Question 1(a)

Candidate 1 evidence

QUESTION NUMBER		DO NOT WRITE IN THIS MARGIN
1.(a)	X = height of boys cm $X \sim N(109, 7^2)$ $P(X>111) = P\left(Z > \frac{111-109}{7}\right)$ $= P(Z > 0.285714)$ $= 0.387549$ from normCdf(0.29, 9E99) $= \underline{38.75\%}$ (2dp)	

Question 1(b)

Candidate 2 evidence

1.(b)

$$X \sim N\left(109, \frac{7^2}{25}\right)$$

$$P(X > 111)$$

$$= 1 - P(X \leq 111)$$

$$= 1 - P\left(2 \leq \frac{111 - 109}{\sqrt{\frac{7^2}{25}}}\right)$$

$$= 1 - P(2 \leq 1.43)$$

$$= 1 - 0.9236$$

$$= 0.0764$$

Candidate 3 evidence

1.(b)

$n = 25$

Sample mean is normally distrubted, ~~n≥20~~ $n \geq 20$

$$X \sim N\left(109, \left(\frac{7}{25}\right)^2\right) = 0 \cdot 07656$$

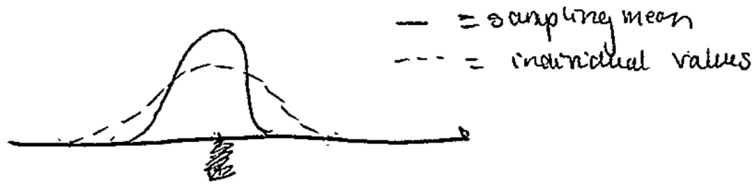
$7 \cdot 656\%$

Question 1(c)

Candidate 4 evidence

NUMBER

1.(c) The variation for the sampling mean is smaller than for the individual values of heights, so it is less likely for the sampling mean to be more than 11 than it is for an individual value.



Candidate 5 evidence

NUMBER

1.(c) Because the sample mean is approx normally distributed and has a smaller spread it is more likely to be close to the population mean than 1 member of the population will be. As a result the probability it will ~~be~~ have a height on average higher than the population mean is lower than the chance an individual will have a height higher than the population mean.

Candidate 6 evidence

QUESTION NUMBER	
1.(c)	It is much less likely that the average of 25 bags will be ~~larger by~~ 0.29 standard deviations above the mean. ~~Also the~~ Also, as the sample size is greater than 20 (n=25) CLT can be invoked.

Question 2

Candidate 7 evidence

QUESTION NUMBER

2.

the distribution is symmetrical

$n = 10$

~~m~~ mean = 300 (one tailed test)

~~-20 -10 1 -16~~

20 10 21 4 ~~-2~~ 28 ~~-4~~ 7 14 ~~-5~~

-2 4 -4 -5 7 10 14 20 21 28

1 +2.5 ~~-2.5~~ -4 5 6 7 8 9 10

$W_+ = 45$

$\boxed{W_- = ~~7~~} ~~+~~ 7.5$

$W = 7 + 0.5$ 7.5

5% level $\therefore$ ~~0.0095~~, 10

Critical value $W = 10$

Since ~~7~~ 7.5 < 10, we reject H_0 at 5% significance level

Interpretation: There is evidence to suggest that step counter from mobile phone overcounts the median number of steps that Duncan takes

Question 3(a)

Candidate 8 evidence

3.(a)

~~200~~

~~B6 %~~

~~B≤2 = B≤1~~

~~B~~ $B \geq 2 = B > 1$

Binomcdf $(20, 0.02, 1)$

$= 0.94$

Candidate 9 evidence

QUESTION NUMBER

3.(a)

let x be rv of number of scots with type B−

$x \sim bin(20, 0.02)$

$P(x \geq 2) = \cancel{P(X \geq 1)}$

$= \cancel{1 - \text{binomcdf}(20, 0.02, 1)}$

$= \cancel{0.0599}$

$= 1 - \text{binomcdf}(20, 0.02, 2)$

$= 0.0071$

Question 3(b)

Candidate 10 evidence

3.(b)

We can use a normal approximation as $np = 25\cdot2 > 5$

$npq = 12\cdot4992 > 5$

$p = 0\cdot504 \quad q = 0\cdot496$

$$X \approx N(25\cdot2,\ 12\cdot4992)$$

$$P(X \leq 30) = P\left(z \leq \frac{30 - 25\cdot2}{\sqrt{12\cdot4992}}\right)$$

$$= P(z \leq 1\cdot3577)$$

$$= 0\cdot9115$$

Candidate 11 evidence

3.(b)

