Stats 1 – January 2006

① a) From calculator: $a = 2.318181\ldots$

$b = 0.886363\ldots$

$\rightarrow \quad y = 2.32 + 0.886x$

b) a = average waiting time of 2.32 minutes if there are 0 customers already seated (ie. $x = 0$)

b = the increase in waiting time of 0.886 mins that each additional customer entering the restaurant causes.

c) i) $x = 5 \rightarrow y = 2.32 + 0.886(5) = 6.75$ mins

ii) $x = 25 \rightarrow y = 2.32 + 0.886(25) = 24.47$ mins

d) i) $x = 5$ is in range of data, and -1.1 to $+1.1$ is fairly small residuals, so result is reliable

$x = 25$ lies outside data range so result is unreliable (extrapolation).

② a) i) $P(L, L, L) = 0.3 \times 0.4 \times 0.2 = 0.024$

ii) $P(L', L', L') = 0.7 \times 0.6 \times 0.8 = 0.336$

iii) $P(L', L', L) = 0.7 \times 0.6 \times 0.2 = 0.084$

b) i) $P(ZL, WL) = 0.2 \times 0.9 = 0.18$

ii) $P(ZL, WL') = 0.2 \times 0.1 = 0.02$

$P(ZL', WL) = 0.8 \times 0.25 = 0.20$

Probability of either $= 0.02 + 0.2 = 0.22$

(3) a) $\bar{x} = \dfrac{286.5}{50} = 5.73$

$s = \sqrt{\dfrac{45.16}{49}} = \sqrt{0.921\ldots} = 0.96001\ldots$

b) 99% (2 tailed) z multiplier = 2.5758

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

$\mu = 5.73 \pm 2.5758 \times \dfrac{0.96001\ldots}{\sqrt{50}}$

$\mu = 5.73 \pm 0.3497\ldots$

$\mu = (5.38,\ 6.08)$

c) The confidence interval does not include 5 or 6½, so both claims apper invalid.

(4) a) From calculator: $\sum fx = 8025$ $\quad n = 100$
(use midpoints for x) $\sum fx^2 = 739975$

Mean ($\bar{x}$) = 80.25

Standard deviation (s) = 31.134889

b) i) Sample size is >30, so using Central Limit Theorem

ii) Mean ($\bar{Y}$) is same = 80.25

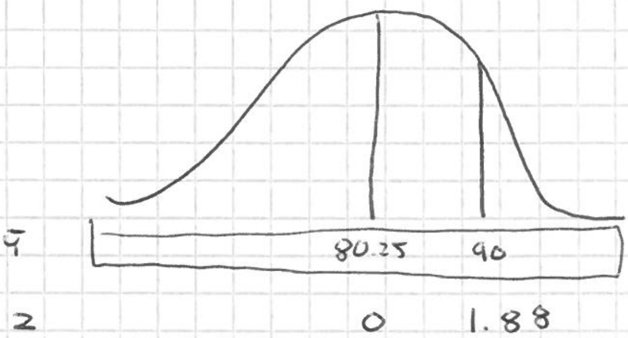
Standard error ($\bar{Y}$) = $\dfrac{s}{\sqrt{n}} = \dfrac{31.13\ldots}{\sqrt{36}} = 5.1891\ldots$

iii) $\bar{Y} \sim N(80.25,\ 5.1819^2)$

1½ mins = 90 secs

$P(\bar{Y} < 90)$

$= P\left(Z < \dfrac{90 - 80.25}{5.181}\right)$



$= P(Z < 1.8818\ldots)$

$\sim P(Z < 1.88) = 0.96995$

(5) a) See Scatter Diagram (not included here!)

b) i) (1) Positive Linear correlation between times

(2) Two outliers, C & D

ii) 0.462 - fairly weak positive

c) C = Free style Champion

D = Backstroke Champion

d) i) From calculator : $r = 0.91242...$

ii) Strong, positive correlation → boys are equally good / bad at both strokes.

(6) $R \sim B(50, 0.2)$

a) i) $P(R \leq 15) = 0.9692$ (from tables)

