

Stats 1 - May/June 2008

① a) From calculator: $\sum x^2 = 3986$

$a = 53.0673\dots$ (intercept)

$b = -1.00336\dots$ (gradient)

$\rightarrow y = 53.07 - 1.00x$

b) $x = 21 \rightarrow y = 53.07 - 1(21) = 32.1$ mins

② a) $P(B) = 160/400$

b) $P(\text{marker}) = 280/400$

c) $P(B \cup \text{Marker}) = \frac{160 + 280 - 119}{400} = \frac{321}{400}$

d) $P(G|H) = \frac{42}{120}$

e) $P(\text{non-M} | R) = 21/90$

Tip: Put totals on table to help calculations

③ a) From calculator: $\sum x^2 = 681.575$

$r = 0.806559\dots$

b) Strong, positive, linear correlation between lengths and widths of ceramic plates

c) See Mark Scheme

d) A to F: zero linear correlation $\rightarrow r = 0$
G to L: zero linear correlation $\rightarrow r = 0$

④ a) In order: 0 0 13 28 35 40 47 51 63 77 a

Median = $\frac{11+1}{2}$ = 6th value = 40

LQ = $\frac{11+1}{4}$ = 3rd value = 13

UQ = $\frac{3(11+1)}{4}$ = 9th value = 63

IQR = $63 - 13 = 50$

b) i) MODE - gives 0 which is not representative of the data

ii) RANGE - cannot calculate as we don't know a

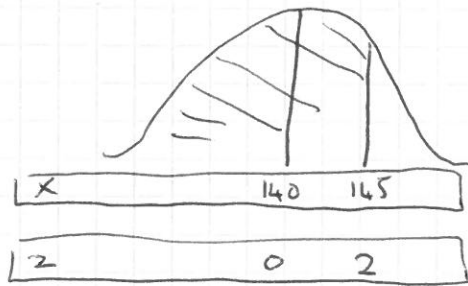
5) $X \sim N(140, 2.5^2)$

a) i) $P(X < 145)$

$$= P\left(Z < \frac{145 - 140}{2.5}\right)$$

$$= P(Z < 2)$$

$$= 0.97725$$



ii) $P(138 < X < 142)$

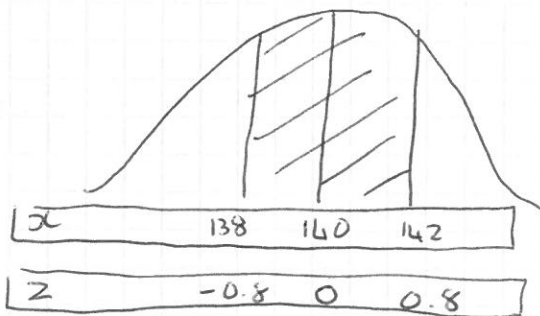
$$= P\left(\frac{138 - 140}{2.5} < Z < \frac{142 - 140}{2.5}\right)$$

$$= P(-0.8 < Z < 0.8)$$

$$= P(Z < 0.8) - P(Z < -0.8)$$

$$\downarrow$$
$$0.78814$$

$$\downarrow$$
$$= P(Z > 0.8)$$
$$= 1 - P(Z < 0.8)$$
$$= 1 - 0.78814 = 0.21186$$



$$\rightarrow 0.78814 - 0.21186 = 0.57628$$

b) $P(X > h) = 0.85$

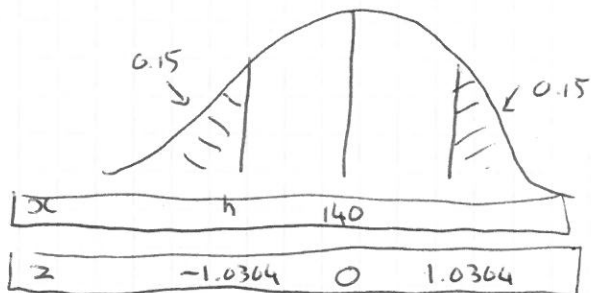
Z value for 0.85 = 1.0364

Standardize:

$$\frac{h - 140}{2.5} = -1.0364$$

$$h = 2.5(-1.0364) + 140$$

$$= 137.409$$

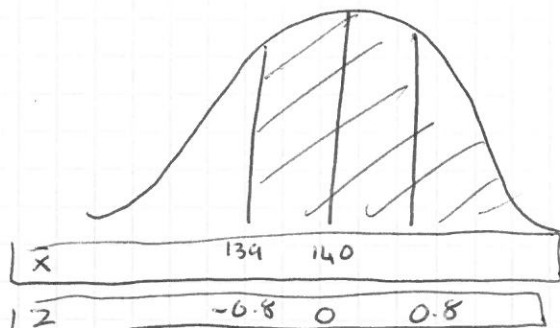


c) $\bar{X} \sim N(140, 2.5^2/4)$

$$P(\bar{X} > 139)$$

$$= P\left(Z > \frac{139 - 140}{2.5/\sqrt{4}}\right)$$

$$= P(Z > -0.8)$$



$$= P(Z < 0.8) = 0.78814$$

(b) a) $M \sim B(40, 0.35)$

i) $P(M \leq 15) = 0.6946$ (from tables)

ii) $P(10 < M < 20)$

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

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

$$= 0.9637 - 0.1215 = 0.8422$$

b) $W \sim B(10, 0.29)$

$$P(W = 3) = \binom{10}{3} \times 0.29^3 \times 0.71^7$$

$$= 0.26618, \dots$$

c) i) $W_{10} \sim B(20, 0.71)$

MEAN: $np = 20 \times 0.71 = 14.2$

VARIANCE: $np(1-p) = 20 \times 0.71 \times 0.29 = 4.118$

ii) Both the means and the variances are different,

This suggests the claim that they are not random samples is justified.

(7) a) i) From calculator: $\sum x^2 = 1452$ $\sum x = 190$

$$\bar{x} = 1.9$$

$$s = 3.31966, \dots$$

use midpoints for x

ii) 60 mins have been subtracted

$$\therefore \text{Mean} = 1.9 + 60 = 61.9$$

$$\text{Standard deviation} = \text{unchanged} = 3.31966, \dots$$

b) $\bar{x} = 61.9$ $s = 3.31966$ $n = 100$

Z value for 98% (2 tails) = 2.3263

$$98\% \text{ CI for } \mu = \bar{x} \pm z \times \frac{s}{\sqrt{n}}$$

$$= 61.9 \pm 2.3263 \times \frac{3.31966}{\sqrt{100}}$$

$$= 61.9 \pm 0.77225 \dots$$

$$= (61.128, 62.672)$$

ii) Mean and Standard Deviation based on grouped data with midpoints used

c) "More than 1 hour on average" : I agree as 60 mins is below lower bound of confidence interval

"More than 1 hour more often than not" : I agree as $\frac{74}{100}$ in our sample were more than 60 mins