Normal Distribution and Sampling

1(a)	By symmetry, $\mu = \dfrac{155+185}{2} = 170$ $\mathrm{P}\left(X < 155\right) = 0.025$ $\Rightarrow \mathrm{P}\left(Z < \dfrac{155-170}{\sigma}\right) = 0.025$ $\Rightarrow -\dfrac{15}{\sigma} = -1.95996$ $\Rightarrow \sigma = 7.6532$ $\therefore \sigma^2 = 58.6$ (3 s.f.)
(i)	$X_1 + X_2 - 2X_3 \sim N\left(0, 6\sigma^2\right)$ $\mathrm{E}\left(X_1 + X_2 - 2X_3\right) = \mathrm{E}\left(X\right) + \mathrm{E}\left(X\right) - 2\mathrm{E}\left(X\right) = 0$ $\mathrm{Var}\left(\mathrm{X}_1 + X_2 - 2X_3\right) = 2\mathrm{Var}\left(X\right) + 4\mathrm{Var}\left(X\right) = 6\sigma^2$ $\mathrm{P}\left(X_1 + X_2 > 2X_3 + 5\right)$ $= \mathrm{P}\left(X_1 + X_2 - 2X_3 > 5\right)$ $= 0.39484$ = 0.394 (3 s.f.)
(ii)	$\overline{X} \sim N\left(\mu, \dfrac{\sigma^2}{50}\right)$ (Since X is normally distributed, there isn't a need to use CLT.) $\mathrm{P}\left(\overline{X} < 172\right)$ $= 0.96769$ = 0.968 (3 s.f.)
(b)	Since n is large, by CLT $\overline{X} \sim N\left(\mu, \dfrac{5^2}{n}\right)$ approx. $\mathrm{P}\left(\left\lvert \overline{X} - \mu \right\rvert < 1\right) > 0.99$ $\Rightarrow \mathrm{P}\left(\lvert Z \rvert < \dfrac{1}{5/\sqrt{n}}\right) > 0.99$ $\Rightarrow \mathrm{P}\left(-\dfrac{1}{5/\sqrt{n}} < Z < \dfrac{1}{5/\sqrt{n}}\right) > 0.99$ $\Rightarrow -0.2\sqrt{n} < -2.5758$ $\Rightarrow 0.2\sqrt{n} > 2.5758$ $\Rightarrow n > 165.87$ Least n = 166

2(a)	Let A be the random variable that denotes the weight, in grams, of a bar of Brand A chocolate $A \sim N(180,100)$ Let B be the random variable that denotes the weight, in grams, of a bar of Brand B chocolate $B \sim N(240,400)$ $A_1 + A_2 + A_3 - 2B \sim N(60,1900)$ $P(A_1 + A_2 + A_3 > 2B) = P(A_1 + A_2 + A_3 - 2B > 0)$ $P(A_1 + A_2 + A_3 - 2B > 0) = 0.916 (\text{to } 3 \text{ s.f.})$
(b)(i)	$\overline{A} \sim N\left(180, \frac{100}{60}\right)$ $P(\overline{A} > 179) = 0.781 (\text{to } 3 \text{ s.f.})$
(ii)	No need to use Central Limit Theorem because the population follows a Normal Distribution
(c)(i)	$0.02A \sim N\left((0.02)(180), (0.02)^2 (100)\right)$ $0.02A \sim N(3.6, 0.04)$ $P(0.02A > 3.80) = 0.1586552596 = 0.159 (\text{to } 3 \text{ s.f.})$

3(i)	$Y = aX + b \sim \text{N} (30a + b , 16a^2)$ $\text{P}(Y < 4) = \text{P}(Y > 16) = 0.06681$ By symmetry(since the two end tailed areas are equal), $30a + b = \frac{4+16}{2} = 10 - - - - (1)$ $\frac{4 - 30a - b}{4a} = -1.5000 \Rightarrow 24a + b - 4 = 0....(2)$ From (1), $b = 10 - 30a$ subt. into (2) $24a + 10 - 30a - 4 = 0$ $a = 1$ and $b = -20$ (Shown) **Alternative solution :** $\text{P} (Z < \frac{4 - 30a - b}{4a}) = 0.06681$ $\frac{4 - 30a - b}{4a} = -1.5000 \Rightarrow 24a + b - 4 = 0....(1)$ $\text{P}(Y > 16) = 0.06681 \Rightarrow \text{P}(Y < 16) = 0.93319$ $\text{P} (Z < \frac{16 - 30a - b}{4a}) = 0.93319$ $\frac{16 - 30a - b}{4a} = 1.5000 \Rightarrow 36a + b - 16 = 0....(2)$

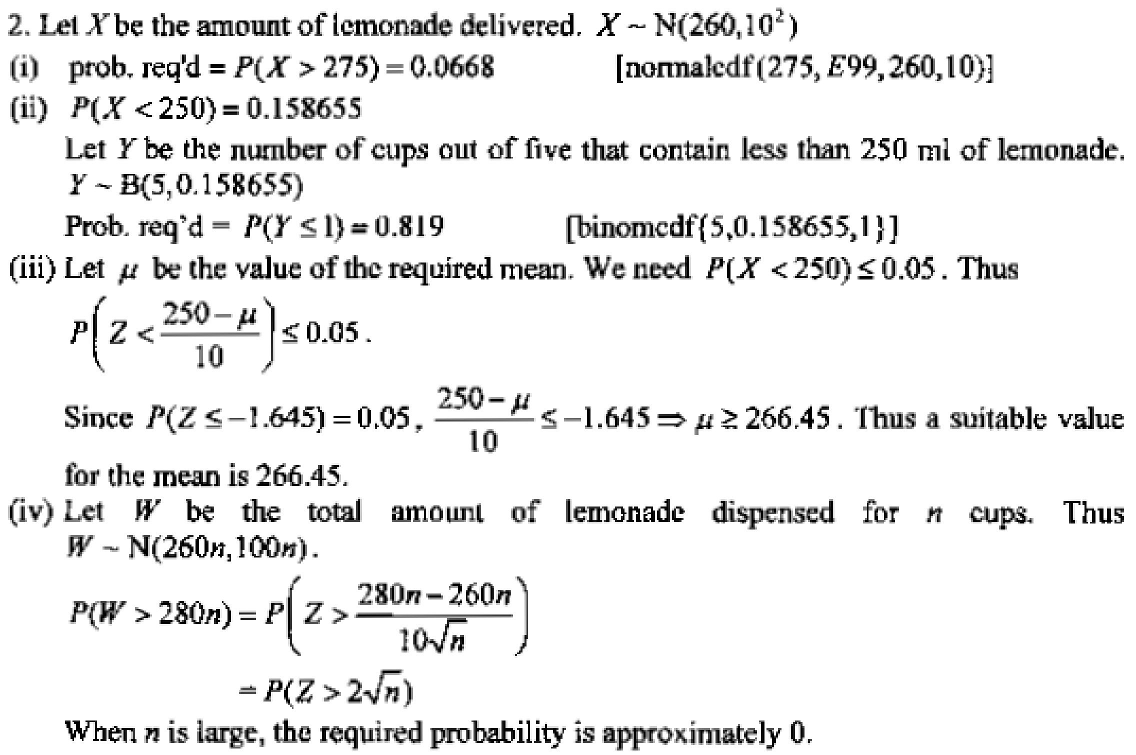
(ii)	$Y = X - 20 \sim \text{N}(10, 16) \Rightarrow \overline{Y} \sim N(10, \frac{16}{n})$ $\text{P}(\overline{Y} < 8) = 0.05$ $\text{P}(z < \frac{8-10}{4/\sqrt{n}}) = 0.05$ $\frac{8-10}{4/\sqrt{n}} = -1.645$ $n \approx 10.8$ $n = 11$

