(1)

S1 June '01 - Solutions

1 a) $\mu = \frac{\Sigma x}{n} = \frac{1075}{25} = \underline{43}$

$\sigma^2 = \frac{\Sigma x^2}{n} - \mu^2 = \frac{44625}{25} - 43^2 =$

b) There is no effect as they are equally far from the mean, 8 above and 8 below.

---

2. a) $S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{n} = 6599600 - \frac{7300^2}{10}$

$= \underline{1270600}$

b) $r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$ $\quad S_{xy} = -13060$ $\quad S_{yy} = 140.9$

$\therefore r = \frac{-13060}{\sqrt{1270600 \times 140.9}} = \underline{-0.976}$ (3sf)

c) A strong negative correlation; the higher above sea level the town is the lower it's temperature.

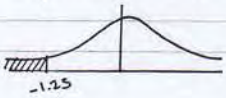
---

3. a) $Z = \frac{Y-100}{\sqrt{256}}$ $\quad Y \sim N(100, 16^2)$ $\quad Z \sim N(0, 1^2)$

$P(Y<80) = P\left(Z < \frac{80-100}{16}\right) = P(Z < -1.25)$

$= \Phi(-1.25) = 1 - \Phi(1.25)$

$= \underline{0.1056}$



(3)

$E(X) = -2\times0.1 + -1\times0.2 + 0\times0.3 + 1\times0.2 + 2\times0.1 + 3\times0.1$

$= -0.2 - 0.2 + 0.2 + 0.2 + 0.3 = 0.3$

$E(3X+4) = 3\times0.3 + 4 = \underline{4.9}$

e) $\text{Var}(2X+3) = 2^2 \text{Var}(X)$ $\quad \text{Var}(X) = E(X^2) - (E(X))^2$

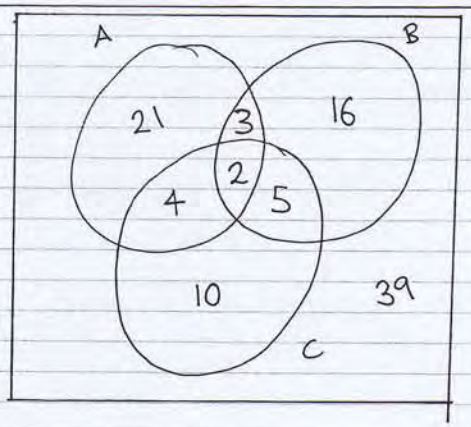
$\text{Var}(X) = \left[(-2)^2\times0.1 + (-1)^2\times0.2 + 0^2\times0.3 + 1^2\times0.2 + 2^2\times0.1 + 3^2\times0.1\right] - 0.3^2$

$= (0.4 + 0.2 + 0.2 + 0.4 + 0.9) - 0.09$

$= 2.01$

$\therefore \text{Var}(2X+3) = 4\times2.01 = \underline{8.04}$

5a)

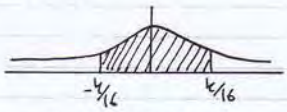


(2)

b) $P(100-k \le Y \le 100+k) = 0.516$

$P\left(\frac{100-k-100}{16} \le Z \le \frac{100+k-100}{16}\right) = 0.516$

$P\left(-\frac{k}{16} \le Z \le \frac{k}{16}\right) = 0.516$



$\Phi\left(\frac{k}{16}\right) - \Phi\left(-\frac{k}{16}\right) = 0.516$

$\Phi\left(\frac{k}{16}\right) - \left(1 - \Phi\left(\frac{k}{16}\right)\right) = 0.516$

