

Sampling Distributions

1 (2 marks)

(a) It describes the whole population, so it is a parameter

B1

Answer: Parameter

1 mark

(b) It is calculated from the sample, so it is a statistic

B1

Answer: Statistic

1 mark

2 (3 marks)

$$E(\bar{X}) = 50$$

B1

$$\text{Var}(\bar{X}) = \frac{36}{9} = 4$$

M1

$$\text{Standard error} = 2$$

A1

Answer: $E(\bar{X}) = 50$, standard error = 2

3 marks

3 (2 marks)

The sampling distribution is the probability distribution of the statistic over all possible random samples of a given size from the population

B1

B1

Answer: Distribution of the statistic over all possible samples

2 marks

4 (2 marks)

$$E(\bar{X}) = 20, \text{Var}(\bar{X}) = \frac{16}{4} = 4$$

M1

$$\bar{X} \sim N(20, 2^2)$$

A1

Answer: $N(20, 2^2)$

2 marks

5 (2 marks)

$$E(\bar{X}) = 12$$

B1

$$\text{Standard error} = \frac{3}{\sqrt{25}} = 0.6$$

B1

Answer: 12, 0.6

2 marks

6 (2 marks)

A census is expensive and time-consuming

B1

Testing may destroy the items, or the population may be infinite/too large to reach

B1

Answer: Cost/time; destructive testing or large population

2 marks

7 (6 marks)

(a) Samples: (1,1) (1,3) (1,5) (3,1) (3,3) (3,5) (5,1) (5,3) (5,5) with means 1, 2, 3, 2, 3, 4, 3, 4, 5

M1 A1

$$P(\bar{X} = 1) = \frac{1}{9}, P(\bar{X} = 2) = \frac{2}{9}, P(\bar{X} = 3) = \frac{3}{9}$$

M1

$$P(\bar{X} = 4) = \frac{2}{9}, P(\bar{X} = 5) = \frac{1}{9}$$

A1

Answer: 1: $\frac{1}{9}$, 2: $\frac{2}{9}$, 3: $\frac{3}{9}$, 4: $\frac{2}{9}$, 5: $\frac{1}{9}$

4 marks

$$(b) E(\bar{X}) = \frac{1 + 4 + 9 + 8 + 5}{9} = \frac{27}{9}$$

M1

$$= 3 = \text{population mean } (1 + 3 + 5)/3$$

A1

Answer: $E(\bar{X}) = 3$

2 marks

8 (4 marks)

The mode is 10 if at least two of the three tiles are 10 cm wide

B1

$$P(M = 10) = 3 \times 0.4^2 \times 0.6 + 0.4^3$$

M1

$$= 0.288 + 0.064 = 0.352$$

A1

$$P(M = 20) = 1 - 0.352 = 0.648$$

A1

Answer: 10: 0.352, 20: 0.648

4 marks

9 (4 marks)

$$E(\hat{p}) = p = 0.3$$

B1

$$\text{Var}(\hat{p}) = \frac{p(1-p)}{n} = \frac{0.3 \times 0.7}{50}$$

M1

$$= 0.0042$$

A1

$$\text{Standard error} = \sqrt{0.0042} \approx 0.0648$$

A1

Answer: 0.3, 0.0042, 0.0648

4 marks

10 (3 marks)

$$\bar{X} \sim N(100, \frac{144}{9}) = N(100, 4^2)$$

M1

$$P(\bar{X} > 104) = P(Z > \frac{104 - 100}{4}) = 1 - \Phi(1)$$

M1

$$= 0.1587$$

A1

Answer: 0.1587

3 marks

11 (4 marks)

$$E(T) = 16 \times 5 = 80$$

B1

$$\text{Var}(T) = 16 \times 4 = 64$$

M1

$$P(T < 70) = P(Z < \frac{70 - 80}{8}) = \Phi(-1.25)$$

M1

$$= 0.1056$$

A1

Answer: 0.1056

4 marks

12 (3 marks)

$$\text{Standard error} = \frac{\sigma}{\sqrt{n}}; \text{ replacing } n \text{ by } 4n \text{ gives } \frac{\sigma}{2\sqrt{n}}$$

M1

So the standard error is halved

A1

The sampling distribution of $\bar{X}$ is less spread out, so the sample mean is likely to be closer to μ

E1

Answer: Halved

3 marks

13 (5 marks)

16 equally likely samples

B1

R = 0: 4 samples; R = 1: 6 samples

M1 A1

R = 2: 4 samples; R = 3: 2 samples

A1

$$E(R) = \frac{0 + 6 + 8 + 6}{16} = \frac{20}{16} = \frac{5}{4}$$

A1

$$\text{Answer: } P(R = 0, 1, 2, 3) = \frac{4}{16}, \frac{6}{16}, \frac{4}{16}, \frac{2}{16}; E(R) = \frac{5}{4}$$

5 marks

14 (7 marks)

There are 10 equally likely samples

B1

Means: 1.5, 2, 2.5, 2.5, 3, 3, 3.5, 3.5, 4, 4.5

M1

$$P(\bar{X} = 1.5, 2, 2.5, 3, 3.5, 4, 4.5) = \frac{1}{10}, \frac{1}{10}, \frac{2}{10}, \frac{2}{10}, \frac{2}{10}, \frac{1}{10}, \frac{1}{10}$$

A1

$$E(\bar{X}) = 3 \text{ and } E(\bar{X}^2) = \frac{1}{10}(2.25 + 4 + 12.5 + 18 + 24.5 + 16 + 20.25) = 9.75$$