~~A = no. blood donors in group O+ or O-~~

~~$A \sim B(50,$~~

M = no. blood donors in group O-

$M \sim B(50, 0.095)$

A = no. blood donors in group O+

$A \sim B(50, 0.409)$

~~Let N be normal approx to M~~

C = no blood donors in group O+ or O-

$C \sim B(50, 0.409 + 0.095)$

$C \sim B(50, 0.504)$

Let Y be normal approx to C $np > 5$ ✓ $nq > 5$ ✓

~~$Y \approx N(25,$~~

$Y \approx N(50 \times 0.504, 50 \times 0.504 \times 0.496)$

3.(b) (cont)

$$\frac{Y}{50} \approx N\left(0.504, \frac{0.504 \times 0.496}{50}\right)$$

$$P(C \leq 30) \approx P\left(\frac{Y}{50} < \frac{30.5}{50}\right) \text{ by c.c.}$$

$$= P\left(Z < \frac{\frac{30.5}{50} - 0.504}{\sqrt{\frac{0.504 \times 0.496}{50}}}\right)$$

$$= P\left(Z < \frac{0.106}{0.070708}\right)$$

$$= P(Z < 1.49911)$$

$= 0.933078$ from normCdf(-9E99, 1.5)

$= \underline{0.9331}$ (4dp)

Question 5(a)(ii)

Candidate 12 evidence

5.(a) (ii)	$E(x) = (0 \times p) + (1 \times p) + (2 \times 2p) + (3 \times 5p) + (4 \times p)$ $= p + 4p + 15p + 4p$ $E(x) = 24p$ $E(x) = 3$, $3 = 24p$ $p = \frac{3}{24}$ $= 0.125$ $V(x) = (Ex)^2 - E(x^2)$ $E(x^2) = 0^2 \times \frac{1}{8} + 1^2 \times \frac{1}{8} + 2^2 \times \frac{2}{8} + 3^2 \times \frac{4}{8} + 4^2 \times \frac{1}{8}$ $= \frac{1}{8} + 1 + 4.5 + 2$ $= 7.625$ $V(x) = 9 - 7.625$ $(Ex)^2 = 3^2 = 9$ $V(x) = 1.375$

Question 6

Candidate 13 evidence

6.

~~Hypothesis~~

~~$H_0: \mu_1 = \mu_2$~~

~~$H_1: \mu_1 < \mu_2$~~

~~μ_1 = population~~

~~μ_2 = basketball players~~

$\frac{x-\mu}{\sigma}$

Z test

$S_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$

$= 198240 - \frac{3840^2}{75}$

$= 1632$

$s = \sqrt{\frac{S_{xx}}{n-1}}$

$= 4.7$

$\sigma = 4.7$

~~$\bar{x} = 50$~~

~~$\bar{x} > 50$~~

~~Hypothesis~~

~~$H_0: \mu = 50$~~ ~~$\alpha = 1\%$~~ ~~1 tail~~

~~$H_1: \mu > 50$~~

~~$P(X > 50) = P(X > 51)$~~

$= P(Z >$

find $Z > 2.33$

Hypotheses

$H_0: \bar{x} = 50$

$H_1: \bar{x} > 50$

$\bar{x} = \frac{\sum x}{n}$

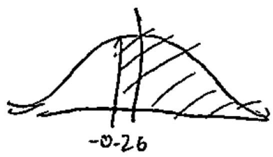
$= \frac{3840}{75}$

$= 51.2$

$P(\bar{x} > 50) = P(Z > \frac{x - \bar{x}}{\sigma})$

$= P(Z > \frac{50 - 51.2}{4.7})$

$= P(Z > -0.26)$



$CI = \bar{x} \pm z(\frac{\sigma^2}{n})$

$CI = 51.2 \pm 2.33 \times (\frac{4.7^2}{75})$

$= 51.2 \pm 0.69$

$CI = (50.51, 51.89)$

Since 50 is outside the interval, the midwifes theory is correct

Candidate 14 evidence

QUESTION NUMBER

DO NOT WRITE IN THIS MARGIN

6.

