Question 1

$$y=\frac{ax+b}{\sqrt{x}}=ax^{1/2}+bx^{-1/2}$$

$$\frac{dy}{dx}=\frac{a}{2\sqrt{x}}-\frac{b}{2x^{3/2}}$$

$$\Rightarrow \text{ at } x=1, \frac{dy}{dx}=\frac{a}{2}-\frac{b}{2}$$

Given equation of normal at $x=1$ is $y=2x-2$

$\Rightarrow \frac{dy}{dx}=-\frac{1}{2}$ and $y=0$ when $x=1$.

$\therefore a-b=-1 \quad \cdots(1)$

$(1,0)$ is on the curve $\Rightarrow a+b=0 \quad \cdots(2)$

Solving (1) and (2), we have $a=-\frac{1}{2}, b=\frac{1}{2}$.

Question 2

$\int_1^n 2^{-\sqrt{x}}\,dx=2.7551$ and

$$\int_2^n \left(2^{-\sqrt{x}}-\frac{1}{2x}\right)dx=1.1727$$

$$\Rightarrow \int_1^n 2^{-\sqrt{x}}\,dx-\int_1^2 2^{-\sqrt{x}}\,dx-\int_2^n \frac{1}{2x}\,dx=1.1727$$

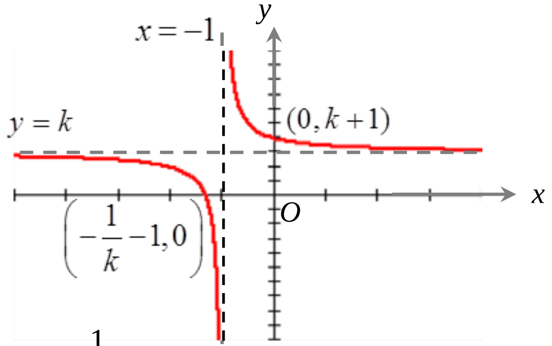
$$\Rightarrow 2.7551-0.431062-\frac{1}{2}\ln x\Big|_2^n=1.1727$$

$$\Rightarrow \ln n-\ln 2=2.302676$$

$$\Rightarrow \ln n\approx 2.9958$$

$$\Rightarrow n\approx e^{2.9958}\approx 20.001=20 \text{ (nearest integer)}$$

Question 3



$$y=k+\frac{1}{x+1}$$

$$\frac{dy}{dx}=-\frac{1}{(x+1)^2}$$

$\therefore$ equation of tangent at $x=p$ is $y-\left(k+\frac{1}{p+1}\right)=-\frac{1}{(p+1)^2}(x-p)$

$$y=-\frac{x-p}{(p+1)^2}+k+\frac{1}{p+1}$$

Tangent passes through (2, 0) $\Rightarrow$ $0=-\frac{2-p}{(p+1)^2}+k+\frac{1}{p+1}$

$$-\frac{1}{(p+1)^2}(2-p)+k+\frac{1}{p+1}=0$$

$$-(2-p)+k(p+1)^2+(p+1)=0$$

$$-2+p+kp^2+2kp+k+p+1=0$$

$$kp^2+(2+2k)p+k-1=0$$

$$\text{discriminant}=(2k+2)^2-4k(k-1)$$

$$=12k+4>0 \text{ for all } k>0$$

Hence, there are 2 distinct values of p.

This implies there will always be 2 tangents to the curve that passes through (2, 0).

Question 4

(i)

$$A_1=\int_1^{\mathrm{e}}\ln\sqrt{x}\ \mathrm{d}x;\quad A_2=\int_1^{\mathrm{e}}\left(\ln x-\ln\sqrt{x}\right)\ \mathrm{d}x$$

Since $\ln\sqrt{x}=\frac{1}{2}\ln x$, $A_1=\frac{1}{2}\int_1^{\mathrm{e}}\ln x\ \mathrm{d}x$ and

$$A_2=\int_1^{\mathrm{e}}\left(\ln x-\frac{1}{2}\ln x\right)\ \mathrm{d}x=\frac{1}{2}\int_1^{\mathrm{e}}\ln x\ \mathrm{d}x=A_1 \text{ (shown)}$$