M1 A1

$$\text{Var}(\bar{X}) = 9.75 - 9 = 0.75$$

A1

$$\text{With replacement } \text{Var}(\bar{X}) = \frac{\sigma^2}{n} = \frac{2}{2} = 1, \text{ which is larger}$$

A1

$$\text{Answer: } \text{Var}(\bar{X}) = \frac{3}{4} < 1$$

7 marks

15 (5 marks)

$M > 0.5$ requires all three values to exceed 0.5

M1

$$P(M > 0.5) = (0.5)^3 = \frac{1}{8}$$

A1

$$P(M > m) = (1 - m)^3$$

M1

$$G(m) = P(M \leq m) = 1 - (1 - m)^3$$

A1

for $0 \leq m \leq 1$

A1

Answer: $\frac{1}{8}; G(m) = 1 - (1 - m)^3$

5 marks

16 (6 marks)

$\hat{p} \geq 0.45$ means at least 18 of the 40 support it

B1

$$X \sim B(40, 0.35): P(X \geq 18) = 1 - P(X \leq 17)$$

M1

$$= 0.124$$

A1

Approximately $N(14, 9.1)$

B1

$$P(X \geq 18) \approx P(Y > 17.5) = P(Z > \frac{3.5}{\sqrt{9.1}}) = P(Z > 1.160)$$

M1

$$\approx 0.123$$

A1

Answer: Exact 0.124; approx 0.123

6 marks

17 (6 marks)

$$\bar{X}_A \sim N(50, 2.5), \bar{X}_B \sim N(48, 2.4)$$

M1 A1

$$D = \bar{X}_A - \bar{X}_B: E(D) = 2$$

B1

$$\text{Var}(D) = 2.5 + 2.4 = 4.9$$

M1

$$P(D > 0) = P(Z > \frac{-2}{\sqrt{4.9}}) = \Phi(0.9035)$$

M1

$$= 0.817$$

A1

Answer: 0.817

6 marks

18 (5 marks)

$$\bar{X} = \frac{1}{n}(X_1 + X_2 + \dots + X_n)$$

B1

$$E(\bar{X}) = \frac{1}{n} \times n\mu = \mu$$

M1 A1

By independence the variances add: $\text{Var}(\bar{X}) = \frac{1}{n^2} \text{Var}(X_1 + \dots + X_n) = \frac{1}{n^2} \times n\sigma^2$

M1

$$= \frac{\sigma^2}{n}$$

A1

Answer: $E(\bar{X}) = \mu, \text{Var}(\bar{X}) = \frac{\sigma^2}{n}$

5 marks

19 (5 marks)

$$\text{Var}(\bar{X}) = \frac{100}{25} = 4, \text{ which is the variance, not the standard error}$$

E1

$$\text{Standard error} = \sqrt{4} = 2$$

A1

$$\text{Halving gives } 1 = \frac{10}{\sqrt{n}}$$

M1

$$\sqrt{n} = 10$$

M1

$$n = 100$$

A1

Answer: Standard error 2; n = 100

5 marks

20 (7 marks)

(a) $P(2p) = \frac{3}{4}$ and $P(10p) = \frac{1}{4}$

B1

(2, 2) gives $Y = 2$ with probability $\frac{9}{16}$; (2, 10) gives $Y = 8$ with probability $\frac{3}{16}$

M1

(10, 2) gives $Y = 4$ with probability $\frac{3}{16}$; (10, 10) gives $Y = 10$ with probability $\frac{1}{16}$

A1

$Y = 2, 4, 8, 10$ with probabilities $\frac{9}{16}, \frac{3}{16}, \frac{3}{16}, \frac{1}{16}$

A1

Answer: 2: $\frac{9}{16}$, 4: $\frac{3}{16}$, 8: $\frac{3}{16}$, 10: $\frac{1}{16}$

4 marks

(b) $E(Y) = \frac{18 + 12 + 24 + 10}{16} = 4$

M1

$E(X) = 2 \times \frac{3}{4} + 10 \times \frac{1}{4} = 4$, so Y is unbiased

A1

$\text{Var}(X) = 4 \times \frac{3}{4} + 100 \times \frac{1}{4} - 16 = 12$, so $\text{Var}(Y) = \frac{1+9}{16} \times 12 = 7.5$

A1

Answer: $\text{Var}(Y) = 7.5$

3 marks

21 (7 marks)

There are $4^3 = 64$ equally likely samples

B1

$M = 1$: at least two 1s: $3 \times 3 + 1 = 10$ samples

M1 A1

$M = 2$: count samples with median 2

M1

Counts: $M = 1$: 10, $M = 2$: 22, $M = 3$: 22, $M = 4$: 10 (total 64)

A1

$E(M) = \frac{10 + 2 \times 22 + 3 \times 22 + 4 \times 10}{64}$

M1

$= \frac{160}{64} = \frac{5}{2}$

A1

Answer: $P(M = 1, 2, 3, 4) = \frac{10}{64}, \frac{22}{64}, \frac{22}{64}, \frac{10}{64}$; $E(M) = \frac{5}{2}$

7 marks

22 (6 marks)

$\bar{X} \sim N(\mu, \frac{64}{n})$, so the standard error is $\frac{8}{\sqrt{n}}$

B1

$P(|\bar{X} - \mu| < 1) = P(|Z| < \frac{\sqrt{n}}{8})$

M1

Need $\frac{\sqrt{n}}{8} \geq 1.96$

M1 A1

$\sqrt{n} \geq 15.68$, $n \geq 245.9$

M1

$n = 246$

A1

Answer: $n = 246$

6 marks

23 (7 marks)