4(i)	Let X = duration of calls to City *A*. Let Y = duration of calls to City *B*. $X \sim N(8, 2.5^2)$ and $Y \sim N(10, 3.2^2)$ $X_1 + X_2 - 3X_3 \sim N\left(2(8) - 3(8), 2(2.5)\right)^2 + 3^2(2.5)^2)$ $\Rightarrow X_1 + X_2 - 3X_3 \sim N(-8, 68.75)$ $P(\lvert X_1 + X_2 - 3X_3 \rvert > 3)$ $= P(X_1 + X_2 - 3X_3 > 3) + P(X_1 + X_2 - 3X_3 < -3)$ = 0.0923122 + 0.726753 = 0.819
(ii)	Total cost $C = 22(X_1 + X_2) + 30Y$ +30 $E(C) = E(22(X_1 + X_2) + 30Y + 30)$ $= E(22(2)(8) + 30(10) + 30)$ = 682 Expected amount of money = 682 cents or \$6.82
(iii)	$Var(C) = Var(22(X_1 + X_2) + 30Y + 30)$ $= Var\left(22^2(2)(2.5)^2 + 30^2(3.2)^2\right)$ = 15266 $C \sim N(682, 15266)$ $P(C > 850) = 0.0869606 = 0.0870$

5(i)(a)	Let X be the random variable for the length of a pickle $X \sim \text{N}(9, 0.8^2)$ P(X < 8) = 0.105650 = 0.106 (to 3sf)
(b)	P(X >10.5) = 0.0303963 = 0.0304 (to 3sf)
(ii)	Since 10.6% of the pickles are already discarded because they are too short, it is impossible to reduce the total rejects to 5%

(iii)	P(both pickles to meet his guaranteed standard $\mid X_1 + X_2 < 16$ cm) $= \frac{\text{P(both pickles fail his g'teed standard \& } X_1 + X_2 < 16)}{\text{P}(X_1 + X_2 < 16)}$ $= \frac{(\text{P}(X < 8))^2}{\text{P}(X_1 + X_2 < 16)}$ …….(***) $= \frac{(0.1056498)^2}{0.03854989} = 0.290$

6(i)	Let X denotes the height of a girl $X \sim N(155,100)$ $\text{P}(155 - a < X < 155 + a) = 0.4$ $\text{P}(X < 155 - a) = 0.3$ $\text{P}(\text{Z} < -\frac{a}{10}) = 0.3$ $-\frac{a}{10} = -0.5244 \Rightarrow a = 5.24$ range of girls' height is (149.76,160.24)
(ii)	$\text{P}(X > 170) = \text{P}(\text{Z} > \frac{170 - 155}{10}) = 0.066807$ $\text{required probabilty} = (0.066807)^2 (1 - 0.066807)^2 \frac{4!}{2!2!}$ $= 0.0233$
(iii)	$X_T \sim N(155n, 100n)$ $\text{P}(X_T > 150n) = \text{P}(\text{Z} > \frac{150n - 155n}{\sqrt{100n}})$ $= \text{P}(\text{Z} > -\frac{1}{2}\sqrt{n})$ $= 1$ (since n is large)

7	 2. Let X be the amount of lemonade delivered. $X \sim \mathrm{N}(260,10^2)$ (i) prob. req'd = $P(X > 275) = 0.0668$ [normalcdf(275, E99, 260, 10)] (ii) $P(X < 250) = 0.158655$ Let Y be the number of cups out of five that contain less than 250 ml of lemonade. $Y \sim \mathrm{B}(5, 0.158655)$ Prob. req'd = $P(Y \le 1) = 0.819$ [binomcdf{5,0.158655,1}] (iii) Let μ be the value of the required mean. We need $P(X < 250) \le 0.05$. Thus $P\left(Z < \frac{250-\mu}{10}\right) \le 0.05$. Since $P(Z \le -1.645) = 0.05$, $\frac{250-\mu}{10} \le -1.645 \Rightarrow \mu \ge 266.45$. Thus a suitable value for the mean is 266.45. (iv) Let W be the total amount of lemonade dispensed for n cups. Thus $W \sim \mathrm{N}(260n, 100n)$. $P(W > 280n) = P\left(Z > \frac{280n - 260n}{10\sqrt{n}}\right)$ $= P(Z > 2\sqrt{n})$ When n is large, the required probability is approximately 0.

8(i)	Let C and W be the random variables denoting the mass of a roll of Cleanex and WoolSoft toilet paper in grams, respectively. $C \sim \mathrm{N}\left(200, 10^2\right)$ $W \sim \mathrm{N}\left(220, 15^2\right)$ $\mathrm{P}\left(202 < C < 208\right) = 0.20888 = 0.209$
(ii)	Let $A = C_1 + C_2 + \ldots + C_{10}$ and $B = W_1 + W_2 + \ldots + W_{10}$ $B - 1.05A \sim \mathrm{N}\left(10(220) - 1.05(10)(200), 10\left(15^2\right) + 1.05^2(10)\left(10^2\right)\right)$ $\sim \mathrm{N}\left(100, 3352.5\right)$ $\mathrm{P}\left(B > 1.05A\right) = \mathrm{P}\left(B - 1.05A > 0\right)$ $= 0.95792$ $= 0.958$
(iii)	$A + B_1 + B_2 + B_3 \sim N\left(10(200) + 3(10)(220),\ 10(10)^2 + 3(10)(15)^2\right)$ $\sim N\left(8600,\ 7750\right)$ $\mathrm{P}\left(A + B_1 + B_2 + B_3 < k\right) = 0.4$ $k = 8577.7$ $= 8580$ (to 3 s.f.)
	Let U be the random variable denoting the mass of a roll of 4-ply toilet paper in grams, respectively.

