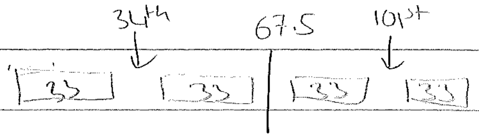
AH Stats Practice Prelim paper (A)

(1)

age	60	61	62	63	64	65	66	67	68	69	70	75	77
freq	10	12	11	8	16	14	9	13	18	15	6	1	1
	10	22	33	41	57	71	80	93	111	126	132	133	134

$n = 134 \qquad LQ = 63 \qquad M = 65 \qquad UQ = 68$

$IQR = 68 - 63 = 5$ 1



upper fence $= UQ + 1.5\,IQR$

$= 68 + 1.5 \times 5$

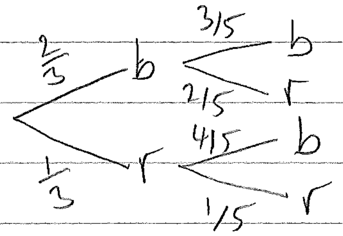
$= 75.5$ 1

(3)

$75 < 75.5$ so not an outlier

$77 > 75.5$ so this person is an outlier. 1

(2) a)



x	0	1	2
$P(X=x)$	$\frac{1}{15}$	$\frac{8}{15}$	$\frac{2}{5}$

2

b) $E(X) = 0 \times \frac{1}{15} + 1 \times \frac{8}{15} + 2 \times \frac{2}{5}$

$= \frac{4}{3}$ 1

$E(X^2) = 0^2 \times \frac{1}{15} + 1^2 \times \frac{8}{15} + 2^2 \times \frac{2}{5}$

$= \frac{32}{15}$

$V(X) = E(X^2) - (E(X))^2$

$= \frac{32}{15} - \left(\frac{4}{3}\right)^2$

$= \frac{16}{45}$ 1

(7)

c) $E(12 - 2X)$

$= 12 - 2E(X)$

$= 12 - 2 \times \frac{4}{3}$

$= 9\frac{1}{3}$ 1

$V(12 - 2X)$

$= 2^2 V(X)$

$= 4 \times \frac{16}{45}$

$= \frac{64}{45} = 1\frac{19}{45}$ 1

$\Rightarrow SD(12 - 2X)$

$= 1.19$ 1

(3) let r.v. X be no. of days with rain

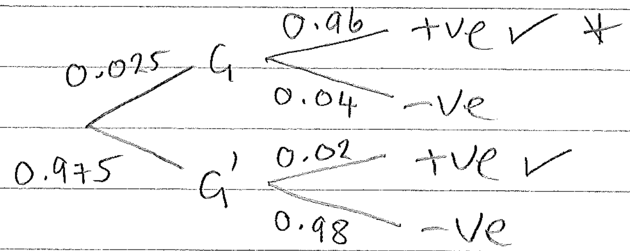
$\Rightarrow \quad X \sim B(14, 0.45)$

a) $P(X > 10) = 1 - P(X \leq 10) = 1 - 0.9886$
$= \underline{\underline{0.0114}}$ 3

9
8
7
6
5
4
3

b) $P(4 \leq X \leq 8) = P(X \leq 8) - P(X \leq 3)$
$= 0.8811 - 0.0632$
$= \underline{\underline{0.8179}}$

(4)



a) $P(+ve) = 0.025 \times 0.96 + 0.975 \times 0.02$
$= \underline{\underline{0.0435}}$ 2

b) $P(G \mid +ve) = \dfrac{P(G \cap +ve)}{P(+ve)}$ $\left(\dfrac{*}{✓ + ✓}\right)$