(a) $M \leq m$ requires both values $\leq m$: $P(M \leq m) = (\frac{m}{5})^2 = \frac{m^2}{25}$

M1 A1

$P(M = m) = \frac{m^2 - (m-1)^2}{25} = \frac{2m-1}{25}$

A1

Answer: $P(M = m) = \frac{2m-1}{25}$, $m = 1, \dots, 5$

3 marks

(b) $E(M) = \frac{1}{25}(1 + 6 + 15 + 28 + 45)$

M1 A1

$= \frac{95}{25} = \frac{19}{5} = 3.8$, which is less than the population maximum 5

A1

M can never exceed 5 but is often smaller, so it underestimates on average

E1

Answer: $E(M) = \frac{19}{5}$

4 marks

24 (6 marks)

For two values a and b , $S^2 = \frac{(a - b)^2}{2}$

B1

$a = b$ (3 samples): $S^2 = 0$

M1

$|a - b| = 2$ (4 samples): $S^2 = 2$; $|a - b| = 4$ (2 samples): $S^2 = 8$

A1

$P(S^2 = 0, 2, 8) = \frac{3}{9}, \frac{4}{9}, \frac{2}{9}$

A1

$E(S^2) = 0 + \frac{8}{9} + \frac{16}{9}$

M1

$= \frac{24}{9} = \frac{8}{3}$, equal to the population variance

A1

Answer: $E(S^2) = \frac{8}{3}$

6 marks

25 (7 marks)

Total of all observations = $10\bar{X}_1 + 40\bar{X}_2$

M1

$C = \frac{10\bar{X}_1 + 40\bar{X}_2}{50} = \frac{\bar{X}_1 + 4\bar{X}_2}{5}$

A1

$E(C) = \frac{\mu + 4\mu}{5} = \mu$

A1

$\text{Var}(C) = \frac{1}{25}(\text{Var } \bar{X}_1 + 16 \text{Var } \bar{X}_2) = \frac{1}{25}\left(\frac{\sigma^2}{10} + \frac{16\sigma^2}{40}\right)$

M1 A1

$= \frac{1}{25} \times \frac{\sigma^2}{2} = \frac{\sigma^2}{50}$

A1

$\text{Var}(\bar{X}_2) = \frac{\sigma^2}{40} > \frac{\sigma^2}{50}$: the combined mean is more precise since it uses more data

E1

Answer: $E(C) = \mu$, $\text{Var}(C) = \frac{\sigma^2}{50}$

7 marks

26 (8 marks)

(a) Possible means: 0, 1, 2, 3, 4

B1

$P(\bar{X} = 0) = 0.5^2 = 0.25$

M1

$P(\bar{X} = 1) = 2 \times 0.5 \times 0.3 = 0.30$, $P(\bar{X} = 2) = 2 \times 0.5 \times 0.2 + 0.3^2 = 0.29$

A1

$P(\bar{X} = 3) = 2 \times 0.3 \times 0.2 = 0.12$

A1

$P(\bar{X} = 4) = 0.2^2 = 0.04$

A1

Answer: 0: 0.25, 1: 0.30, 2: 0.29, 3: 0.12, 4: 0.04

5 marks

(b) $E(X) = 0.6 + 0.8 = 1.4$, $E(X^2) = 1.2 + 3.2 = 4.4$, $\text{Var}(X) = 4.4 - 1.96 = 2.44$

M1 A1

$E(\bar{X}^2) = 0.30 + 1.16 + 1.08 + 0.64 = 3.18$, $\text{Var}(\bar{X}) = 3.18 - 1.96 = 1.22 = \frac{1}{2} \times 2.44$

A1

Answer: $\text{Var}(\bar{X}) = 1.22$

3 marks

27 (9 marks)

(a) $P(M \leq m) = \left(\frac{m}{\theta}\right)^n$

M1 A1

Differentiating: $g(m) = \frac{nm^{n-1}}{\theta^n}$

A1

Answer: $g(m) = \frac{nm^{n-1}}{\theta^n}$, $0 \leq m \leq \theta$

3 marks

(b) $E(M) = \int_0^\theta m \times \frac{nm^{n-1}}{\theta^n} dm = \frac{n}{\theta^n} \left[\frac{m^{n+1}}{n+1} \right]_0^\theta$

M1 A1

$= \frac{n\theta}{n+1}$

A1

Answer: $\frac{n\theta}{n+1}$

3 marks

(c) $P(M > 0.95\theta) = 1 - (0.95)^n > 0.9$

M1

$(0.95)^n < 0.1$, $n > \frac{\ln 0.1}{\ln 0.95} \approx 44.9$

M1

$n = 45$

A1

Answer: $n = 45$ **3 marks****28 (11 marks)**

$$(a) f_S(s) = \int_0^s \lambda e^{-\lambda u} \times \lambda e^{-\lambda(s-u)} du$$

M1 A1

$$= \lambda^2 e^{-\lambda s} \int_0^s du$$

M1

$$= \lambda^2 s e^{-\lambda s}$$

A1**Answer: $f_S(s) = \lambda^2 s e^{-\lambda s}$, $s > 0$** **4 marks**

$$(b) P(\bar{X} \leq x) = P(S \leq 2x)$$

M1

$$= \int_0^{2x} \lambda^2 s e^{-\lambda s} ds = [-\lambda s e^{-\lambda s} - e^{-\lambda s}]_0^{2x}$$