	$U \sim \mathrm{N}\left(\mu, \sigma^2\right)$ $\mathrm{P}(U < 220) = 0.04$ $\mathrm{P}\left(Z < \frac{220-\mu}{\sigma}\right) = 0.04$ $\frac{220-\mu}{\sigma} = -1.7507$ $\mu - 1.7507\sigma = 220$ ------(1) $\mathrm{P}(U > 230) = 0.80$ $\mathrm{P}\left(Z > \frac{230-\mu}{\sigma}\right) = 0.80$ $\frac{230-\mu}{\sigma} = -0.84162$ $\mu - 0.84162\sigma = 230$ ------(2) $\mu = 239.26 = 239$ $\sigma = 11.000 = 11.0$

9(i)	8 let X -- time taken for giving order Y - time taken to fulfill an order Then $X \sim N(60, 20^2)$, $Y \sim N(40, 10^2)$. i) $P(Y>60) = 0.0228$ ※ [normalcdf(60, e^{99}, 40, 10)]
(ii)	ii) $\bar{Y} \sim N(40, \frac{10^2}{8})$ $P(35 < \bar{Y} < 55) = 0.921$ ※ [normalcdf (35, 55, 40, $\frac{10}{\sqrt{8}}$)] iii) $Y_1 + Y_2 + Y_3 + X_1 + X_3 = T \sim N(240, 1100)$ ∴ $P(T < 180) = 0.0352$ ※ [normalcdf (-e^{99}, 180, 240, $\sqrt{1100}$)]

10(i)	Let X be the waiting time on a randomly chosen day. X~ N(8,5) Let Y be the journey time on a randomly chosen day. Y~N(11,4) Let T be the total time taken on a randomly chosen day. T=X+Y~N(19,9) P(T>20) = 0.369 (to 3 sf)
(ii)	Expected no. of days late in a month = 30 x 0.369 = 11.07 ≈ 11 days

(iii)	Let c be the time taken. $P(T > c) < 0.05$ $P(T < c) > 0.95$ $c > 23.9346$ Least value of c = 24 min Thus, the latest time he can leave his house = 7:40am – 24 min = 7:16am
(iv)	Let $\overline{T}$ be the average time taken in a month. $\overline{T} \sim N(19, 9/30)$ $P(15<\overline{T}<20) = 0.966$ (to 3sf)

11(a)	No. The distribution would be asymmetric or skewed. If it can be modeled by Normal distribution, then there will be approximately the same number of employees earning above and below the mean salary
(b)(i)	$T_A \sim N(55, 25)$, $T_B \sim N(53, 16)$ Let $\overline{T} \sim N(54, 10.25)$ where $\overline{T} = \dfrac{T_A + T_B}{2}$ $P(50 <\overline{T} <60) = 0.864$ (3 sig figures).
(ii)	$T_B \sim N(53, 16)$ $P(53 - a \le T_B \le 53 + a) \le 0.6$ $\Rightarrow P(T_B \le 53 + a) \le 0.8$ $\Rightarrow 53 + a \le 56.366$ (from GC) $\Rightarrow a \le 3.366$ Hence greatest value of a is 3.36

12(i)	Let r.v. A be the mass of a snapper fish and r.v. B be the mass of a pomfret fish. $A \sim N(1, 0.1^2)$; $B \sim N(0.6, 0.05^2)$ (a)(i) $A_1 + A_2 + A_3 + B_1 + B_2 \sim N(4.2, 0.035)$ $P[A_1 + A_2 + A_3 + B_1 + B_2 > 4.5] = 0.0544$
(ii)	$A_1 + A_2 + A_3 - 2B \sim N(1.8, 0.04)$ $P[A_1 + A_2 + A_3 - 2B > 1.85] = 0.401$
(iii)	$12A + 7(B_1 + B_2) \sim N(20.4, 1.685)$ $P[12A + 7(B_1 + B_2) > 21] = 0.322$ $12(A_1 + A_2 + \ldots + A_n) + 7(B_1 + B_2 + \ldots + B_{15-n}) \sim N(63 + 7.8n, 1.8375 + 1.3175n)$ $P[12(A_1 + A_2 + \ldots + A_n) + 7(B_1 + B_2 + \ldots + B_{15-n}) > 150] < 0.7$. Largest $n = 11$

13(i)	Let *A* be the mass of an avocado. $A \sim N\left(115, 9^2\right)$ $\mathrm{P}\left(110 < X < 115\right) = 0.21074$ Required prob= $3\times\left(0.21074\right)\left(0.5\right)^2$ = 0.158
(ii)	Let *K* be the mass of a kiwi. $K \sim N\left(82, \sigma^2\right)$ $\mathrm{P}\left(X > 90\right) = 0.1055$ $\mathrm{P}\left(Z > \frac{90-82}{\sigma}\right) = 0.1055$ $\frac{8}{\sigma} = 12508$ $\sigma = 6.40$
(iii)	$\frac{K_1+K_2+A_1+A_2+A_3}{5} \sim N\left(\frac{1}{5}\left(509\right), \frac{1}{5^2}\left(324.92\right)\right)$ $\frac{K_1+K_2+A_1+A_2+A_3}{5} \sim N\left(101.8, 12.9968\right)$ $P\left(\frac{K_1+K_2+A_1+A_2+A_3}{5} > 100\right) = 0.691$
(iv)	Let $T = 0.012\left(K_1+K_2\right)+0.015\left(A_1+A_2+A_3\right)$ $T \sim N\left(0.012\left(164\right)+0.015\left(345\right), 0.012^2\left(6.40^2\right)+0.015^2\left(9^2\right)\right)$ $\therefore T \sim N\left(7.143, 0.024123\right)$ $P\left(T \le a\right) \ge 0.99$ $a \ge 7.5403$ lease value of *a* = 8

14(i)	$X \sim N\left(10.2, 1.2^2\right)$ $\mathrm{P}\left(11 < Y < 15\right) = \mathrm{P}\left(11 < \frac{1}{2}\left(30 - X\right) < 15\right)$ $= \mathrm{P}\left(0 < X < 8\right)$ = 0.03376 = 0.0334(to 3 sf) Alternative Method:

	$Y \sim N\left(\frac{1}{2}(30-10.2),\left(\frac{1}{2}\right)^2 1.2^2\right) \Rightarrow Y \sim N(9.9,0.36)$ $\mathrm{P}(11<Y<15)=0.03376=0.0334$ (to 3 sf)
(ii)	P(letter will fit into envelope) $=\mathrm{P}(X<11 \text{ and } Y<11)$ $=\mathrm{P}\left(X<11 \text{ and } \frac{1}{2}(30-X)<11\right)$ $=\mathrm{P}(8<X<11)$ $=0.714$
(iii)	$X-Y=X-\frac{1}{2}(30-X)=\frac{3}{2}X-15$ $\mathrm{Var}(X-Y)=\mathrm{Var}\left(\frac{3}{2}X-15\right)=\left(\frac{3}{2}\right)^2\mathrm{Var}(X)=\frac{9}{4}(1.2)^2=3.24$ $\mathrm{Var}(X)+\mathrm{Var}(Y)=1.2^2+\mathrm{Var}\left(15-\frac{1}{2}X\right)==1.2^2+\left(\frac{1}{2}\right)^2\mathrm{Var}(X)=1.8$ Therefore $\mathrm{Var}(X-Y)\neq\mathrm{Var}(X)+\mathrm{Var}(Y)$ The rule does not hold since *X* and *Y* are not independent.