X = Length baby bifth cm

$X \sim N(50, \sigma^2)$

$n = 75$

H_0: mean baby length pop = mean baby length basketball

H_1: mean baby length pop < mean baby length basketball

Assume H_0 to be true

$\alpha = 1\%$ One-tail test

$\bar{X} \approx N(50, \frac{\sigma^2}{75})$ by CLT as $n \geq 20$

as we don't know σ^2, we estimate with S^2 and use the t_{74} distribution

$s = \sqrt{\frac{S_{xx}}{n-1}}$ ~~[illegible]~~

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

~~$= 3840$~~

$= 198240 - \frac{(3840)^2}{75}$

$= 1632$

~~s~~ $= \sqrt{\frac{1632}{74}}$

$= 4.69617$

$s^2 = 22.0541$

~~$\bar{X} \approx N(50, \frac{22.0541}{75})$~~

$\bar{x} = 51.2$

p-value $= P(\bar{X} \geq 50)$

$= P\left(t_{74} \geq \frac{51.2 - 50}{\sqrt{\frac{22.0541}{75}}}\right)$

$= P(t_{74} > 2.21293)$

$= 0.014993$ from tcdf(2.21,9E99,74)

continued back a page

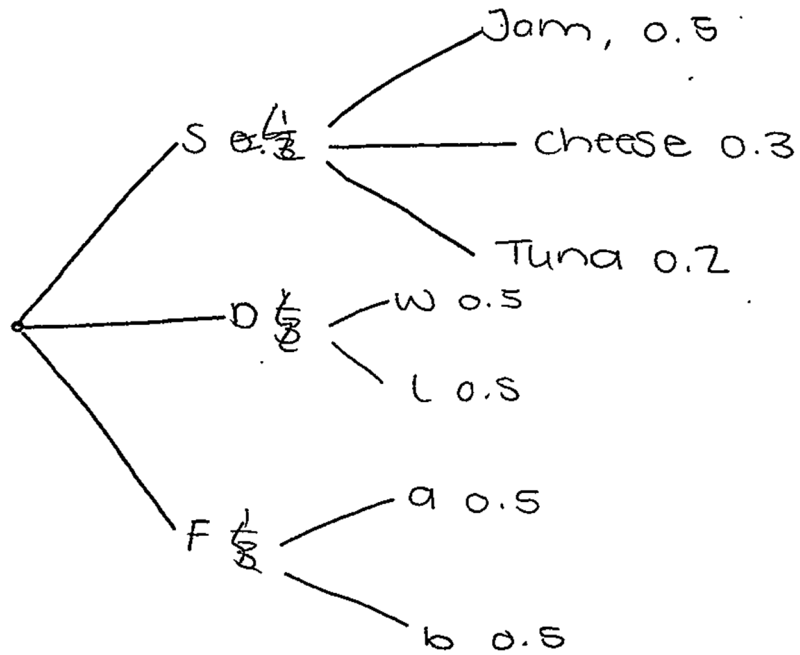
6)	$= 0.014993$ > 0.01, we are not in the critical region we do not have evidence to reject H_0, so we conclude the midwife theory is incorrect at the 1% level	

Question 7(a)

Candidate 15 evidence

QUESTION NUMBER

7.(a)



$P(W \cap \text{Tuna})$

$= P(W) \times P(\text{Tuna})$

$= 0.5 \times 0.2$

$= 0.1$

Question 7(b)

Candidate 16 evidence

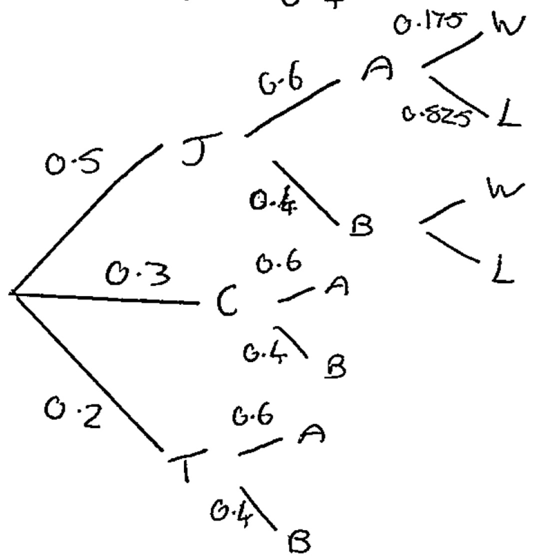
7.(b)