M1 A1

$$= 1 - (1 + 2\lambda x)e^{-2\lambda x}$$

A1**Answer: $G(x) = 1 - (1 + 2\lambda x)e^{-2\lambda x}$, $x \geq 0$** **4 marks**

$$(c) P(\bar{X} < 1) = 1 - 3e^{-2}$$

M1

$$= 0.594$$

A1

$P(X < 1) = 1 - e^{-1} = 0.632$: the sample mean is more likely to be small, but unlike the individual values its distribution is more concentrated about the mean 1

E1**Answer: 0.594****3 marks****29 (8 marks)**

$$(a) \mu = \frac{34}{6} = \frac{17}{3}$$

B1

$$\Sigma x^2 = 4 + 16 + 16 + 36 + 64 + 100 = 236$$

M1

$$\sigma^2 = \frac{236}{6} - \left(\frac{17}{3}\right)^2 = \frac{65}{9}$$

A1**Answer: $\mu = \frac{17}{3}$, $\sigma^2 = \frac{65}{9}$** **3 marks****(b) $E(\bar{X}) = \mu$ since each value appears equally often****B1**

$$E(\bar{X}^2) = \frac{1}{20} \times \frac{6040}{9} = \frac{302}{9}$$

M1

$$\text{Var}(\bar{X}) = \frac{302}{9} - \frac{289}{9} = \frac{13}{9}$$

A1

$$\frac{\sigma^2}{n} \times \frac{N-n}{N-1} = \frac{65}{27} \times \frac{3}{5}$$

M1

$$= \frac{13}{9}, \text{ agreeing}$$

A1**Answer: $\text{Var}(\bar{X}) = \frac{13}{9}$** **5 marks****30 (10 marks)****(a) $P(R \leq r)$ = area of the region $|x_1 - x_2| \leq r$ in the unit square****M1**The excluded region is two right-angled triangles with legs $1 - r$ **M1 A1**

$$\text{Area} = 1 - (1 - r)^2 = 2r - r^2$$

A1**Answer: $G(r) = 2r - r^2$** **4 marks**

$$(b) g(r) = G'(r) = 2 - 2r = 2(1 - r)$$

M1 A1

$$E(R) = \int_0^1 2r(1 - r) dr = 2\left(\frac{1}{2} - \frac{1}{3}\right)$$

M1

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

A1**Answer: $g(r) = 2(1 - r)$, $E(R) = \frac{1}{3}$** **4 marks**

$$(c) P(R > 0.8) = (1 - 0.8)^2 = 0.04$$

B1

$$2m - m^2 = \frac{1}{2}, m = 1 - \frac{1}{\sqrt{2}} \approx 0.293$$

B1**Answer: 0.04; 0.293****2 marks**

Central Limit Theorem

1 (2 marks)

For a random sample of size n from any population with mean μ and finite variance σ^2 ,

B1

$\bar{X}$ is approximately $N(\mu, \frac{\sigma^2}{n})$ when n is large

B1

Answer: $\bar{X} \approx N(\mu, \sigma^2/n)$ for large n

2 marks

2 (2 marks)

$$E(\bar{X}) = 10, \text{Var}(\bar{X}) = \frac{16}{64} = 0.25$$

M1

$$\bar{X} \approx N(10, 0.25)$$

A1

Answer: $N(10, 0.25)$

2 marks

3 (2 marks)

$$E(T) = 500, \text{Var}(T) = 100 \times 4 = 400$$

M1

$$T \approx N(500, 400)$$

A1

Answer: $N(500, 400)$

2 marks

4 (2 marks)

By the central limit theorem the distribution of $\bar{X}$ is approximately normal for large n
 $n = 50$ is large, whatever the shape of the parent distribution

B1

B1

Answer: Central limit theorem

2 marks

5 (3 marks)

$$\bar{X} \approx N(20, \frac{36}{36}) = N(20, 1)$$

M1

$$P(\bar{X} > 21) = P(Z > 1)$$

M1

$$= 0.1587$$

A1

Answer: 0.1587

3 marks

6 (3 marks)

$$E(X) = 3$$

B1

$$\text{Var}(X) = \frac{36}{12} = 3, \text{ so } \text{Var}(\bar{X}) = \frac{3}{48} = \frac{1}{16}$$

M1

$$\bar{X} \approx N(3, \frac{1}{16})$$

A1

Answer: $N(3, \frac{1}{16})$

3 marks

7 (4 marks)

$\text{Exp}(0.5)$ has mean 2 and standard deviation 2

B1

$$\bar{X} \approx N(2, \frac{4}{50}), \text{ standard error } \frac{2}{\sqrt{50}} = 0.2828$$