15(i)	Let *X* denote the waiting time (in minutes) for SuperX Ride Let *Y* denote the waiting time (in minutes) for SuperY Ride $X \sim \mathrm{N}(50,10^2)$ $Y \sim \mathrm{N}(45,12^2)$ $\mathrm{E}(X_1+X_2-2Y)=2(50)-2(45)=10$ $\mathrm{Var}(X_1+X_2-2Y)=2(10^2)+4(12^2)=776$ $\therefore X_1+X_2-2Y \sim \mathrm{N}(10,776)$ P (sum of waiting times of Jay and Kay is less than twice the waiting time of Candy) = $\mathrm{P}(X_1+X_2<2Y)=\mathrm{P}(X_1+X_2-2Y<0)=0.360$
(ii)	Let *W* denotes the number of tourists, out of 20, who purchase souvenir photos. $W \sim \mathrm{B}(20,0.4)$ $\mathrm{P}(W \geq k)>0.5$ $\mathrm{P}(W \leq k-1)<0.5$ $\mathrm{P}(W \leq 7)=0.41589<0.5$ $\mathrm{P}(W \leq 8)=0.59560>0.5$ Largest value of $k-1=7$ $k=8$

(iii)	Let M denote the mean amount spent by 60 tourists. Since n is large, by Central Limit Theorem, $M \sim \mathrm{N}\left(35, \frac{9.4^2}{60}\right)$ approximately $\mathrm{P}(M > 36.50) = 0.108$

16(i)	Let X and Y be random variables denoting the waiting times, in minutes, for one random order of Chili Crab and one random order of Pepper Crab respectively. Then $X \sim \mathrm{N}(\mu , 3^2)$ and $Y \sim \mathrm{N}(10 , 2^2)$ $Y - X \sim \mathrm{N}(10 - \mu,\ 3^2 + 2^2)$ $Y - X \sim \mathrm{N}(10 - \mu,\ 13)$ $\mathrm{P}(Y - X > 0) = 0.290$ $\Rightarrow \quad \mathrm{P}\left(Z > \frac{0 - (10 - \mu)}{\sqrt{13}} \right) = 0.290$ $\Rightarrow \quad \frac{\mu - 10}{\sqrt{13}} = 0.5533847152$ $\Rightarrow \quad \mu = 11.99525697 \approx 12.0$ (to 3 s.f.) **(shown)**
(ii)	Carpark charges \$0.05 per minute Let T be the random variable denoting the carpark charges of one random customer in cents. $T \sim \mathrm{N}(5 \times 10 + 5 \times 20 , 5^2 \times 2^2 + 5^2 \times 4^2)$ $T \sim \mathrm{N}(150, 500)$ P(T > \$2) = P(T > 200 cents) = 0.0126736174 ≈ 0.0127
(iii)	Since $Y \sim \mathrm{N}(10 , 2^2)$ So, $\bar{Y} \sim \mathrm{N}\left(10, \frac{2^2}{n} \right)$ Given $\mathrm{P}(\bar{Y} > 10.5) \leq 0.0385$ **Method 1: Using G.C. table** n $\mathrm{P}(\bar{Y} > 10.5)$ 50 0.03855 51 0.0371 52 0.03571 Hence, least value of n is 51. **Method 2: Using standardization** $\mathrm{P}(\bar{Y} > 10.5) \leq 0.0385$ $\Rightarrow \quad \mathrm{P}\left(Z > \frac{10.5 - 10}{\frac{2}{\sqrt{n}}} \right) \leq 0.0385$

	$\Rightarrow \qquad \frac{0.5\sqrt{n}}{2} \geq 1.768364425$ $\Rightarrow \qquad n \geq 50.03380381$ Hence, least value of n is 51. No assumption needed since Y is normally distributed

17(a)	Let T s be the time taken to complete a race. I think 5 s is the wrong standard deviation. Suppose T~N(10, 5^2). If T~N(10, 5^2), then 99.7% of the values should lie within $10 \pm 3(5)$, i.e. $(-5, 25)$, which contains a significant range of negative values. Then P($T < 0$) ≈ 0.0228 which is significantly large. However T is a non-negative quantity
(b)(i)	Now let T~N(10, 2^2). P($T < t$) = 0.99 ⇒ $t \approx 14.7$
(ii)	Let Y be the number of races out of 20 races that are less than t seconds. $Y \sim B(20, 0.99)$ Required probability $= P(Y \leq 18) = 0.0169$
(iii)	Let $W = \frac{T_1 + T_2 + ... + T_5}{5} - 2T$. W ~ N(-10, 16.8) $\text{P}(\frac{T_1 + T_2 + ... + T_5}{5} > 2T) = \text{P}(W > 0) \approx 0.00735$
18(i)	Let X be rv representing the time in minutes taken by a boy in completing the obstacle course. $\therefore \; X \sim N(33, 5^2)$ Given $\text{P}(\lvert X - 33 \rvert < k) = 0.80$ $\text{P}\left(\lvert Z \rvert < \frac{k}{5}\right) = 0.80$ $\text{P}\left(Z < \frac{k}{5}\right) = 0.90$ or $\text{P}\left(Z < -\frac{k}{5}\right) = 0.10$ Using GC, $\frac{k}{5} = 1.28155$ $\Rightarrow \; k = 6.40776 = 6.41$ (3 s.f.)
(ii)	Let Y be the rv representing time in minutes taken by a girl in completing the obstacle course $\therefore \; Y \sim N(48, 7^2)$ $Y_1 + Y_2 - 3X \sim N(96 - 99, \; 2(49) + 3^2(25))$

	$Y_1+Y_2-3X \sim N(-3,\ 323)$ $\mathrm{P}\left(Y_1+Y_2>3X\right)=\mathrm{P}\left(Y_1+Y_2-3X>0\right)=0.43371=0.434$