ii) $P(R = 10) = {}^{50}C_{10} \times 0.2^{10} \times 0.8^{40}$

$= 0.13981...$

iii) $P(5 < R < 15)$

can be : 6, 7, 8 ... 14,

so need $P(R \leq 14) - P(R \leq 5)$

$= 0.9393 - 0.0480 = 0.8913$

b) $\boxed{\text{MEAN}}$ Population mean $(\mu) = np = 50 \times 0.2 = 10$

Sly claims 10.5

Mean is similar to expected

$\boxed{\text{VARIANCE}}$ Population Variance $(\sigma) = np(1-p)$

$= 50 \times 0.2 \times 0.8 = 8$

Sly claims 20.41

This is very different from expected

Therefore we have reason to doubt Sly's claims

(7) $X \sim N(406, 4.2^2)$

a) i) $P(X < 400)$

$= P\left(Z < \frac{400-406}{4.2}\right)$

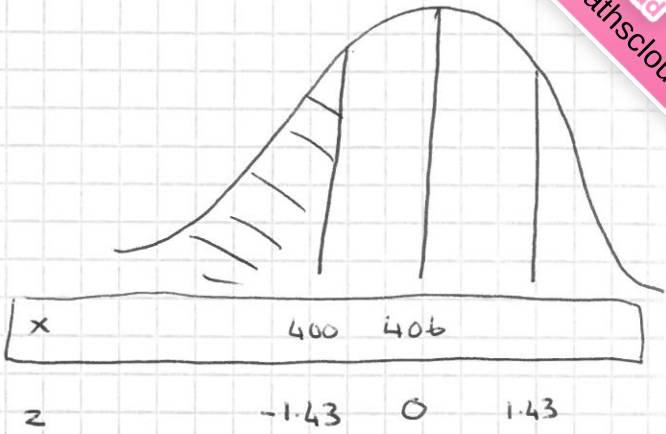
$= P(Z < -1.43)$

$= P(Z > 1.43)$

$= 1 - P(Z < 1.43)$

$= 1 - 0.92364$

$= 0.07636$



a) ii) $P(402.5 < X < 407.5)$

$= P\left(\frac{402.5-406}{4.2} < Z < \frac{407.5-406}{4.2}\right)$

$= P(-0.83 < Z < 0.36)$

Need: $P(Z < 0.36) - P(Z < -0.83)$

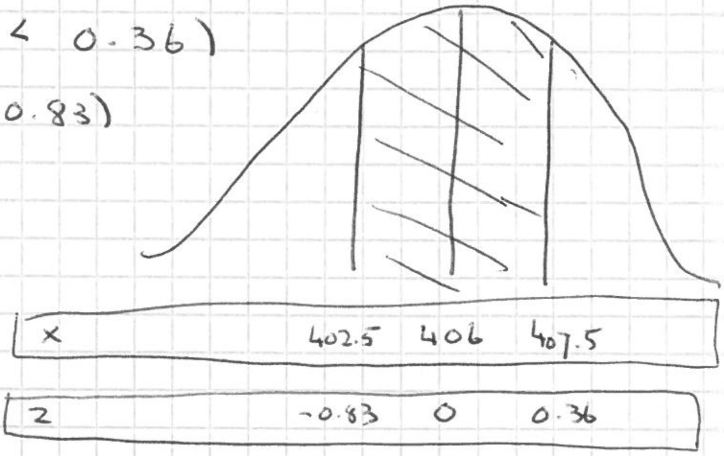
0.64058

$= P(Z > 0.83)$

$= 1 - P(Z < 0.83)$

$= 1 - 0.79673$

$= 0.20327$



$\rightarrow 0.64058 - 0.20327 = 0.43731$

b) i) $P(Y < 310) = 0.975$

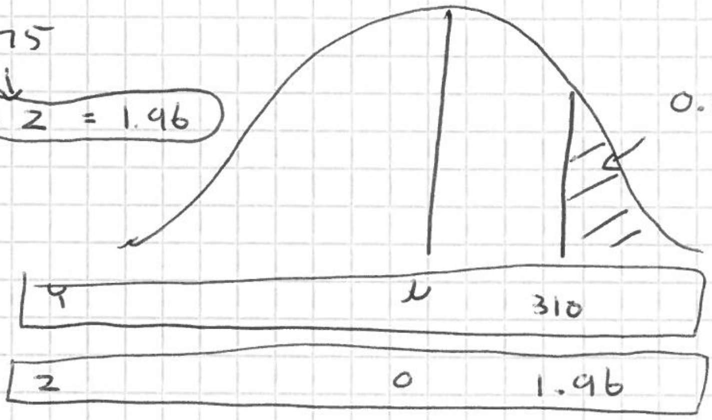
$\rightarrow Z = 1.96$

Standardise:

$\frac{310 - \mu}{\sigma} = 1.96$

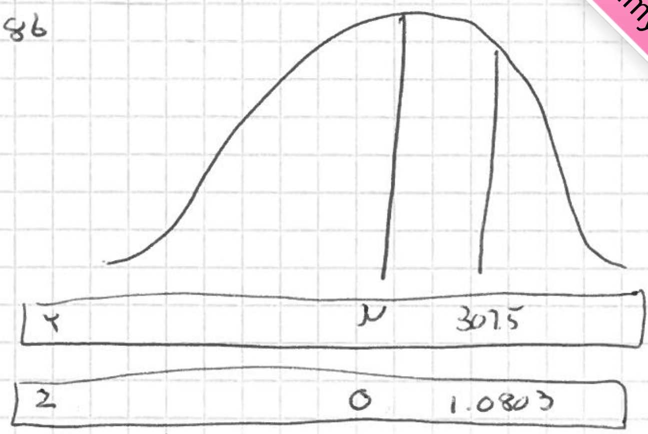
$310 - \mu = 1.96\sigma$

0.025



ii) $P(Y < 307.5) = 0.86$

$0.86 \rightarrow z = 1.0803$



Standardise:

$$\frac{307.5 - \mu}{\sigma} = 1.0803$$

$$307.5 - \mu = 1.0803\sigma$$

Simultaneous Equations

$$\begin{array}{rl} & 310 - \mu = 1.96\sigma \\ - & 307.5 - \mu = 1.0803\sigma \\ \hline & 2.5 = 0.8797\sigma \end{array}$$

$\rightarrow \sigma = 2.5 \div 0.8797 = 2.84$ (2dp)

$310 - \mu = 1.96(2.8418\ldots)$

$310 - \mu = 5.570\ldots$

$\mu = 310 - 5.570\ldots$

$= 304.43$ (2dp)