M1

$$P(\bar{X} > 2.4) = P(Z > 1.414)$$

M1

$$\approx 0.0786$$

A1

Answer: ≈ 0.0786

4 marks

8 (4 marks)

$$T \approx N(40 \times 12.5, 40 \times 2.25) = N(500, 90)$$

M1 A1

$$P(T > 510) = P(Z > \frac{10}{\sqrt{90}}) = P(Z > 1.054)$$

M1

$$\approx 0.146$$

A1

Answer: ≈ 0.146

4 marks

9 (4 marks)

$$E(X) = 4, \text{Var}(X) = 4$$

B1

$$\bar{X} \approx N(4, \frac{4}{60})$$

M1

$$P(\bar{X} < 3.8) = P(Z < \frac{-0.2}{0.2582}) = P(Z < -0.775)$$

M1

$$\approx 0.219$$

A1

Answer: ≈ 0.219

4 marks

10 (5 marks)

$$\text{For one roll } E(X) = 3.5, \text{Var}(X) = \frac{35}{12}$$

B1

$$\text{Total} \approx N(105, 87.5)$$

M1 A1

$$P(T > 115) = P(T \geq 116) \approx P(Y > 115.5)$$

M1

$$z = \frac{10.5}{\sqrt{87.5}} = 1.122, \text{probability} \approx 0.131$$

A1

Answer: ≈ 0.131

5 marks

11 (4 marks)

$$\hat{p} \approx N(0.2, \frac{0.2 \times 0.8}{100}) = N(0.2, 0.04^2)$$

M1 A1

$$P(\hat{p} > 0.27) = P(Z > \frac{0.07}{0.04}) = P(Z > 1.75)$$

M1

$$= 0.0401$$

A1

Answer: 0.0401

4 marks

12 (5 marks)

$$E(X) = 200 \times 0.35 = 70$$

B1

$$\text{Var}(X) = 200 \times 0.35 \times 0.65 = 45.5$$

M1

$$\bar{X} \approx N(70, \frac{45.5}{50}) = N(70, 0.91)$$

M1

$$P(\bar{X} > 71) = P(Z > \frac{71 - 70}{\sqrt{0.91}}) = P(Z > 1.048)$$

M1

$$= 0.1473$$

A1

Answer: 0.1473

5 marks

13 (5 marks)

$$\bar{X} \approx N(\mu, \frac{9}{n})$$

B1

$$P(|\bar{X} - \mu| < 0.5) = P(|Z| < \frac{0.5\sqrt{n}}{3}) \geq 0.95$$

M1

$$\frac{0.5\sqrt{n}}{3} \geq 1.96$$

M1 A1

$$\sqrt{n} \geq 11.76, n \geq 138.3, \text{so } n = 139$$

A1

Answer: $n = 139$

5 marks

14 (6 marks)

$$E(X) = 0.15 + 0.5 + 0.6 + 1.2 + 0.5 = 2.95$$

M1 A1

$$E(X^2) = 0.15 + 1 + 1.8 + 4.8 + 2.5 = 10.25, \text{Var}(X) = 10.25 - 2.95^2 = 1.5475$$

M1 A1

$$\bar{X} \approx N(2.95, \frac{1.5475}{60})$$

M1

$$P(\bar{X} > 3.2) = P(Z > \frac{0.25}{0.1606}) = P(Z > 1.557) \approx 0.0598$$

A1

Answer: ≈ 0.0598

6 marks

15 (5 marks)

Each lifetime has mean 30 and variance 900

B1

Total $T \approx N(1500, 45000)$

M1 A1

$$P(T > 1600) = P\left(Z > \frac{100}{\sqrt{45000}}\right) = P(Z > 0.471)$$

M1

$$\approx 0.319$$

A1

Answer: ≈ 0.319

5 marks

16 (6 marks)

$$\bar{X}_A \approx N\left(20, \frac{25}{40}\right), \bar{X}_B \approx N\left(19, \frac{16}{50}\right)$$

M1 A1

$$D = \bar{X}_A - \bar{X}_B \approx N(1, 0.625 + 0.32) = N(1, 0.945)$$

M1 A1

$$P(D > 0) = P\left(Z > \frac{-1}{\sqrt{0.945}}\right) = P(Z > -1.029)$$

M1

$$\approx 0.848$$

A1

Answer: ≈ 0.848

6 marks

17 (5 marks)Total $T \sim \text{Po}(120) \approx N(120, 120)$

M1 A1

$$P(T \geq 130) \approx P(Y > 129.5)$$

M1

$$z = \frac{9.5}{\sqrt{120}} = 0.867$$

A1

$$P \approx 0.193$$

A1

Answer: ≈ 0.193

5 marks

18 (5 marks)

$$E(X) = \frac{1}{0.4} = 2.5$$

B1

$$\text{Var}(X) = \frac{1 - 0.4}{0.4^2} = 3.75$$

M1 A1

$$\bar{X} \approx N\left(2.5, \frac{3.75}{120}\right) = N(2.5, 0.03125)$$

M1

$$P(\bar{X} > 2.8) = P\left(Z > \frac{2.8 - 2.5}{\sqrt{0.03125}}\right) = P(Z > 1.697) = 0.0448$$

A1

Answer: 0.0448

5 marks

19 (6 marks)

$$(a) E(X) = 5.5 \text{ and } \text{Var}(X) = \frac{10^2 - 1}{12} = 8.25$$

M1 A1

$$\bar{X} \approx N\left(5.5, \frac{8.25}{50}\right) = N(5.5, 0.165)$$

M1

$$\frac{k - 5.5}{\sqrt{0.165}} = -1.6449$$

M1

$$k = 4.83$$

A1

Answer: 4.83

5 marks

(b) The score on one roll is not normally distributed, but with $n = 50$ the sample mean is approximately normally distributed

E1

Answer: The sample mean is approximately normal for large n

1 mark

20 (5 marks)

The central limit theorem applies to the sample mean (or total), not to individual observations

E1

The shape of the distribution of individual rods is unchanged

E1

$$\bar{X} \approx N\left(40, \frac{9}{100}\right), \text{ standard error } 0.3$$

M1

$$P(\bar{X} > 40.5) = P(Z > 1.667)$$

M1

$$= 0.0478$$

A1

Answer: 0.0478

5 marks

21 (6 marks)