$= \dfrac{0.025 \times 0.96}{0.0435}$ 5

$= \underline{\underline{0.5517}}$

5) Let X be no of fabric flaws in 5m of rope

a) $X \sim Po(5 \times 1.5)$ $\quad P(X=6) = P(X \le 6) - P(X \le 5)$

$\Rightarrow X \sim Po(7.5)$ $\quad = 0.3782 - 0.2414$

$= \underline{\underline{0.1368}}$

b) Let Y be no of weaving flaws in 2m of rope

$Y \sim Po(2 \times 3)$ $\quad P(Y > 8) = 1 - P(Y \le 8)$

$\Rightarrow Y \sim Po(6)$ $\quad = 1 - 0.8472$

$= \underline{\underline{0.1528}}$

c) Let T be total flaws in 10m of rope

$T \sim Po(10 \times 1.5 + 10 \times 3)$

$\Rightarrow T \sim Po(45)$

since $\lambda > 10$, normal approx is appropriate

$\Rightarrow T \approx N(45, 45)$

$$P(T < 50) = P\left(Z < \frac{49.5 - 45}{\sqrt{45}}\right)$$

9

$= P(Z < 0.67)$

$= \underline{\underline{0.7486}}$

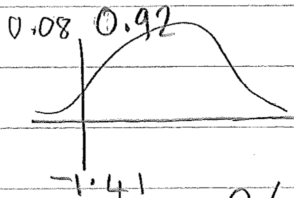
6) Let r.v X be weight of a tomato

$X \sim N(120, 300)$

a) $P(X \geq 135) = P\left(Z \geq \frac{135-120}{\sqrt{300}}\right)$

$= P(Z \geq 0.87)$

$= 0.1922 \Rightarrow \underline{\underline{19.2\%}}$

b) 

$P(Z < 1.41) = 0.92$

$\Rightarrow P(Z > -1.41) = 0.92$ 2 7

$P(X < a) = 0.08$

$\Rightarrow \frac{a-120}{\sqrt{300}} = -1.41$

$a = 95.6g.$

max weight of grade D tomato is $\underline{\underline{95.6g}}$

7) a) e.g. Police forces might cover areas with different numbers of offences and so might not be represented proportionally in the sample.

b) $\hat{p} = \frac{22}{50} = 0.44$ $\hat{p} \pm 1.64\sqrt{\frac{\hat{p}\hat{q}}{n}}$

$\hat{q} = 0.56$ $= 0.44 \pm 1.64 \times \sqrt{\frac{0.44 \times 0.56}{50}}$

$= 0.44 \pm 0.115$ 6

$\Rightarrow \underline{\underline{(0.325, 0.555)}}$

c) 31% lies outwith this interval so there is evidence to suggest the proportion has changed.

(8) let r.v. X be no of ryegrass seeds

$X \sim B(400, 0.35)$ 1

$np = 140 > 5$
$nq = 260 > 5$ normal approx appropriate

$\Rightarrow X \approx N(140, 91)$ 1

$\mu = np = 140$
$\sigma^2 = npq = 91$ 1

$P(120 \le X \le 150)$

$= P(119.5 \le X \le 150.5)$ 1

$= P\left(Z \le \frac{150.5 - 140}{\sqrt{91}}\right) - P\left(Z \le \frac{119.5 - 140}{\sqrt{91}}\right)$

1 6

$= P(Z \le 1.10) - P(Z \le -2.15)$

$= \underline{\underline{0.8486}}$ 1

(9) a) approx 49 years 1

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

$= 18477 - \frac{509^2}{15}$

$= 1204.93$

$b = \frac{S_{xy}}{S_{xx}}$

$= \frac{-1215.47}{1204.93}$

$= -1.009$ 1

$a = \bar{y} - b\bar{x}$

$= \frac{2738}{15} - (-1.009) \times \frac{509}{15}$

$= 216.7$ 1

$S_{xy} = \Sigma xy - \frac{\Sigma x \Sigma y}{n}$

$= 91694 - \frac{509 \times 2738}{15}$

$= -1215.47$

$\Rightarrow \underline{\underline{y = 216.7 - 1.009x}}$ 5

c) $\boxed{x = 45}$ n=16 $y = 209 - 0.727 \times 45$
$= 176.285$ 1

n=15 $y = 216.7 - 1.009 \times 45 = 171.295$

$\Rightarrow$ the volunteer would be predicted a lower average peak heart rate.

1

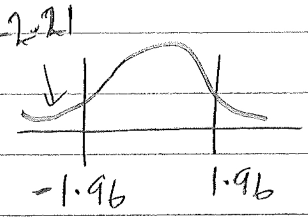
10 a) $\mu = 1015 \quad \sigma = 4 \quad n = 10 \quad \bar{x} = 1012.2$

$$\left.\begin{array}{l} H_0 : \mu = 1015 \\ H_1 : \mu \neq 1015 \end{array}\right] \begin{array}{l} \text{2-tail} \\ 0.05 \end{array}$$

(pvalue = 0.0271)

$$\bar{X} \sim N\left(1015, \frac{4^2}{10}\right)$$

$$Z = \frac{1012.2 - 1015}{4/\sqrt{10}}$$



$$= -2.21 < -1.96$$

C.V. $= -1.96$

$\Rightarrow$ reject H_0 there is significant evidence to suggest the mean volume has changed.

Assumption: standard deviation is unchanged.