(ii)

$$A_3=\int_0^1 x\ \mathrm{d}y$$

$$=\int_0^1 \mathrm{e}^y\ \mathrm{d}y$$

$$=\mathrm{e}^y\Big|_0^1=\mathrm{e}-1$$

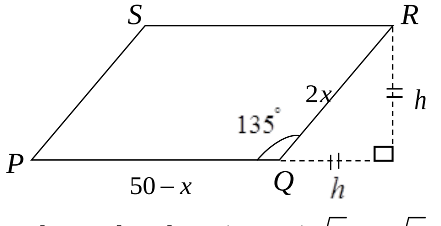
(iii)

$$A_1+A_2+A_3=\mathrm{e}\times 1$$

$$2A_1+\mathrm{e}-1=\mathrm{e}$$

$$\therefore A_1=\frac{1}{2}$$

Question 5



$$2h^2=(2x)^2$$

$$\therefore h=\sqrt{2}x$$

$A=\text{base}\times\text{height}=(50-x)\sqrt{2}x=\sqrt{2}x(50-x)$ (shown)

$A=\sqrt{2}x(50-x)=\sqrt{2}\left(50x-x^2\right)$ $\frac{\mathrm{d}A}{\mathrm{d}x}=\sqrt{2}(50-2x)=0\Rightarrow x=25$

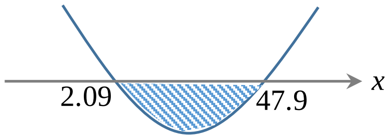
Sign test on $\frac{\mathrm{d}A}{\mathrm{d}x}$:

x	25^-	25	25^+
Sign of $\frac{\mathrm{d}A}{\mathrm{d}x}$	+	0	-
Slope	/	—	\

When $x = 25, A = 625\sqrt{2}$ is a maximum value.

$$A > 100\sqrt{2} \Rightarrow 50x - x^2 > 100$$
$$x^2 - 50x + 100 < 0$$

From GC, $x^2 - 50x + 100 = 0$ when $x = 2.09$ or 47.9



$2.09 < x < 47.9$

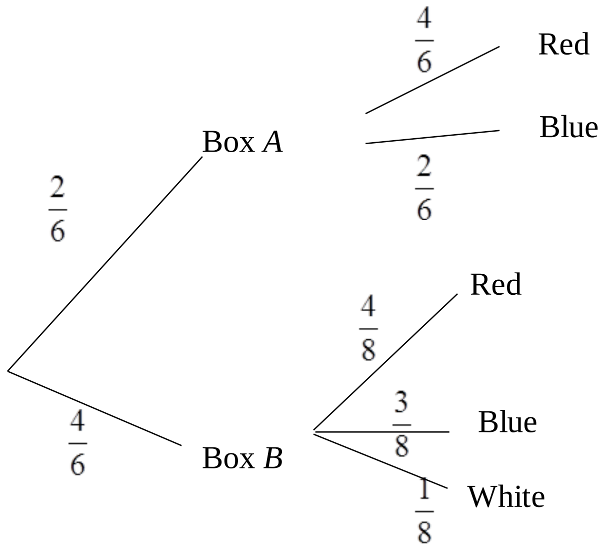
The largest value of x is 47.

Section B

Question 6
(i) Obtain a list of all the shops in each category in the shopping mall from the directory. Use a random sampling to select from each category a number which is proportional to the number of shops in the category. For example, if there are 60 shops in the fashion category, select 10 fashion shops. A stratified sample of 10 fashion shops can be obtained by using a random number generator to obtain 10 distinct numbers and then select the 10 shops which correspond to the numbers generated. This procedure is repeated for the remaining 3 categories.
(ii) It is difficult to obtain the sampling frame i.e. the number of shoppers in the shopping mall, thus, it would be difficult to use a stratified sampling.
(iii) Quota sampling. The manager would not be able to obtain a random sample as the manager might select shoppers based on his preference. Hence not everyone has an equal chance of being selected.