$P(C \cap B) = 0.12$

$0.3 \times P(\text{Banana}) = 0.12$

$P(\text{Banana}) = \dfrac{0.12}{0.3}$

$P(B) = 0.4$



$P(\text{Jam Sandwich and apple}) = P(J \cap A \cap W) + P(J \cap A \cap L)$

$P(\text{Tuna} \cap \text{Water}) = 0.035$

$P(T \cap A \cap W) + P(T \cap B \cap W) = 0.035$

$0.12w + 0.08w = 0.035$

$w = 0.175$

$P(J \cap A) = (0.5 \times 0.6 \times 0.175) + (0.5 \times 0.6 + 0.825)$

$= 0.0525 + 0.2475$

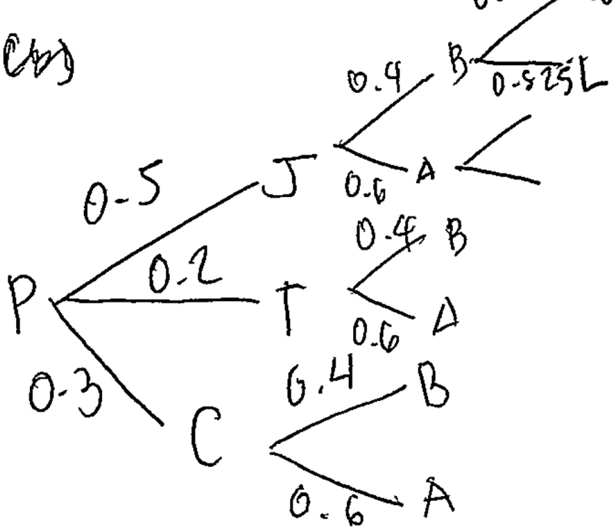
$P(J \cap A) = \underline{\underline{0.3}}$

Candidate 17 evidence

7.(b)

$P(C \cap B) = 0.12$

(b)



$0.2 \times X\,0.035$

$P(J) = 0.5$

$\cancel{0.5x}$

$\cancel{0.3}$

$P(J|A) = 0.3$

Question 9(a)

Candidate 18 evidence

QUESTION NUMBER 9.(a)

overall $n = 12$ ($m = 3$ $n = 9$)

1	2	3	4	5	6	7	8	9	10	11	12	runner
9	4	10	12	2	7.5	5	6	11	3	7.5	1	ranking

$W_m = 20.5$

$m(m+n+1) - W_m$

$3(3+9+1) - 20.5$

$= 18.5$

$W = 18.5$

$H_0: \mu_d = 0$

$H_1: \mu_d > 0$

~~two~~ One-tailed

$\alpha = 0.05$

$z = \frac{W - \mu}{\sigma}$

$cv = 1.64$

$= \frac{18.5 - 0.45}{0.927}$

~~1.96~~ 1.64 < 19.47 so reject H_0 as evidence at the 5% level

$= 19.47$

that the difference doesn't equal zero and the runners ran further when wearing a tracker

Candidate 19 evidence

9.(a)	~~WA~~ ~~Runners~~ H_0: Runners ran the same distance H_1: Runners ran further with a fitness tracker 1 - tail $0.05 = \alpha$ $t = \dfrac{\bar{x}_d}{\frac{s_d}{\sqrt{n}}} = \dfrac{0.45}{\frac{0.927}{\sqrt{12}}} = 1.6816$ $t_{11, 0.95} = 1.796$ As $1.796 > 1.6816$ there is insufficient evidence to reject H_0 at the 5% significance level and we must conclude that ~~the runners~~ there is not enough evidence to suggest the runners ran faster when wearing a fitness tracker.

Question 9(b)(i)

Candidate 20 evidence

9.(b) (i)	This normality assumption is not plausible, as the histogram is not symmetrical, and there is a heavy positive skew.

Question 9(b)(ii)

Candidate 21 evidence

9.(b) (ii)	Wilcoxon signed rank Requires the distribution of differences to be symmetrical. From this graph ~~this~~ assumption is dubious...

Question 10

Candidate 22 evidence

QUESTION NUMBER	
10.	$p_1 = 0.6240$ $p_2 = 0.6154$ $q_1 = 0.3760$ $q_2 = 0.3846$ $\alpha = \cancel{0.0025}$ 0.005 $CV = \cancel{3.29}$ 2.81 $H_0: p_1 = p_2$ $H_1: p_1 \neq p_2$ Increase 37878 homeless in 55 studies so assume same ~~[illegible]~~ for both 2010-2017 and 2018 $\hat{p} = \dfrac{37878 \times 0.624 + 37878 \times 0.6154}{75756} = \dfrac{23635.872 + 23310.1212}{75756} = \dfrac{46945.9932}{75756}$ $\hat{p} = 0.6197$ $\hat{q} = 0.3803$ $Z = \dfrac{0.624 - 0.6154}{\sqrt{0.6197 \times 0.3803\left(\frac{1}{37878} + \frac{1}{37878}\right)}}$ $= \dfrac{0.0086}{\sqrt{0.23567\left(\frac{1}{18939}\right)}}$ $= 2.44$ $\cancel{\approx} 2.437 = 2.44 < 2.81$ so don't reject H_0 so ~~there is~~ no evidence at 0.5% level that there is a significant difference between homeless people in sheltered accommodation in 2018 than in the previous 8 years

Question 11

Candidate 23 evidence

11.