19(i)	$\mathrm{P}\left(X>\mu-a\right)+\mathrm{P}\left(X>\mu+a\right)+\mathrm{P}\left(\mu<X<\mu+2a\right)=1.38$ $1+\mathrm{P}\left(\mu<X<\mu+2a\right)=1.38$ $\mathrm{P}\left(\mu<X<\mu+2a\right)=0.38$ $\mathrm{P}\left(X>\mu+2a\right)=0.5-0.38$ $\mathrm{P}\left(X>\mu+2a\right)=0.12$
(ii)(a)	$\mathrm{P}\left(\lvert X-\mu\rvert\le L\right)=0.4$ $\mathrm{P}\left(\left\lvert\frac{X-\mu}{\sigma}\right\rvert\le\frac{L}{\sigma}\right)=0.4$ $\mathrm{P}\left(-\frac{L}{\sigma}\le\frac{X-\mu}{\sigma}\le\frac{L}{\sigma}\right)=0.4$ $\frac{L}{\sigma}=0.5244005=0.524$ (to 3 s.f)
(ii)(b)	$\mathrm{P}\left(X-\mu\ge 2L\right)=\ \mathrm{P}\left(\frac{X-\mu}{\sigma}\ge\frac{2L}{\sigma}\right)$ $=\ \mathrm{P}\left(Z>\frac{2L}{\sigma}\right)=\mathrm{P}\left(Z>2(0.5244005)\right)$ $=\mathrm{P}\left(Z>1.048801\right)$ $=0.14713488=\ 0.147$ (to 3 s.f)

20	P(Wise misses the bus) = P(X – Y < 0) $X-Y \sim N\left(25-15,9+4\right)$ i.e. $X-Y \sim N\left(10,13\right)$ Thus, probability required = 0.0027728 ≈ 0.00277 (3 s.f.) T = 45 –（W + X） E(T) = 45 – 30 – 25 = – 10 Var(T) = 3 + 9 = 12 P(the bus arriving after 8.30 a.m.) = P(T < 0) = 0.99805 ≈ 0.998 (3 s.f.)

21(i)	$X \sim \mathrm{N}\left(\mu,\sigma^2\right)$

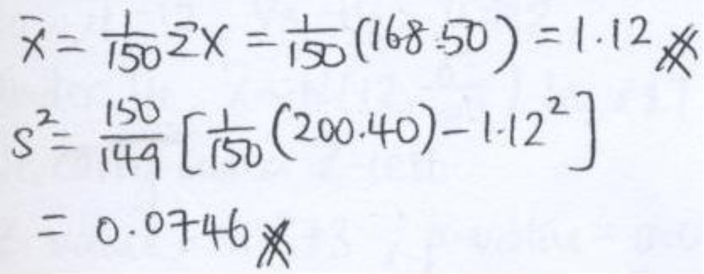
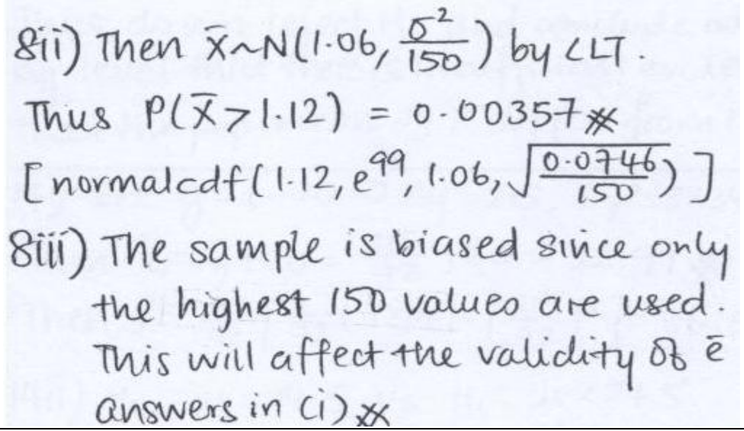
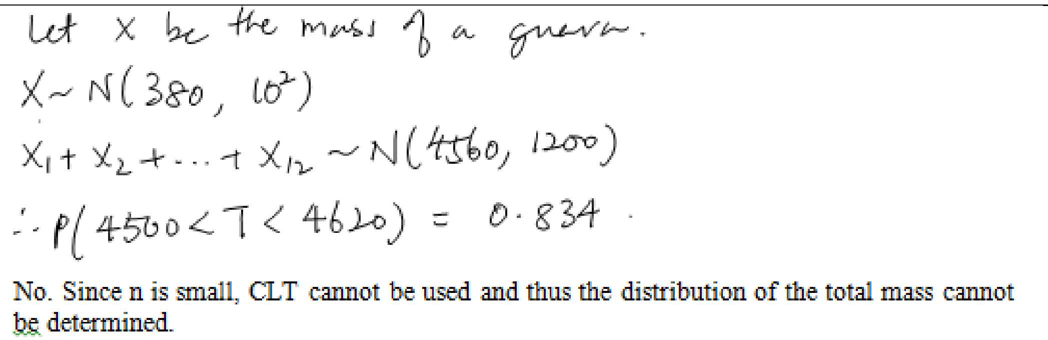
	$\mathrm{P}\left(X > 2a\right) = 0.10$ $\mathrm{P}\left(X < 2a\right) = 0.90$ $\mathrm{P}\left(Z < \frac{2a-\mu}{\sigma}\right) = 0.90$ $\frac{2a-\mu}{\sigma} = 1.28155$ $2a-\mu = 1.28155\sigma$ ----------- (1) $\mathrm{P}\left(X < a\right) = 0.30$ $\mathrm{P}\left(Z < \frac{a-\mu}{\sigma}\right) = 0.30$ $\frac{a-\mu}{\sigma} = -0.52440$ $a-\mu = -0.52440\sigma$ --------- (2) $(1)-(2):\quad a = 1.80595\sigma$ $\sigma = \frac{a}{1.80595} = 0.553725a$ $a-\mu = -0.52440 \times \frac{a}{1.80595}$ $= -0.29037a$ $\mu = 1.29a$ $\mathrm{E}(X) = \mu = 1.29a$ $\mathrm{Var}(X) = \sigma^2 = 0.553725^2 a^2 = 0.30661a^2 = 0.307a^2$
(ii)	$\mathrm{E}(X_1 + X_2 - 2X_3) = \mathrm{E}(X) + \mathrm{E}(X) - 2\mathrm{E}(X) = 0$ $\mathrm{Var}(X_1 + X_2 - 2X_3) = \mathrm{Var}(X) + \mathrm{Var}(X) + 4\mathrm{Var}(X)$ $= 6\mathrm{Var}(X)$ $= 1.83966526a^2$ $X_1 + X_2 - 2X_3 \sim \mathrm{N}(0, 1.83966526a^2)$ $\mathrm{P}\left(X_1 + X_2 - 2X_3 > a\right) = \mathrm{P}\left(Z > \frac{a-0}{\sqrt{1.83966526a^2}}\right)$ $= \mathrm{P}\left(Z > 0.737276848\right)$ ≈ 0.230

22(i)	Let X = number of customers who pay by credit card in sample of 4 Let Y = number of customers who pay by credit card in sample of 15