X has a negative binomial distribution with $r = 3$ and $p = \frac{1}{6}$

B1

$$E(X) = \frac{r}{p} = 18 \text{ and } \text{Var}(X) = \frac{r(1-p)}{p^2} = 3 \times \frac{5}{6} \times 36 = 90$$

M1 A1

$$\bar{X} \approx N(18, \frac{90}{40}) = N(18, 2.25)$$

M1

$$P(\bar{X} < 17.5) = P(Z < \frac{17.5 - 18}{1.5}) = \Phi(-0.333)$$

M1

$$= 0.3694$$

A1

Answer: 0.3694

6 marks

22 (7 marks)

$$E(X) = \int_0^2 \frac{3x^3}{8} dx = \frac{3}{2}$$

M1 A1

$$E(X^2) = \int_0^2 \frac{3x^4}{8} dx = \frac{12}{5}$$

M1 A1

$$\text{Var}(X) = \frac{12}{5} - \frac{9}{4} = \frac{3}{20}$$

A1

$$\bar{X} \approx N(\frac{3}{2}, \frac{3}{1000})$$

M1

$$P(\bar{X} > 1.6) = P(Z > \frac{0.1}{0.05477}) = P(Z > 1.826) \approx 0.0339$$

A1

Answer: ≈ 0.0339

7 marks

23 (6 marks)

$$E(\text{total}) = 150 + 160 = 310$$

B1

$$\text{Var}(\text{total}) = 30 \times 1 + 20 \times 4 = 110$$

M1 A1

Total $\approx N(310, 110)$ (large number of items)

M1

$$P(\text{total} > 330) = P(Z > \frac{20}{\sqrt{110}}) = P(Z > 1.907)$$

M1

$$\approx 0.0283$$

A1

Answer: ≈ 0.0283

6 marks

24 (7 marks)

Exp: mean 2, standard deviation 2

B1

$$\bar{X} \approx N(2, \frac{4}{50}), P(\bar{X} > 2.1) = P(Z > \frac{0.1}{0.2828}) = P(Z > 0.354)$$

M1 A1

$$p \approx 0.362$$

A1

$$Y \sim B(5, p), P(Y \geq 2) = 1 - P(Y = 0) - P(Y = 1)$$

M1

$$= 1 - (1 - p)^5 - 5p(1 - p)^4$$

M1

$$\approx 0.594$$

A1

Answer: ≈ 0.594

7 marks

25 (6 marks)

$$\hat{p} \approx N(0.3, \frac{0.21}{150}), \text{ standard error } 0.03742$$

M1 A1

$$P(\hat{p} > 0.38) = P(Z > 2.138)$$

M1

$$\approx 0.0163$$

A1

$$\frac{\sqrt{0.21}}{\sqrt{n}} < 0.03, \text{ so } n > \frac{0.21}{0.0009} = 233.3$$

M1

$$n = 234$$

A1

Answer: ≈ 0.0163 ; $n = 234$

6 marks

26 (7 marks)

(a) $E(X) = 1, \text{Var}(X) = 1$

B1

$\bar{X} \approx N(1, \frac{1}{30})$

M1

$P(\bar{X} > 1.3) = P(Z > \frac{0.3}{0.1826}) = P(Z > 1.643)$

M1

≈ 0.0502

A1

Answer: ≈ 0.0502

4 marks

(b) $\text{Error} = \frac{0.0502 - 0.0591}{0.0591} \times 100$

M1

$\approx -15\%$ (an underestimate)

A1

The exponential distribution is strongly skewed and $n = 30$ is only moderate, so the sample mean is still slightly skewed

E1

Answer: -15%

3 marks

27 (8 marks)

$T \approx N(12 \times 72, 12 \times 196) = N(864, 2352)$

M1 A1

$P(T > 1000) = P(Z > \frac{136}{\sqrt{2352}}) = P(Z > 2.804)$

M1

≈ 0.00252

A1

For n passengers need $\frac{1000 - 72n}{14\sqrt{n}} \geq 1.645$

M1 A1

$n = 12: \frac{136}{48.5} = 2.80 \geq 1.645$ satisfied; $n = 13: \frac{64}{50.5} = 1.27 < 1.645$ fails

M1

Largest $n = 12$

A1

Answer: 0.00252; $n = 12$

8 marks

28 (8 marks)

$\bar{X} \approx N(50, \frac{100}{n}), \text{standard error } \frac{10}{\sqrt{n}}$

B1

$P(|\bar{X} - 50| > 2) = P(|Z| > \frac{2\sqrt{n}}{10}) = P(|Z| > 0.2\sqrt{n})$

M1 A1

$= 2[1 - \Phi(0.2\sqrt{n})]$

A1

Need $\Phi(0.2\sqrt{n}) > 0.995$

M1

$0.2\sqrt{n} > 2.576$

A1

$\sqrt{n} > 12.88, n > 165.9$

M1

$n = 166$

A1

Answer: $n = 166$

8 marks

29 (10 marks)

Mean and standard deviation of one lifetime are both 100

B1

$T \approx N(100n, 10000n)$

M1 A1

$P(T \geq 5000) \geq 0.95$ requires $\frac{5000 - 100n}{100\sqrt{n}} \leq -1.645$

M1 A1

$n - 1.645\sqrt{n} - 50 \geq 0$

A1

Quadratic in $\sqrt{n}$: $\sqrt{n} \geq \frac{1.645 + \sqrt{1.645^2 + 200}}{2} = 7.94$

