Stats 1 - June 2007

(1) a) From calc: $r = -0.525443...$

b) Weak negative linear correlation between length and maximum diameter of honeydew melons.

(2) a) i) $7/50$

ii) $\frac{14+8}{50} = \frac{22}{50}$

iii) $1 - \frac{22}{50} = \frac{28}{50}$

iv) TOTAL BACKS = 23
BACKS + IRISH = 6 $\rightarrow \frac{6}{23}$

v) TOTAL NOT SCOTTISH = 46
FORWARD + NOT SCOTTISH = 25 $\rightarrow \frac{25}{46}$

b) P(English) = $22/50$

P(4 - without replacement) = $\frac{22}{50} \times \frac{21}{49} \times \frac{20}{48} \times \frac{19}{47} = 0.03176$

(3) a) $\bar{x} = 234$ $s = 25.1$ $n = 50$

95% multiplier (2 tailed) = 1.96

$\mu = \bar{x} \pm z \times s/\sqrt{n}$

$\mu = 234 \pm 1.96 \times 25.1/\sqrt{50}$

$\mu = 234 \pm 6.9573...$

$= (227.04, 240.96)$

b) Customers likely to choose big potatoes!

(4) a) i) MODE = 2

RANGE = 15 - 0 = 15

ii) Cumulative frequency:

BOOKS	0	1	2	3	4	5 - 9	10 - 14	15
CF	4	17	41	58	~~73~~	84	89	95

Median = $\frac{95+1}{2}$ = 48th person = 3 books

LQ = $\frac{95+1}{4}$ = 24th person = 2 books

UQ = $\frac{3(95+1)}{4}$ = 72nd person = 4 books

IQR = UQ − LQ = 4 − 2 = 2 books

iii) Use midpoints for x values: 5–9 = 7, 10–14 = 12

From calc: $\Sigma x^2 = 3111$, $\Sigma x = 399$

$\rightarrow \bar{x} = 4.2$, $s = 3.9074$

b) i) $\bar{x}$ and s are estimated using midpoints for 2 groups. These unknown values have no impact on median or IQR.

ii) Mean & SD use all available data.
Mode & range do not.

⑤ a) Time taken depends on temperature.

b) From calc: $a = 5.950909...$

$b = -0.0872727...$

$\rightarrow y = 5.95 - 0.087x$

c) i) Each 1° increase in temperature leads to a 0.087 minute decrease in the time taken to dissolve.

ii) a is the value when $x = 0$
water is frozen at this temperature!

d) i) $x = 30 \rightarrow y = 5.95 - 0.087(30)$

$= 3.33$ mins

ii) when $x = 75 \rightarrow y = 5.95 - 0.087(75)$

$= -0.575$

It is impossible for time to be negative.

(6) a.i) $T \sim B(10, 0.4)$

$P(T \leq 3) = 0.3823$ (from tables)

ii) $T \sim B(40, 0.4)$

$P(10 < T < 20)$

can be: 11, 12, 13 ... 18, 19

$\rightarrow P(T \leq 19) - P(T \leq 10)$

$= 0.8702 - 0.0352 = 0.835$

b) i) $\text{mean} = np = 0.4 \times 5 = 2$

$\text{var} = np(1-p) = 0.4 \times 0.6 \times 5 = 1.2$

$\rightarrow sd = \sqrt{1.2} = 1.09544...$

ii) From calc: $\Sigma x^2 = 68$, $\Sigma x = 26$

$\bar{x} = 2$

$s = 1.154700...$

iii) Mean is identical, standard deviations are very similar.

$\therefore$ the claim appears valid.

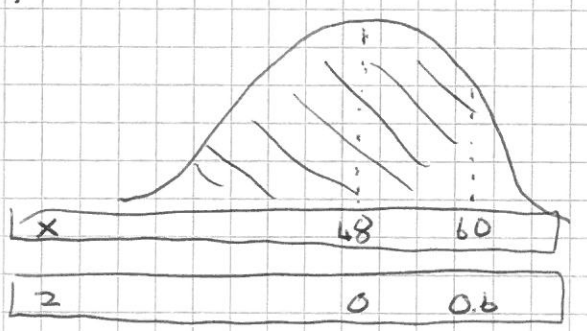
(7) a) $X \sim N(48, 20^2)$

i) $P(X < 60)$

$= P\left(Z < \frac{60-48}{20}\right)$

$= P(Z < 0.6)$

$= 0.72575$

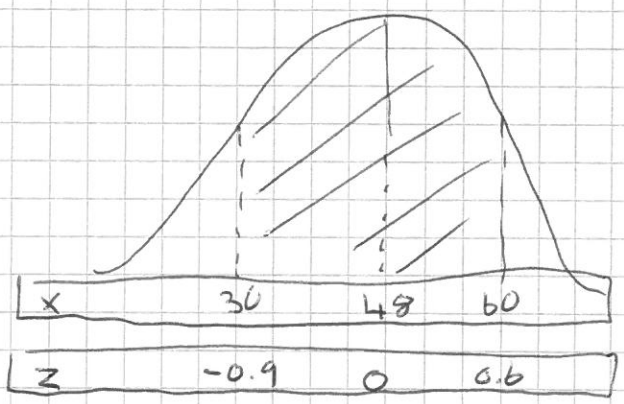


ii) $P(30 < X < 60)$

$= P(X < 60) - P(X < 30)$

$= P(Z < 0.6) - P\left(Z < \frac{30-48}{20}\right)$

$= P(Z < 0.6) - P(Z < -0.9)$



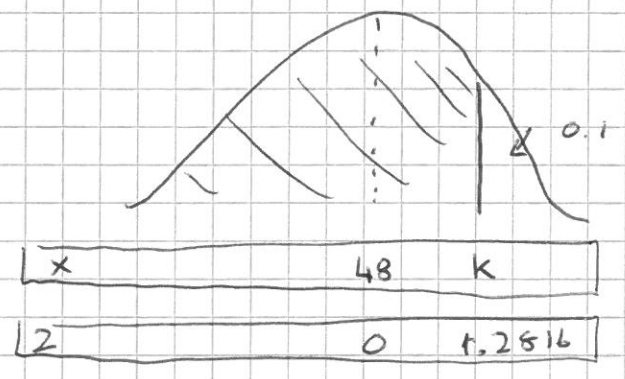
$= P(Z < 0.6) - [1 - P(Z < 0.4)]$

$= 0.72575 - [1 - 0.65542]$

$= 0.72575 - 0.34458$

$= 0.38117$

iii) $P(X < K) = 0.9$



Look up 2 of 0.9

$= 1.2816$

Now standardise:

$\frac{K - 48}{20} = 1.2816$

$\rightarrow K = 1.2816 \times 20 + 48$

$= 73.632$

b) $Y \sim N(37, 25^2)$

i) Large standard deviation.

So, if we do 2 SDs from mean we get:

$37 - 2(25) = -13$, which implies negative times.

ii) Central Limit Theorem can be used as $n > 30$

iii) $\bar{Y} \sim N(37, 25^2/35)$

$P(\bar{Y} > 40)$

$= P\left(Z > \frac{40 - 37}{25/\sqrt{35}}\right)$

$= P(Z > 0.70492)$

$= 1 - P(Z < 0.70492)$

$= 1 - P(Z < 0.71)$

$= 1 - 0.76115$

$= 0.23885$