$2\Phi\left(\frac{k}{16}\right) - 1 = 0.516$

$\Phi\left(\frac{k}{16}\right) = 0.758$

$\frac{k}{16} = 0.7$

$\underline{k = 11.2}$

---

4 a) $\alpha = \underline{0.2}$

b) $P(-1 < X \le 2) = P(X = 0, 1, 2) = \underline{0.6}$

c) $F(-0.4) = P(X = -1, -2) = \underline{0.3}$

d) $E(3X+4) = 3E(X) + 4$

(4)

b) $\frac{61}{100}$ c) $\frac{21}{100}$ d) $\frac{47}{100}$ e) $\frac{21}{47}$

---

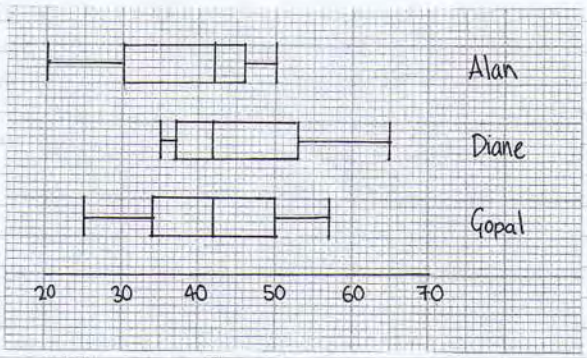
6. a) $n = 46$

Q_1: $\frac{46}{4} = 11.5 \Rightarrow 12^{th}$ value $= \underline{30}$

Q_2: $\frac{46}{2} = 23 \Rightarrow \frac{23^{rd} + 24^{th}}{2} = \frac{41+43}{2} = \underline{42}$

Q_3: $\frac{46}{4}\times3 = 34.5 \Rightarrow 35^{th}$ value $= \underline{46}$

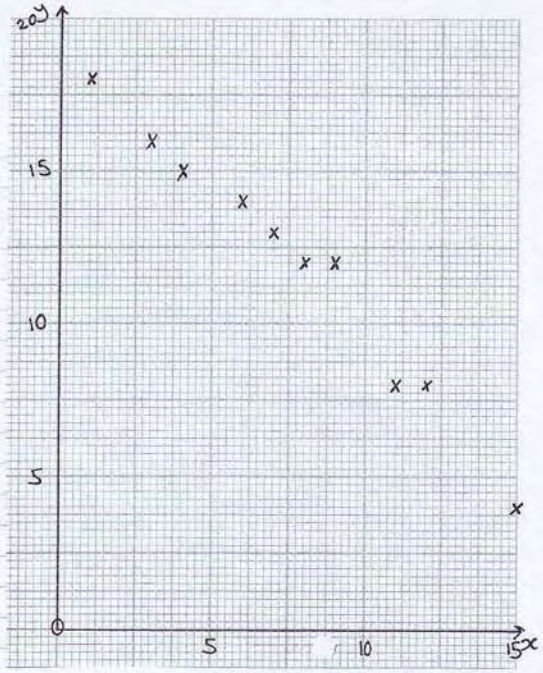
b)



c)
- All three have the same median.
- All three have the same IQR.
- A shows negative skew, D positive skew and G shows no skew.
- A and D share the same range, G has the biggest range.

(5)

7 a)



b) $y = a + bx \quad a = \bar{y} - b\bar{x} \quad b = \frac{S_{xy}}{S_{xx}}$

$$S_{xy} = \Sigma xy - \frac{\Sigma x \Sigma y}{n}$$

$$= 749 - \frac{76 \times 120}{10} = -163$$

$$S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{n} = 746 - \frac{76^2}{10} = 168.4$$

$$b = \frac{-163}{168.4} = -0.968 \text{ (3sf)}$$

(6)

$$a = \frac{\Sigma y}{n} - b\frac{\Sigma x}{n} = \frac{120}{10} + 0.968 \times \frac{76}{10}$$

$$= 19.4 \text{ (3sf)}$$

$\therefore$ $y = 19.4 - 0.968x$

c)
- The more hours you study, the less errors you make.
- If you don't study you would expect to make around 19 errors.

d) I) It is reasonable because there are a good spread of x values.

II) It is not reasonable. For all x values you need to have a sample that covers all values of x.