M1

$n \geq 63.06$

A1

Check: $n = 63$ gives $63 - 1.645\sqrt{63} = 49.94 < 50$ (fails); $n = 64$ gives $64 - 13.16 = 50.84 \geq 50$

M1

$n = 64$

A1

Answer: $n = 64$

10 marks

30 (9 marks)

$$(a) E(X) = \frac{1}{2}, \text{Var}(X) = \frac{1}{12}, \text{ so } \bar{X} \approx N\left(\frac{1}{2}, \frac{1}{12n}\right)$$

B1 M1

$$P(0.45 < \bar{X} < 0.55) = P(|Z| < \frac{0.05\sqrt{n}}{0.2887}) \geq 0.9$$

M1

$$\frac{0.05\sqrt{n}}{0.2887} \geq 1.645, \sqrt{n} \geq 9.50$$

A1

$$n \geq 90.3, \text{ so } n = 91$$

A1**Answer: n = 91****5 marks**

$$(b) E(T) = 12 \times \frac{1}{2} - 6 = 0$$

B1

$$\text{Var}(T) = 12 \times \frac{1}{12} = 1$$

M1 A1

By the central limit theorem T is approximately $N(0, 1)$, so adding 12 uniform random numbers and subtracting 6 gives approximately standard normal values

E1**Answer: T $\approx$ N(0, 1)****4 marks****Point Estimation****1 (2 marks)**

An estimator is a rule, or a random variable, calculated from the sample (e.g. $\bar{X}$)

B1

An estimate is the numerical value the estimator takes for one observed sample

B1**Answer: Estimator = rule/random variable; estimate = observed value****2 marks****2 (3 marks)**

$$\Sigma x = 35, \Sigma x^2 = 271; \bar{x} = 7$$

B1

$$s^2 = \frac{1}{n-1}(\Sigma x^2 - n\bar{x}^2) = \frac{1}{4}(271 - 245)$$

M1

$$= 6.5$$

A1

$$\text{Answer: } \bar{x} = 7, s^2 = 6.5$$

3 marks**3 (2 marks)**

$$\bar{x} = \frac{\Sigma x}{n} = \frac{240}{30}$$

M1

$$= 8$$

A1**Answer: 8****2 marks****4 (2 marks)**

$$\hat{p} = \frac{18}{60}$$

M1

$$= 0.3$$

A1**Answer: 0.3****2 marks****5 (2 marks)**

The parameter λ is the mean, estimated by $\bar{x}$

M1

$$\hat{\lambda} = \frac{15}{5} = 3$$

A1**Answer: 3****2 marks****6 (2 marks)**

$$E(X) = \frac{1}{\lambda}, \text{ so estimate } \frac{1}{\hat{\lambda}} = 4$$

M1

$$\hat{\lambda} = 0.25$$

A1**Answer: 0.25****2 marks**

7 (4 marks)

$$\bar{x} = \frac{1640}{40} = 41$$

B1

$$s^2 = \frac{1}{39}(70000 - 40 \times 41^2)$$

M1

$$= \frac{1}{39}(70000 - 67240)$$

A1

$$= 70.8$$

A1

Answer: $\bar{x} = 41$, $s^2 = 70.8$

4 marks

8 (5 marks)

$$\bar{y} = \frac{75}{25} = 3, \text{ so } \bar{x} = 103$$

M1 A1

$$s_y^2 = \frac{1}{24}(1000 - 25 \times 9)$$

M1

$$= \frac{775}{24} = 32.3$$

A1

Coding by subtraction does not change the variance, so $s_x^2 = 32.3$

A1

Answer: $\bar{x} = 103$, $s^2 = 32.3$

5 marks

9 (5 marks)

$$\Sigma fx = 0 + 20 + 22 + 15 + 8 = 65$$

M1

$$\bar{x} = \frac{65}{50} = 1.3$$

A1

$$\hat{\lambda} = 1.3$$

B1

$$P(X = 0) \approx e^{-1.3}$$

M1

$$= 0.273$$

A1

Answer: 0.273

5 marks

10 (4 marks)

$$\bar{x} = \frac{20.4}{5} = 4.08$$

B1

$$E(X) = \frac{\theta}{2}, \text{ so } \theta = 2\bar{x}$$

M1

$$\theta = 8.16$$

A1

This exceeds the largest observation 5.1, as it must since $\theta \geq 5.1$

E1

Answer: $\theta = 8.16$

4 marks

11 (3 marks)

$$SE = \frac{s}{\sqrt{n}}$$

M1

$$s = \sqrt{12.96} = 3.6$$

A1

$$SE = \frac{3.6}{6} = 0.6$$

A1

Answer: 0.6

3 marks

12 (3 marks)

$$\text{Combined total} = 20 \times 15 + 30 \times 16 = 300 + 480 = 780$$

M1 A1

$$\text{Estimate} = \frac{780}{50} = 15.6$$

A1

Answer: 15.6

3 marks

13 (3 marks)

$$\text{Mean number of trials} = \frac{20}{5} = 4$$

B1

$$E(X) = \frac{1}{p}, \text{ so } \frac{1}{p} = 4$$

M1

$$\hat{p} = 0.25$$

A1

Answer: 0.25

3 marks

14 (5 marks)

