S1 January 2018 (IAL) (MA)

Q1a) $\frac{\sum x}{15} = 61$

$\frac{t}{15} = 61 \quad \therefore t = 61 \times 15 = \boxed{915}$

b) $Var(A) = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2 = \frac{59610}{10} - (77)^2$

$= \boxed{32}$

$Var(B) = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2 = \frac{58035}{15} - (61)^2$

$= \boxed{148}$

c) $\boxed{B}$ since $Var(B) > Var(A)$
(so larger spread of data)

di) $\mu = \frac{n_A \bar{x}_A + n_B \bar{x}_B}{n_A + n_B} = \frac{770 + 915}{25} = \boxed{67.4}$

ii) $\sigma^2 = \frac{\sum x^2}{n} - (\bar{x})^2 = \frac{(59610) + (58035)}{25} - (67.4)^2$

$= \boxed{163}$

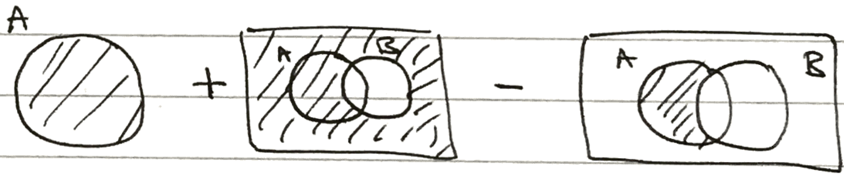
ei) No effect - addition / subtraction to all values won't affect variance.

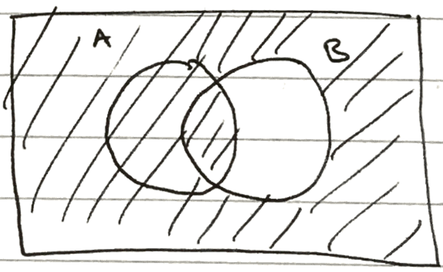
$$\left[\text{ie } Var(X+a) = Var(X) \right]$$

ii) mean will increase since Σx has increased for all students.

iii) Total variance increases - mean of A was originally 16 more than the mean of B. This will now increase meaning greater overall variance since mean of A is further away from mean of B.

Q2a) $P(A \cup B') = P(A) + P(B') - P(A \cap B')$

$=$ 

$=$ 

bi) $P(C \cup D) = P(C) + P(D) - P(C \cap D)$ ← 0 (mutually exclusive)

$$= \frac{1}{5} + \frac{3}{10} - 0 = \boxed{\frac{1}{2}}$$

ii) $P(C|D) = \dfrac{P(C \cap D)}{P(D)} = \dfrac{0}{\frac{3}{10}} = \boxed{0}$

c i) $P(G)$.

$P(F \cup G) = P(F) + P(G) - P(F \cap G)$

$\dfrac{3}{8} = \dfrac{1}{6} + P(G) - P(F) \times P(G)$

$P(G)\left[1 - \dfrac{1}{6}\right] = \dfrac{5}{24}$

$\therefore P(G) = \dfrac{\frac{5}{24}}{\frac{5}{6}} = \boxed{\dfrac{1}{4}}$

ii) $P(F|G') = \dfrac{P(F \cap G')}{P(G')} = \dfrac{\frac{1}{6} \times \cancel{(1-\frac{1}{4})}}{\cancel{(1-\frac{1}{4})}}$

$= \boxed{\dfrac{1}{6}}$

<u>alt</u> : F and G are independent

so $P(F|G') = P(F) = \dfrac{1}{6}$

and total probabilities $= 1$

$\Rightarrow a + b + a + b + 0.2 = 1$

$\Rightarrow 2a + 2b = 0.8$

$\Rightarrow a + b = 0.4 \quad — (2)$

(1) - (2) :

$$\begin{array}{r} 3a + b = 1 \\ -[\,a + b = 0.4\,] \\ \hline 2a + 0 = 0.6 \end{array}$$

$\therefore \boxed{a = 0.3}$

$\hookrightarrow b = 0.4 - 0.3 = \boxed{0.1}$

c) $Var(1-3X) = (-3)^2 Var(X) = 9\,Var(X)$

$= 9E(X^2)$

$E(X^2) = \sum x^2 P(X = x)$

$= 16a + 9b + a + 4b + 25(0.2)$

$= 17(0.3) + 13(0.1) + 25(0.2) = 11.4$ //

hence $9E(X^2) = 9 \times 11.4 = \boxed{103} = Var(1-3X)$

di) $P(Y<0) = P(1-X<0) = P(X>1)$

$= P(X=2) + P(X=5) = 0.1 + 0.2 = \boxed{0.3}$

Q3a) $r \simeq -1$ $\therefore$ strong $-$ve correlation.

b) The dependent variable.

c) $b = \frac{S_{ch}}{S_{cc}} = \frac{-3034.6}{303448} = -0.01$

$h = a + \underline{b}c$

$100 \times -0.01 = -1$ (hence the data support this statement)

for 100mg of caffeine

d) $a = \bar{h} - b\bar{c} = \frac{126}{20} - 0.01\left(\frac{3660}{20}\right)$

$= 8.13$

$h = a + bc$

when $c = 0$, $h = a = \boxed{8.13}$

Q4a) $E(X) = \sum_i x P(X = x) = \boxed{-4a - 3b + a + 2b + 1}$

bi) $Var(X) = E(X^2) - [E(X)]^2$

so if $Var(X) = E(X^2)$ then $\boxed{E(X) = 0}$

ii) $\Rightarrow -3a - b + 1 = 0$

$\Rightarrow 3a + b = 1$ ~ 1

ii) $P(y < u) = 0 \cdot 2$

$P(1 - x < u) = 0 \cdot 2$

$P(x > 1 - u) = 0 \cdot 2$

largest value of u is when x is smallest.