10

b) 95% c.i. $\quad \bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}$

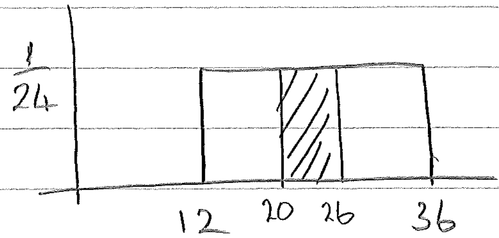
$$= 1012.2 \pm 1.96 \times \frac{4}{\sqrt{10}}$$

$$= 1012.2 \pm 2.48$$

$$\Rightarrow (1009.72, 1014.68)$$

since $\mu = 1015$ lies outwith the interval, this confirms the evidence that suggests the mean volume has changed.

11 $X \sim U(12, 36)$



$P(20 < X \leq 26)$

$= 6 \times \frac{1}{24}$

$= \frac{1}{4}$

(2)

12 a) e.g. * no sampling frame is available for the popn so random sampling methods cannot be used

(need 2 appropriate comments)

* respondents may be reluctant to be truthful about their snacking habits
* the amount of time needed is likely to be an issue.

2

b) advantage

e.g. * telephone surveys are relatively cheap & quick to do since there is no travel

(4)

disadvantage

e.g. * Using only respondents with access to a phone is potential source of bias since not everyone has a phone (landline / mobile).

Unacceptable

* better response rate by phone
* more likely to talk to you than in the street
* not everyone has a phone
* not everyone might be in when you phone

13 a) $X \sim N(46.3, 0.6^2)$

$S = X_1 + X_2 + X_3 + X_4$ $\quad E(S) = 4 \times 46.3 = \underline{\underline{185.2}}$ $\quad V(S) = 4 \times 0.6^2 = \underline{\underline{1.44}}$

$\Rightarrow SD(S) = 1.2$

b) $E \sim N(184.5, 1.3^2)$

$P(S < E) = P(S - E < 0)$

$E(S-E) = 185.2 - 184.5 = 0.7$

$\Rightarrow S - E \sim N(0.7, 3.13)$

$V(S-E) = 1.3^2 + 1.2^2 = 3.13$

$\Rightarrow P(S - E < 0)$

$= P\left(Z < \frac{0 - 0.7}{\sqrt{3.13}}\right)$

7

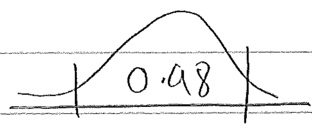
$= P(Z < -0.40)$

$= \underline{\underline{0.3446}}$

14 $X \sim N(136.5, 0.25^2)$

a) $P(136.5 - a < X < 136.5 + a) = 0.98$

$\Rightarrow P(X < 136.5 + a) = 0.99$



$\Rightarrow P\left(Z < \frac{136.5 + a - 136.5}{0.25}\right) = 0.99$

4

$\Rightarrow \frac{a}{0.25} = 2.33$

$a = 0.58$

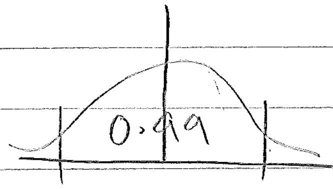
$\Rightarrow$ 98% of lengths lie between 135.92 cm and 137.08 cm.

b) $X \sim N(136.5, \sigma^2)$

$$P(136.2 < X < 136.8) = 0.99$$

$$\Rightarrow P\left(\frac{136.2 - 136.5}{\sigma} < Z < \frac{136.8 - 136.5}{\sigma}\right) = 0.99$$

$$\Rightarrow P\left(Z < \frac{0.3}{\sigma}\right) = 0.995$$



$$\frac{0.3}{\sigma} = 2.58$$

4

$$\underline{\underline{\sigma = 0.116 \text{ cm}}}$$

15 a) taking $\bar{p} = 0.6807$ for the first 60 days as p:

Control limits are $p \pm 3\sqrt{\frac{pq}{n}}$

$$= 0.6807 \pm 3\sqrt{\frac{0.6807 \times 0.3193}{50}}$$

$$= \underline{\underline{(0.4829, 0.8785)}}$$

b) The fact that a number of points lie above the upper limit provides evidence that the proportion of units free from non conformities has increased.

c) $p = 0.8648$

$$\Rightarrow UCL = 0.8648 + 3\sqrt{\frac{0.8648 \times 0.1352}{50}}$$

$$= 1.01$$

A proportion cannot exceed 1.

d) $p + 3\sqrt{\frac{pq}{n}} < 1$

$\Rightarrow \quad 3\sqrt{\frac{pq}{n}} < 1 - p$

$3\sqrt{\frac{pq}{n}} < q$

9

$9\frac{pq}{n} < q^2$

$n > 9\frac{p}{q}$

$n > 9 \times \frac{0.8648}{0.1352} = 57.6$

$\Rightarrow$ a sample size of at least 58 is required.