Question 7

(i) (a)



(ii) $\text{P(box A} \mid \text{blue)} = \dfrac{\text{P(box A} \cap \text{blue)}}{\text{P(blue)}}$

$= \dfrac{\left(\frac{2}{6}\right)\left(\frac{2}{6}\right)}{\left(\frac{2}{6}\right)\left(\frac{2}{6}\right)+\left(\frac{4}{6}\right)\left(\frac{3}{8}\right)} = \dfrac{\left(\frac{2}{6}\right)\left(\frac{2}{6}\right)}{\frac{13}{36}} = \dfrac{4}{13}$

Required probability $= \left(\frac{2}{6}\right)\left(\frac{4}{6}\right)\left(\frac{3}{5}\right)\left(\frac{2}{4}\right)+\left(\frac{4}{6}\right)\left(\frac{4}{8}\right)\left(\frac{3}{7}\right)\left(\frac{2}{6}\right)+\left(\frac{4}{6}\right)\left(\frac{3}{8}\right)\left(\frac{2}{7}\right)\left(\frac{1}{6}\right) = \dfrac{53}{420}$

Question 8

Let X be the number of spoiled eggs in a carton. $X \sim \text{B}(10, p)$.

$\text{P}(X = 0) = 0.48$

$\binom{10}{0} p^0 (1-p)^{10} = 0.48$

$(1-p)^{10} = 0.48$

$p = 0.07076807$

$\text{P}(X = 1)$

$= \binom{10}{1}(0.07076807)^1(1-0.07076807)^9$

≈ 0.365556

$= 0.3656$ (4 d.p)

$X \sim \text{B}(10,\ 0.07076807)$

$2\text{E}(X) = 2(10 \times 0.07076807) = 1.41536$

$P(X \geq 1.41536)$

$= P(X \geq 2)$

$= 1 - P(X \leq 1)$

$= 0.1544435$

Required prob = $(0.1544435)^3 = 0.0036839 = 0.00368(3 \text{ s.f.})$

Let Y be the number of cartons, out of 30, with no spoiled eggs.
$Y \sim B(30, 0.48)$
$np = 14.4 > 5$; $n(1-p) = 15.6$; $np(1-p) = 7.488$
Since $np > 5$ and $n(1-p) > 5$, $Y \sim N(14.4, 7.488)$ approximately.

$P(20 < Y \leq 25)$

$= P(20.5 < Y < 25.5)$

$= 0.0128757$

$= 0.0129$ (3 s.f)

Question 9

Let $w = x - 3$, then , $\sum w = 45$, $\sum w^2 = 425$ and

Unbiased estimate of μ is $\bar{x} = 3 + \bar{w} = 3 + \frac{45}{60} = 3.75$

Unbiased estimate of σ^2 is $s_x^2 = s_w^2 = \frac{1}{n-1}\left[\sum x^2 - \frac{\left(\sum x\right)^2}{n}\right]$

$= \frac{1}{59}\left[425 - \frac{(45)^2}{60}\right] = 6.631356 \approx 6.63$

$H_0 : \mu = 3$

$H_1 : \mu > 3$

Level of significance:5%

Test Statistic: Since n = 60 is sufficiently large, so s_x^2 is a good estimate of σ^2 and by Central Limit Theorem, $\bar{X}$ is approximately normal.

$\therefore \bar{X} \sim N\left(3, \frac{s_x^2}{n}\right)$ approximately when H_0 is true.

$\therefore Z = \frac{\bar{X} - 3}{s_x / \sqrt{n}} \sim N(0,1)$.

Rejection region: $z \geq 1.6449$

Computation: $\bar{x} = 3.75$, $n = 60$, $s_x = \sqrt{6.631356}$

$\therefore z = 2.25598 \approx 2.26$

$p - \text{value} = 0.0120358 \approx 0.0120$

Conclusion: Since p – value = 0.0120 < 0.05, $\therefore H_0$ is rejected at 5% significance level. Hence there is sufficient evidence to conclude that the machine is not working correctly at the 5% significance level.

Yes. The test is valid since $n = 60$ is sufficiently large, by Central Limit Theorem, the sample mean length of a nail ($\bar{X}$) is approximately normally distributed.