$P(X > 2) = 0.2$ //

2 is the smallest value where this is true.

so $1 - u = 2$

$\boxed{u = -1}$

Q5a) $-2 = \dfrac{x - 20000}{1000}$

$\therefore x = -2000 + 20000 = \boxed{£18000}$

b) $PMCC = \dfrac{S_{wy}}{\sqrt{S_{ww} \times S_{yy}}}$

$S_{wy} = \sum wy - \dfrac{(\sum w)(\sum y)}{n}$

$S_{wy} = 2490 - \dfrac{(81)(405)}{9} = -1155$ //

$\therefore r = \dfrac{-1155}{\sqrt{660 \times 2500}} = \boxed{-0.899}$

c) -0.899 [as PMCC is unaffected by coding.]

d) $y = 60.75 - 1.75w$

$y = 60.75 - 1.75\left(\frac{x - 20000}{1000}\right)$

$y = 60.75 + 1.75(20) - \frac{1.75x}{1000}$

$\boxed{y = 95.75 - 0.00175x}$ //

e) $\underline{x = 21000}$: $y = 95.75 - 0.00175(21000)$

$= \boxed{59}$ min

f) Data range is from $x = 18k$ to $x = 45k$.

$25k - 40k$ is within this range so any estimates are interpolations so they are <u>reliable</u>.

Q6a) 0 - can't get both blue socks in one selection.

b) $P(S>2) = 1 - P(S \leq 2) = 1 - [P(S=2)]$

$= 1 - \left(\frac{2}{5}\right)\left(\frac{1}{4}\right) = \boxed{\frac{9}{10}}$

c) $P(S=3) = \left(\frac{2}{5}\right)\left(\frac{3}{4}\right)\left(\frac{1}{3}\right) \times 2 = \boxed{\frac{1}{5}}$

(2 possible ways for $S=3$)

d) $P(S=3 \mid \text{second sock is blue}) = \dfrac{P(S=3 \cap 2^{nd} \text{ sock blue})}{P(2^{nd} \text{ sock blue})}$

$= \dfrac{\left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{1}{3}\right)}{\left(\frac{3}{5}\right)\left(\frac{2}{4}\right) + \left(\frac{2}{5}\right)\left(\frac{1}{4}\right)} = \boxed{\frac{1}{4}}$

$P(BB) + P(B'B)$

if 2^{nd} sock is blue and $S=3$ then socks 2 and 3 have to be blue

e) $P(S=5) = P(BB'B'B'B)$
$+ P(B'BB'B'B)$
$+ P(B'B'BB'B)$
$+ P(B'B'B'BB)$

these will all be the same since socks are chosen at random

5^{th} sock must be blue

$= 4 \times \left(\frac{3}{5}\right)\left(\frac{2}{4}\right)\left(\frac{1}{3}\right)\left(\frac{2}{2}\right)\left(\frac{1}{1}\right) = \boxed{\frac{2}{5}}$

Q7a) $G \sim N(180, 15^2)$

$P(G > 174) = P\left(Z > \frac{174-180}{15}\right) = P(Z > -0.4)$

$= P(Z < 0.4) = \boxed{0.6554}$

b) $P(k < G < 174) = P(G < 174) - P(G < k)$

$= (1 - 0.6554) - P(G < k)$

$\Rightarrow \quad 0.3446 - P(G < k) = 0.3196$

$\Rightarrow \quad P(G < k) = 0.025$

$\Rightarrow \quad P\left(Z < \frac{k-180}{15}\right) = 0.025$

$\Rightarrow \quad P\left(Z > \frac{180-k}{15}\right) = 0.025$

but $\quad P(Z > 1.96) = 0.025$

so $\quad 1.96 = \frac{180-k}{15}$

$\therefore k = 180 - 15(1.96) = \boxed{150.6}$

ci) $B \sim N(216, 30^2)$

$P(G > w) = P(B < w)$

$P\left(Z > \frac{w-180}{15}\right) = P\left(Z < \frac{w-216}{30}\right)$

so $P\left(Z > \frac{w-180}{15}\right) = P\left(Z > \frac{216-w}{30}\right)$

hence $\frac{w-180}{15} = \frac{216-w}{30}$

$\underline{\times 30}$: $2w - 360 = 216 - w$

$3w = 576$

$\therefore w = \frac{576}{3} = \boxed{192}$

ii) $P(G > w) = p = P\left(Z > \frac{192-180}{15}\right)$

$= P(Z > 0.8)$

$= 1 - P(Z < 0.8)$

$= 1 - 0.7881 = \boxed{0.2119}$