGrow. $X \sim \mathrm{N}\left(8.0,\ \sigma^2\right)$ Let Y be the sleeping hours of a randomly chosen baby who drinks InfanGrow. $Y \sim \mathrm{N}\left(6.5,\ 0.795^2\right)$ $\mathrm{P}(X < 9) = 0.85$ $\mathrm{P}\left(Z < \frac{9-8}{\sigma}\right) = 0.85$ $\frac{9-8}{\sigma} = 1.0364$ $\sigma = 0.96488$ $= 0.965$ (shown)
(ii)	Let $T = X_1 + X_2 + X_3 \sim \mathrm{N}\left(3\times 8,\ 3\times 0.965^2\right)$ $\Rightarrow T \sim \mathrm{N}(24,\ 2.7937)$ $4Y \sim \mathrm{N}\left(4\times 6.5,\ 4^2 \times 0.795^2\right) \Rightarrow 4Y \sim \mathrm{N}(26,\ 10.112)$ $\therefore T - 4Y \sim \mathrm{N}(24-26,\ 2.7937{+}10.112)$ $\Rightarrow T - 4Y \sim \mathrm{N}(-2,\ 12.906)$ $P(T - 4Y > 1)$ $= 0.20184 = 0.202$ (3 s.f.)
(iii)	Let W be the number of babies who drank BabyGrow who slept more than 7 hours, out of 12 babies. $\mathrm{P}(X > 7) = 0.85$ (by symmetry) $W \sim B(12, P(X > 7))$ $\Rightarrow W \sim B(12, 0.85)$ $\mathrm{P}(W \geq 10)$ $= 1 - \mathrm{P}(W \leq 9)$ $= 1 - 0.26418$ $= 0.73582 = 0.736$ (3 s.f.)
(iv)	$W \sim \mathrm{B}(12, 0.85)$ E(W) = 12(0.85) = 10.2 Var (W) = 12(0.85)(1 – 0.85) = 1.53 Since n = 50 is large, by Central Limit Theorem, $\overline{W} \sim \mathrm{N}\left(10.2, \frac{1.53}{50}\right)$ approximately. $\mathrm{P}\left(\overline{W} > \alpha\right) < 0.9$ $1 - \mathrm{P}\left(\overline{W} \leq \alpha\right) < 0.9$ $\mathrm{P}\left(\overline{W} \leq \alpha\right) > 0.1$ By GC, inverse norm,

	$\alpha > 9.97582$ $\alpha > 9.98$ (3 s.f.)
10(i)	An unbiased estimate is an estimate in which the <u>expectation</u> of the estimate is equal to the <u>population parameter.</u>
(ii)	Unbiased estimate for population mean $= \dfrac{\sum(x-18)}{50} + 18$ $= 21.306$ Unbiased estimate of population variance σ^2 $= s^2 = \dfrac{1}{49}\left[\sum(x-18)^2 - \dfrac{\left(\sum(x-18)\right)^2}{50}\right]$ $= 6.7351$ $= 6.74$ (3 s.f.)
(iii)	Let X be the mass of each bag of beans from Factory A, with population mean μ kg. To test $H_0 : \mu = 22$ Against $H_1 : \mu \neq 22$ Conduct a two-tailed test at 5% level of significance $(\alpha = 0.05)$: Under H_0, since $n = 50$ is large, by Central Limit Theorem, $\overline{X} \sim N\left(22, \dfrac{6.7351}{50}\right)$ approximately Using a Z-test, p-value = 0.058635 Since p-value > 0.05, we do not reject H_0 and conclude that there is insufficient evidence at the 5% level of significance that the claim is not valid.  

(iv)	p-value is the lowest level of significance for which the null hypothesis of the mean mass of the bag of beans of 22 kg will be rejected.

Let Y be the mass of each bag of beans from Factory B, with population mean μ kg.