$X \sim N(\mu, \sigma^2)$

$P(X > 24) = 0.05$

$P(X < 17) = 0.1$

$P\left(Z > \frac{24-\mu}{\sigma}\right) = 0.05$

$1 - P\left(Z \leq \frac{24-\mu}{\sigma}\right) = 0.05$

$P\left(Z \leq \frac{24-\mu}{\sigma}\right) = 0.95$

$\frac{24-\mu}{\sigma} = 1.64485$

$\cancel{1 - P} \quad \cancel{P(Z < 17)}$

$P\left(Z < \frac{17-\mu}{\sigma}\right) = 0.1$

$\cancel{1-P} \; \frac{17-\mu}{\sigma} = -1.28155$

$24 - \mu = 1.64485\sigma$

$17 - \mu = -1.28155\sigma$

$\frac{24-\mu}{17-\mu} = \frac{1.64485\sigma}{-1.28155\sigma}$

$-(24-\mu) \times 1.28155 = 1.64485(17-\mu)$

$1.28155\mu + 1.64485\mu = 17 \times 1.64485 + 1.28155 \times 24$

$\mu = 20.06549$

$\underline{\mu = 20.1}$

$\therefore \sigma = \frac{24 - 20.1}{1.64485}$

$\underline{\underline{\sigma = 2.371}}$

Question 12(a)

Candidate 24 evidence

12.(a)

$X \sim B(100, 0.55)$

$X \sim N(100 \times 0.55, 100 \times 0.55 \times (1-0.55))$

$X \sim N(55, 24.75)$

$99\%\ CI = \bar{x} \pm t_{99, 0.995}\sqrt{\sigma^2/n}$

$= 55 \pm 2.57583\sqrt{24.75/100}$

$= (53.7185, 56.2815)$

$= (53.72, 56.28)$

Only an approx interval as ~~not~~ actual interval cannot be determined. There will always be votes for both candidates

Candidate 25 evidence

NUMBER

12.(a)

$$99\% \text{ CI} = \hat{p} \pm \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$0{\cdot}55 \pm \sqrt{\frac{0{\cdot}55 \times (1-0{\cdot}55)}{100}}$$

$$0{\cdot}55 \pm 2{\cdot}58 \times 0{\cdot}04975$$

$$0{\cdot}55 \pm 0{\cdot}12835$$

$$= \underline{\underline{(0{\cdot}422,\ 0{\cdot}678)}}$$

This is only an approximate interval as it involves a Binomial to Normal approximation. (discrete to continuous data).

Question 12(b)

Candidate 26 evidence

QUESTION NUMBER

12.(b)

$$\hat{p} \pm k\sqrt{\frac{pq}{n}}$$

$$0.55 \pm 2.58\sqrt{\cancel{\frac{pq}{n}}\frac{0.55\times 0.45}{n}}$$

when $n = 700$ min $= 0.5015$
$n = 600$ min $= 0.4976$
$n = 650$ min $= 0.4999$
$n = 660$ min $= 0.50004$
$n = 659$ min $= 0.50000$
$n = 658$ min $= 0.49996$

for a 99% CI the smallest sample ~~si~~ size to indicate candidate would win is ~~65~~ ~~659~~. a sample of 659 voters if same sample proportion remains

Candidate 27 evidence

QUESTION NUMBER

12.(b)

$$\cancel{0{\cdot}5 = \hat{p} + k\sqrt{\frac{\hat{p}\hat{q}}{n}}}$$

$$0{\cdot}5 \le \hat{p} + k\sqrt{\frac{\hat{p}\,\hat{q}}{n}}$$

$$0{\cdot}5 - 0{\cdot}55 \le 2{\cdot}58\sqrt{\frac{0{\cdot}55 \times 0{\cdot}45}{n}}$$

$$\frac{-0{\cdot}05}{2{\cdot}58} \le \sqrt{\frac{0{\cdot}55 \times 0{\cdot}45}{n}}$$

$$\left(-\frac{5}{258}\right)^2 \le \frac{0{\cdot}55 \times 0{\cdot}45}{n}$$

$$\frac{25}{66564} \le \frac{0{\cdot}55 \times 0{\cdot}45}{n}$$

$$n \le \frac{0{\cdot}55 \times 0{\cdot}45}{\left(\frac{25}{66564}\right)}$$

$$n \le 658{\cdot}9836$$

$$n \le 659.$$