	$X \sim B(4,0.3)$, $Y \sim B(15,0.3)$ Required probability $= P(X=1)P(Y>4)\times 0.3$ $= P(X=1)\left[1-P(Y\le 4)\right]\times 0.3$ = 0.05982 = 0.0598
(ii)	Let $\bar{Y}$ = average number of customers who pay by credit card $Y \sim B(15,0.3)$ $E(Y)=15(0.3)=4.5$ and $Var(Y)=15(0.3)(0.7)=3.15$ By CLT, $\bar{Y} \sim N\left(4.5,\dfrac{3.15}{50}\right)$ $P\left(4<\bar{Y}<a\right)=0.95$ $\Rightarrow P\left(\bar{Y}<a\right)-P\left(\bar{Y}<4\right)=0.95$ $\Rightarrow P\left(\bar{Y}<a\right)=0.95+P\left(\bar{Y}<4\right)$ $\Rightarrow P\left(\bar{Y}<a\right)=0.97318$ $\Rightarrow a=4.98$

23(i)	n = 50, large using Central Limit theorem $\bar{X} \sim \text{N}(20\ 000, \dfrac{2200^2}{50})$ P($\left\|\bar{X}-20\ 000\right\| < 800$) = P(19200≤ $\bar{X}$ ≤20 800) = 0.990
(ii)	Let Y be the random variable for number of samples with sample mean within \$800 of the population mean $Y \sim \text{B}(5, 0.989868)$ P(Y= 3) = 9.96×10^{-4} Assume : P(success) = 0.989868 is a constant for each trial

24	(i) $\bar{x} = 11.3$ $s^2 = \dfrac{\Sigma(x-\bar{x})^2}{n-1} = \dfrac{14400}{99} = 145.45$ (ii) Since $n=100$ is large, $\bar{X} \sim N\left(12, \dfrac{145.45}{100}\right)$ by CLT $\therefore P(\bar{X} > 11.3) = 0.719.$

25(i)	 $\bar{x} = \frac{1}{150}\Sigma x = \frac{1}{150}(168.50) = 1.12$ ※ $s^2 = \frac{150}{149}\left[\frac{1}{150}(200.40) - 1.12^2\right]$ $= 0.0746$ ※  8ii) Then $\bar{X} \sim N(1.06, \frac{\sigma^2}{150})$ by CLT. Thus $P(\bar{X} > 1.12) = 0.00357$ ※ [normalcdf(1.12, e99, 1.06, $\sqrt{\frac{0.0746}{150}}$)] 8iii) The sample is biased since only the highest 150 values are used. This will affect the validity of ē answers in (i) ※
26	 Let X be the mass of a guava. $X \sim N(380, 10^2)$ $X_1 + X_2 + \ldots + X_{12} \sim N(4560, 1200)$ $\therefore P(4500 < T < 4620) = 0.834$. No. Since n is small, CLT cannot be used and thus the distribution of the total mass cannot be determined.

27(i)	If the amount spent by the person, X, is normally distributed, i.e. if X~N(6.5, 3.76^2), then 99.7% of the values should lie within $6.5 \pm 3(3.76)$, ie, $(-4.78, 17.78)$ which contains a significant range of negative values. OR There may be some very large purchases leading to an asymmetric distribution.
(ii)	Let X be the amount spent by a person. Since n = 250 is large, by CLT, $X_1 + \ldots + X_{250} \sim N(1625, 3534.4)$ $P(X_1 + \ldots + X_{250} > 1658) = 0.289$
28(i)	Unbiased estimate of the mean μ = $\frac{4000}{50} = 80$ Unbiased estimate of the variance = $s^2 = \frac{1}{49}\left(360250 - \frac{(4000)^2}{50}\right) = \frac{5750}{7} = 821\frac{3}{7}$

(ii)	Since n = 50 is large, by Central Limit Theorem, $\bar{X} \sim N\left(80, \frac{115}{7}\right)$ approximately
(iii)	$P\left(\bar{X} > 85\right)$ = 0.109 (3 sig. fig.)
(iv)	Let n be the sample size. $\bar{X} \sim N\left(80, \frac{5750}{7n}\right)$ approximately $P(\bar{X} > 85) < 0.03$ $P\left(Z > \frac{5}{\sqrt{\frac{5750}{7n}}}\right) < 0.03$ $5\sqrt{\frac{7n}{5750}} > 1.8808$ $\sqrt{n} > \frac{1.8808}{5}\sqrt{\frac{5750}{7}} = 10.781$ $n > 116.2$ Hence least number of sample size is 117.

29	(i) Let L be the amount of drink in a large cup and S be the amount of drink in a small cup. $L \sim \mathrm{N}(405, 74)$ $S \sim \mathrm{N}(202, 21)$ $L_1 + L_2 - 4S \sim \mathrm{N}(2, 484)$ $\mathrm{P}(L_1 + L_2 - 4S < 0) = 0.464$ (ii) He can buy 3 small cups of drink OR 1 large cup and 1 small cup of drink. 3 small cups of drink : $S_1 + S_2 + S_2 \sim \mathrm{N}(606, 63)$ $\mathrm{P}(S_1 + S_2 + S_2 > 600) = 0.775$ 1 large cup and 1 small cup of drink : $L + S \sim \mathrm{N}(607, 95)$ $\mathrm{P}(L + S > 600) = 0.764$ He should buy 3 small cups. (Note: the amount he needs to spend is the same for both cases) (iii)

Let $T = \dfrac{S_1 + S_2 + \ldots + S_{20} + L_1 + L_2 + \ldots + L_n}{20+n}$

$$T \sim N\left(\frac{20(202)+n(405)}{20+n}, \frac{20(21)+n(74)}{(20+n)^2}\right)$$

$$T \sim N\left(\frac{4040+405n}{20+n}, \frac{420+74n}{(20+n)^2}\right)$$

$\text{P}(T > 350) > 0.8$

Using GC, when n = 54, $\text{P}(T > 350) = 0.5598 < 0.8$

when n = 55, $\text{P}(T > 350) = 0.834 > 0.8$

Least value of n = 55

OR

$$\text{P}\left(Z > \frac{350 - \frac{4040+405n}{20+n}}{\sqrt{\frac{420+74n}{(20+n)^2}}}\right) > 0.8$$

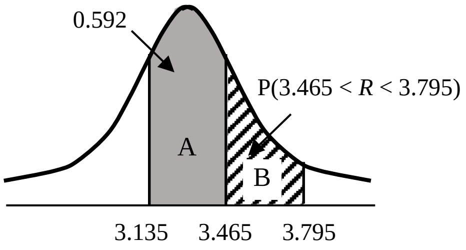
$$\frac{350 - \frac{4040+405n}{20+n}}{\sqrt{\frac{420+74n}{(20+n)^2}}} < -0.8416212$$

Using GC, Least value of n = 55

30

Let R be the radius of a randomly chosen tennis ball. $R \sim \text{N}\left(3.3,\ 0.2^2\right)$

(i) $\text{P}\left(3.135 < R < 3.465\right) = 0.59063 = 0.591$ (3 s.f.)



From diagram, $\text{P}\left(3.135 < R < 3.465\right) > \text{P}\left(3.465 < R < 3.795\right)$

since area A is larger than area B given that the widths of the two intervals $3.135 < R < 3.465$ and $3.465 < R < 3.795$ are the same.