To test $\mathrm{H}_0 : \mu = \mu_0$

Against $\mathrm{H}_1 : \mu < \mu_0$

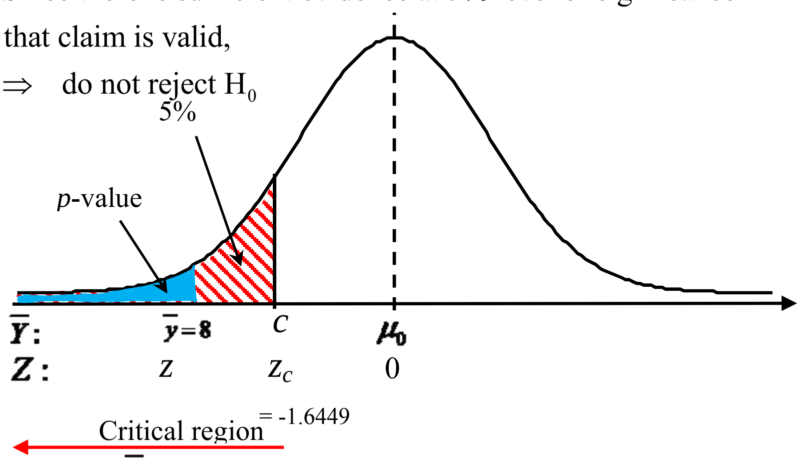
Under H_0,

$$\bar{Y} \sim \mathrm{N}\left(\mu_0, \frac{4^2}{50}\right)$$

Using test statistic, $z = \dfrac{8-\mu_0}{\sqrt{\dfrac{4^2}{50}}}$,

Since there is sufficient evidence at 5% level of signficance that claim is valid,

$\Rightarrow$ do not reject H_0



Reject H_0 : $\bar{y} < c$, $z < z_c$

Method 1:

$p - \text{value} > 0.05$

$\mathrm{P}\left(\bar{Y} < 8\right) > 0.05$

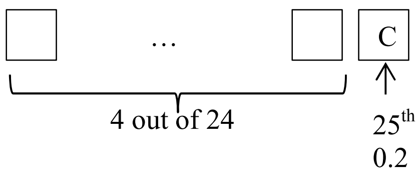
$$\Rightarrow \quad \frac{8-\mu_0}{\sqrt{\dfrac{4^2}{50}}} > -1.6449$$

$$\Rightarrow \quad 8-\mu_0 > -1.6449\sqrt{\frac{4^2}{50}}$$

$$\Rightarrow \quad \mu_0 < 8 + \frac{6.5796}{\sqrt{50}}$$

$$\Rightarrow \quad \mu_0 < 8.9305$$

$\therefore$ Set of values of $\mu_0 = \left\{\mu_0 \in \mathbb{R} : \mu_0 < 8.93\right\}$

	Method 2: $\Rightarrow$ critical region : $z > -1.6449$ $\Rightarrow \dfrac{8-\mu_0}{\sqrt{\dfrac{4^2}{50}}} > -1.6449$ $\Rightarrow \mu_0 < 8.9305$ $\therefore$ Set of values of $\mu_0 = \{\mu_0 \in \mathbb{R} : \mu_0 < 8.93\}$
11(i)	**Condition 1:** The event where a student is sick is independent of another student. Reason why the above condition may not be met: As some sickness is contagious, it may cause students who stay close to the infected student to also get infected. **Condition 2:** The probability where a student is sick is constant at 0.2. Reasons why the above condition may not be met: As each student has different immunity, the probability that each student fall sick may differ.
(ii)	Let X be the number of students who are sick, out of 24 students. X~B(24, 0.2)  Required probability = P(X = 4) (0.2) = 0.19601 (0.2) = 0.039203… = 0.0392 (3s.f.)
(iii)	Let Y be the number of students who are sick, out of n students. Y~B(n, 0.2) For sample size n, P($Y \geq$ 6) > 0.95 1 – P($Y \leq$ 5) > 0.95 P($Y \leq$ 5) < 0.05 When n = 49, 1 – P($Y \leq$ 5) = 1 – 0.05467…= 0.9453… When n = 50, 1 – P($Y \leq$ 5) = 1 – 0.04802…= 0.9519… > 0.95 When n = 51, 1 – P($Y \leq$ 5) = 1 – 0.04212…= 0.9578… Least n = 50.

(iv)	Let W be the number of students who are sick, out of 2400 students. W~B(2400, 0.2) Since n = 2400 is large, $np = (2400)(0.2) = 480 > 5$, $nq = (2400)(0.8) = 1920 > 5$, W~N(480, 384) approximately $21\% \times 2400 = 504$ students $\mathrm{P}(W < 504) \xrightarrow{\text{continuity correction}} \mathrm{P}(W < 503.5)$ $= 0.88478$ $= 0.885$ (3s.f.)
(v)	Expected number of months = (12)(6)(0.88478) = 63.704 = 63.7 (3 s.f.)