If $\sigma = 0.1$, then when H_0 is true, $\bar{X} \sim N\left(3, \frac{\sigma^2}{n}\right)$

P (presuming machine has gone wrong when in fact it is working correctly) = 0.01

$\therefore P\left(\bar{X} > a \text{ when } H_0 \text{ is true}\right) = 0.01$

$\Rightarrow P\left(Z > \frac{a-3}{0.1/\sqrt{n}}\right) = 0.01$

From GC: $P\left(Z > 2.32635\right) = 0.01$

$\therefore \frac{a-3}{0.1/\sqrt{n}} = 2.32635$

$\Rightarrow a = 3 + 2.32635\left(\frac{0.1}{\sqrt{n}}\right)$

$\approx 3 + \frac{0.233}{\sqrt{n}}$

Question 10

Let B kg be the mass of a randomly chosen Butternut pumpkin.

$B \sim N(\mu, \left(\frac{\mu}{8}\right)^2)$

$P\left(B \geq 0.9\mu\right)$

$= P\left(Z \geq \frac{0.9\mu - \mu}{\mu/8}\right)$

$= P\left(Z \geq -0.8\right)$

$= 0.7881447$

$= 0.788$ (3 s.f.)

78.8% of the Butternut pumpkins have mass at least 0.9 of the mean mass.

$B \sim N(0.8, 0.1^2)$

$\bar{B} = \frac{B_1 + B_2 + ... + B_5}{5}$

$\bar{B} \sim N\left(0.8, \frac{0.1^2}{5}\right)$

$P\left(\bar{B} \leq 0.9\right) = 0.987326 = 0.987$ (3 s.f.)

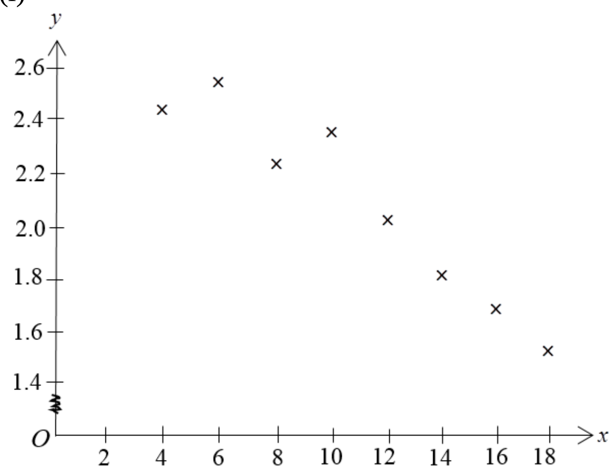
Let J kg be the mass of a randomly chosen Japanese pumpkin.
$J \sim N(1, 0.15^2)$
Cost of one Butternut pumpkin, $X = 2.5B$
$E(X) = 2.5(0.8) = 2$ $\text{Var}(X) = 2.5^2(0.1^2) = 0.0625$
$X \sim N(2, 0.0625)$

Cost of one Japanese pumpkin, $Y = 1.67J$
$E(Y) = 1.67(1) = 1.67$
$\text{Var}(Y) = 1.67^2(0.15^2) = 0.06275025$
$Y \sim N(1.67, 0.06275025)$
$C = X_1 + X_2 + Y_1 + Y_2 + Y_3 \sim N\left(9.01, 0.31325075\right)$
$P\left(8.50 < C < 9.50\right) = 0.628$ (3 s.f.)

Assume that the masses of all the pumpkins are independent of one other.

Question 11

(i)



(ii) $r = -0.960$ (to 3 s.f.)

Since $r = -0.960$ is close to -1 and the points seem to lie close to a straight line with negative gradient are indications of a strong negative linear relationship between the charge (x) and the average number of vehicles entering the city centre per day (y). This means that as x increases, y tends to decrease at a constant rate.

(iii) Since the values of x are fixed (or controlled), hence x is an independent variable. So the least squares regression lines, y on x should be used.

(iv) The equation of the regression line of y on x is $y = 2.8226 - 0.070238x$ i.e $y = 2.82 - 0.0702x$ (to 3 s.f.)

(v) When there is no congestion charge i.e. $x = 0$, so the average number of vehicles which will enter the city centre per day is 2820000 (or 2.82 million).

Since $x = 0$ is out of the range of the data $4 \le x \le 18$, $\therefore$ 2.82 million does not necessarily give the expected average number of vehicles entering the city centre per day.

(vi) $x = 2 + 2w$

$\therefore r = -0.960$ (same as the product moment correlation coefficient between x and y found in **(i)**) as the product moment correlation coefficient is unaffected by change of scale and location.