(ii) Required probability = $\left[\mathrm{P}\left(R<3.4\right)\right]^3=0.331$ (3 sf)

(iii) Let *H* be the height of a randomly chosen cylindrical tube. $H\sim\mathrm{N}\left(20,\ 0.3^2\right)$

Gap, $G=H-2\ (R_1+R_2+R_3)$

$G\sim\mathrm{N}(\ 20-2(3.3\times3),\ 0.3^2+2^2(0.2^2\times3)\ )$

i.e. $G\sim\mathrm{N}(\ 0.2,\ \sqrt{0.57}^2\ )$

$\mathrm{P}(G\geq k)\ \geq\ 0.15$

Using GC, $0<k\leq0.982$

31(i)	Let random variable *A* be the BFE of Brand *A* mask. Since $\mathrm{P}(A<95.7)=\mathrm{P}(A>95.78)$, $\mu=\frac{95.7+95.78}{2}=95.74$ $\mathrm{P}(A<95.7)=0.0912$ $\mathrm{P}(Z>\frac{95.7-95.74}{\sigma})=0.0912$ $\frac{-0.04}{\sigma}=-1.333401746$ $\sigma=0.0299984608=0.03$ (2 d.p)
31(ii)	Let random variable *B* be the BFE of Brand *B* mask. $B\sim\mathrm{N}\left(92.19,0.03^2\right)$ $A\sim\mathrm{N}\left(91.09,0.08^2\right)$ $\overline{A}\sim\mathrm{N}\left(91.09,\frac{0.08^2}{n}\right)$ $B-\overline{A}\sim\mathrm{N}\left(1.1,0.03^2+\frac{0.08^2}{n}\right)$

$\mathrm{P}(|B-\overline{A}| \le 1.15) \ge 0.9405$

$\mathrm{P}(-1.15 \le B-\overline{A} \le 1.15) \ge 0.9405$

$$\mathrm{P}\left(\frac{-1.15-1.1}{\sqrt{0.03^2+\frac{0.08^2}{n}}} \le Z \le \frac{1.15-1.1}{\sqrt{0.03^2+\frac{0.08^2}{n}}} \right) \ge 0.9405$$

Using GC,

n	$\mathrm{P}(\lvert B-\overline{A}\rvert \le 1.15)$
49	0.9403
50	**0.9406**
51	0.9408

$\therefore$ Least $n = 50$.

31(iii)

$X \sim \mathrm{N}\left(203, \sigma_1{}^2\right) \qquad Y \sim \mathrm{N}\left(203, \sigma_2{}^2\right)$

$3X - \left[Y+Y_2+Y_3\right] \sim \mathrm{N}\left(0,\ 9\sigma_1{}^2 + 3\sigma_2{}^2\right)$

$$\mathrm{P}\left(3X > Y+Y_2+Y_3\right) = \mathrm{P}\left(3X - \left[Y+Y_2+Y_3\right] > 0\right)$$
$$= 0.5$$

32:

(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 $X \sim \mathrm{N}(\mu , \sigma^2)$,

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

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

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

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

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

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

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

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

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.

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

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

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

$\quad \sim \mathrm{N}(-69.99\ ,\ 258.56)$

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

$\qquad = 0.54974(5\ \text{S.F.})$

(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.

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

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

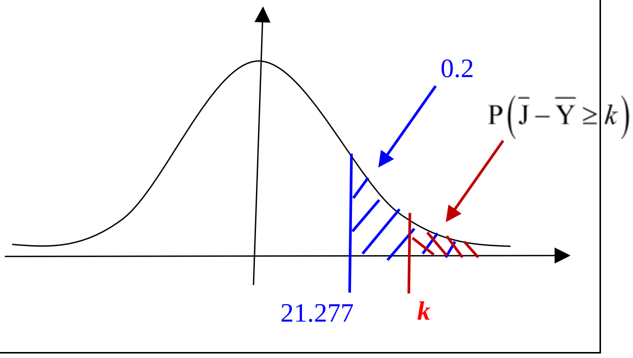
$\overline{J} \sim \mathrm{N}\left(220, \dfrac{30.25}{24}\right)$ and $\overline{Y} \sim \mathrm{N}\left(200, \dfrac{5^2}{24}\right)$

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

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

$\qquad k \geq 21.277\ \ (5\ \text{s.f.})$

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



33 (i)	Let F be the random variable denoting the mass of a packet of fettucine in grams. $F \sim \mathrm{N}\left(450\ ,\ \sigma^2\right)$ Given that

	$\mathrm{P}\left(F \le 437\right) = 0.1$ $\mathrm{P}\left(\frac{F-450}{\sigma} \le \frac{437-450}{\sigma}\right) = 0.1$ $\mathrm{P}\left(Z \le \frac{-13}{\sigma}\right) = 0.1$, where $Z \sim \mathrm{N}(0,1)$ $-\frac{13}{\sigma} = -1.2816$ $\sigma = 10.144 = 10.1$ (3 sig fig)
(ii)	Let S be the random variable denoting the mass of a packet of spaghetti produced in g. $S \sim \mathrm{N}(500\,, 9.2^2)$ $F \sim \mathrm{N}\left(450\,, 10.144^2\right)$ Required probability $=\mathrm{P}\left(\left\lvert S_1+S_2+S_3+S_4+S_5+S_6-2\left(F_1+F_2\right)\right\rvert \ge 1250\right)$ Let $A = S_1+S_2+S_3+S_4+S_5+S_6-2\left(F_1+F_2\right)$ $\mathrm{E}(A)$ $= \mathrm{E}\left[S_1+S_2+S_3+S_4+S_5+S_6-2\left(F_1+F_2\right)\right]$ $= \mathrm{E}\left(S_1+S_2+S_3+S_4+S_5+S_6\right)-2\mathrm{E}\left(F_1+F_2\right)$ $= 6\mathrm{E}\left(S\right)-2(2)\mathrm{E}\left(F\right)$ $= 6\left(500\right)-4(450)$ $=1200$ $\mathrm{Var}(A)$ $= \mathrm{Var}\left[S_1+S_2+S_3+S_4+S_5+S_6-2\left(F_1+F_2\right)\right]$ $= \mathrm{Var}\left(S_1+S_2+S_3+S_4+S_5+S_6\right)+4\mathrm{Var}\left(F_1+F_2\right)$ $= 6\mathrm{Var}\left(S\right)+4(2)\mathrm{Var}\left(F\right)$ $= 6\left(9.2^2\right)+8(10.144^2)$ $=1331.045$ $A \sim \mathrm{N}(1200\ , 1331.045)$

$$\mathrm{P}\left(|A| \geq 1250\right)$$
$$= 1 - \mathrm{P}\left(-1250 < A < 1250\right)$$
$$= 0.085268$$
$$= 0.0853 \text{ (to 3 sf)}$$

(iii)

The mass of each packet of pasta (for both spaghetti and fettuccine) must be independent of each other.
[Reject if did not consider intermingling or within the pasta group]

(iv)

P(495 < S < 505) = 0.41320

Let X be the random variable denoting the number of packets of spaghetti out of n purchased that has a mass between 495 g and 505 g each.

X ~ B(n , 0.41320)

Given $\mathrm{P}(X > 7) \geq 0.85 \Rightarrow \mathrm{P}(X \geq 8) \geq 0.85$

$$1 - \mathrm{P}(X \leq 7) \geq 0.85$$
$$\mathrm{P}(X \leq 7) \leq 0.15$$

Using GC

n	$\mathrm{P}(X \leq 7)$
24	0.1582>0.15
25	0.1241<0.15
26	0.0963<0.15

Hence, the least value of n is 25.

(v)

From (iv), X ~ B(25 , 0.41320)

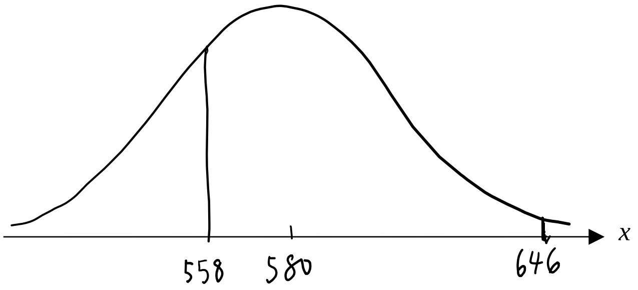
$$\mathrm{E}\left(X\right) = np$$
$$= 25(0.41320)$$
$$= 10.33$$
$$\mathrm{Var}(X) = np(1-p)$$
$$= 25(0.41320)(1-0.41320)$$
$$= 6.0616$$

Since the sample size = 24 is sufficiently large, by Central Limit Theorem,

$\overline{X} \sim N\left(10.33\,,\, \dfrac{6.0616}{24}\right)$ approximately

$$\mathrm{P}(\overline{X} \leq 11) = 0.90876 = 0.909 \text{ (to 3 sf)}$$

<table>
<tr><td>34(i)</td><td>Let X be the number of standard size packets (out of 10 in a family pack) that contains a winning coupon. $X \sim \mathrm{B}(10,\ p)$
Most likely number of winning coupons in a family pack is 2
$\Rightarrow$ mode of $X = 2$
Thus $\mathrm{P}(X=2) > \mathrm{P}(X=1)$ and $\mathrm{P}(X=2) > \mathrm{P}(X=3)$
$\binom{10}{2} p^2 (1-p)^8 > \binom{10}{1} p (1-p)^9$ and $\binom{10}{2} p^2 (1-p)^8 > \binom{10}{3} p^3 (1-p)^7$
$45p > 10(1-p)$ and $45(1-p) > 120p$
$p > \frac{2}{9}(1-p)$ and $(1-p) > \frac{8}{3}p$
$p > \frac{2}{11}$ and $p < \frac{3}{11}$
Thus $\frac{2}{11} < p < \frac{3}{11}$</td></tr>
<tr><td>(ii)</td><td>$X \sim \mathrm{B}(10, 0.2)$
P(a family pack has at least 1 winning coupon) $= \mathrm{P}(X \geq 1) = 1 - \mathrm{P}(X = 0) = 0.89263$ (5 s.f.)
Let Y be the number of family packs (out of N packs) with at least 1 winning coupon.
$Y \sim \mathrm{B}(N, 0.89263)$
Given: $\mathrm{P}(Y \geq 30) > 0.99$
$\Rightarrow$ $1 - \mathrm{P}(Y \leq 29) > 0.99$
Using GC,
<table>
<tr><th>N</th><th>$1 - \mathrm{P}(Y \leq 29)$</th></tr>
<tr><td>38</td><td>0.9829 < 0.99</td></tr>
<tr><td>39</td><td>0.9931 > 0.99</td></tr>
</table>
Least $N = 39$</td></tr>
<tr><td>(iii)(a)</td><td>Let W be the number of standard size packets (out of $12 \times 10 = 120$ packets in a carton) that contains a winning coupon.
$W \sim \mathrm{B}(120, 0.2)$
P(a carton contains exactly 28 winning coupons)
$= \mathrm{P}(W = 24) = 0.0907$ (3 s.f.)</td></tr>
<tr><td>(iii)(b)</td><td>P(every family pack in a carton contains exactly 2 winning coupons each)
$= [\mathrm{P}(X = 2)]^{12} = 5.75 \times 10^{-7}$ (3 s.f.)</td></tr>
<tr><td>(iii)(c)</td><td>The answer for (b) is smaller than that for (a) because the case in (b) is only one of the many cases for (a).
In addition to the case in (b), (a) includes many other cases such as 4 family packs containing 10 winning coupons, 1 family pack containing 8 winning coupons and the remaining family packs do not contain any winning coupons.</td></tr>
</table>

Qn	**Suggested Solution**
35(a)	
(b)	$X \sim \mathrm{N}(580, 22^2)$ Expected number $= 300 \times \mathrm{P}(X > 600)$ $= 300 \times 0.18165$ $= 54.495$ $= 54.5$ (3 s.f.)
(c)	No. By combining the masses, it would give a distribution with 2 peaks instead of a single peak.
(d)	Let K and L be the selling price of a randomly chosen rock melon and watermelon respectively. $K = 0.003X, \quad L = 0.0028Y$ $K \sim \mathrm{N}(0.003 \times 580,\ 0.003^2 \times 22^2)$ $K \sim \mathrm{N}(1.74, 0.004356) \Rightarrow \bar{K} \sim \mathrm{N}\left(1.74, \frac{0.004356}{4}\right)$ $L \sim \mathrm{N}(0.0028 \times 870,\ 0.0028^2 \times 30^2)$ $L \sim \mathrm{N}(2.436, 0.007056)$ $\bar{K} - L \sim \mathrm{N}(-0.696, 0.008145)$ $\mathrm{P}(\left\|\bar{K} - L\right\| \le 0.60)$ $= \mathrm{P}(-0.60 \le \bar{K} - L \le 0.60)$ $= 0.14373$ $= 0.144$ (3s.f.)

(e)	$K_1+...+K_n$ ~ $\text{N}(1.74n,\ 0.004356\text{n})$ $L_1+...+L_{20-n}$ ~ $\text{N}(2.436(20-n),\ 0.007056(20-n))$ Let W be the total cost of the 20 melons. $W=K_1+...+K_n+L_1+...+L_{20-n}$ W ~ $\text{N}(1.74n+2.436(20-n),\ 0.004356n+0.007056(20-n))$ $\text{P}(W>38)>0.95$ Using GC table, $n=13,\quad \text{P}(W>38)=1>0.95$ $n=14,\quad \text{P}(W>38)=0.9988>0.95$ $n=15,\quad \text{P}(W>38)=0.8113<0.95$ Greatest n = 14