$$E(X) = \int_0^1 (\theta + 1)x^{\theta+1} dx = \frac{\theta+1}{\theta+2}$$

M1 A1

$$\text{Equate to the sample mean: } \frac{\theta+1}{\theta+2} = 0.7$$

M1

$$\theta + 1 = 0.7\theta + 1.4$$

A1

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

A1

$$\text{Answer: } \theta = \frac{4}{3}$$

5 marks

15 (5 marks)

$$\Sigma x = -0.2, \Sigma x^2 = 34.64$$

B1

$$s^2 = \frac{1}{7}(\Sigma x^2 - 8\bar{x}^2) = 4.948$$

M1 A1

$$\frac{\theta^2}{3} = s^2, \text{ so } \theta = \sqrt{3s^2}$$

M1

$$\theta = 3.85$$

A1

$$\text{Answer: } \theta \approx 3.85$$

5 marks

16 (6 marks)

$$L(\lambda) = \prod \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} = \frac{e^{-8\lambda} \lambda^{12}}{\prod x_i!}$$

M1 A1

$$\ln L = -8\lambda + 12 \ln \lambda - \ln \prod x_i!$$

M1

$$\frac{d \ln L}{d\lambda} = -8 + \frac{12}{\lambda} = 0$$

M1

$$\hat{\lambda} = \frac{12}{8} = 1.5 = \bar{x}$$

A1

$$\frac{d^2 \ln L}{d\lambda^2} = -\frac{12}{\lambda^2} < 0, \text{ so this is a maximum}$$

A1

$$\text{Answer: } \hat{\lambda} = 1.5$$

6 marks

17 (5 marks)

$$\hat{\lambda} = \frac{62}{40} = 1.55$$

M1 A1

$$P(X \geq 3) = 1 - e^{-1.55} \left(1 + 1.55 + \frac{1.55^2}{2} \right)$$

M1

$$= 1 - 0.2122 \times 3.7513$$

A1

$$\approx 0.204$$

A1

$$\text{Answer: } 0.204$$

5 marks

18 (6 marks)

$$\hat{p}_A = \frac{45}{120} = 0.375, \hat{p}_B = \frac{30}{100} = 0.3$$

B1 B1

$$\text{Estimated difference} = 0.075$$

B1

$$\text{Var}(\hat{p}_A - \hat{p}_B) = \frac{p_A q_A}{120} + \frac{p_B q_B}{100}$$

M1

$$= \frac{0.375 \times 0.625}{120} + \frac{0.3 \times 0.7}{100} = 0.001953 + 0.0021$$

A1

$$\text{Standard error} = \sqrt{0.004053} \approx 0.0637$$

A1

$$\text{Answer: } 0.075; \text{ SE } \approx 0.0637$$

6 marks

19 (6 marks)

$$\Sigma x = 120, \text{ new } \Sigma x = 135, \text{ new mean} = \frac{135}{11} = 12.3$$

M1 A1

$$\Sigma x^2 = 9s^2 + 10\bar{x}^2 = 36 + 1440 = 1476$$

M1

$$\text{New } \Sigma x^2 = 1476 + 225 = 1701$$

A1

$$\text{New } s^2 = \frac{1}{10}(1701 - 11 \times (135/11)^2)$$

M1

$$= \frac{1}{10}(1701 - 1656.8) = 4.42$$

A1

Answer: mean ≈ 12.3 , $s^2 \approx 4.42$

6 marks

20 (5 marks)

$$\bar{x} = \frac{45}{5} = 9, \Sigma(x - \bar{x})^2 = 9 + 1 + 4 + 4 + 16 = 34$$

M1 A1

Dividing by n gives a biased estimator that underestimates σ^2 on average

E1

$$\text{Unbiased estimate } s^2 = \frac{34}{4} = 8.5$$

M1

$$s = \sqrt{8.5} \approx 2.92$$

A1

Answer: $s^2 = 8.5$, $s \approx 2.92$

5 marks

21 (6 marks)

Mid-values 5, 15, 25, 35, 45

B1

$$\Sigma fx = 40 + 330 + 875 + 875 + 450 = 2570, \bar{x} = 25.7$$

M1 A1

$$\Sigma fx^2 = 200 + 4950 + 21875 + 30625 + 20250 = 77900$$

M1

$$s^2 = \frac{1}{99}(77900 - 100 \times 25.7^2)$$

M1

$$= 120$$

A1

Answer: $\bar{x} = 25.7$, $s^2 = 120$

6 marks

22 (8 marks)

$$L(\lambda) = \prod \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \Sigma x_i}$$

M1 A1

$$\ln L = n \ln \lambda - \lambda \Sigma x_i$$

M1

$$\frac{d \ln L}{d \lambda} = \frac{n}{\lambda} - \Sigma x_i = 0$$

M1

$$\hat{\lambda} = \frac{n}{\Sigma x_i} = \frac{1}{\bar{x}}$$

A1

Second derivative $-\frac{n}{\lambda^2} < 0$, so a maximum

B1

$$\bar{x} = \frac{10}{5} = 2, \text{ so } \hat{\lambda} = 0.5$$

A1

$$P(X > 3) \approx e^{-0.5 \times 3} = e^{-1.5} = 0.223$$

A1

Answer: $\hat{\lambda} = 0.5$; 0.223

8 marks

23 (8 marks)

$$L(p) = p^n (1 - p)^{T - n}$$

M1 A1

$$\ln L = n \ln p + (T - n) \ln(1 - p)$$

M1

$$\frac{d \ln L}{dp} = \frac{n}{p} - \frac{T - n}{1 - p} = 0$$

M1

$$n(1 - p) = (T - n)p, \text{ so } n = Tp$$

A1

$$\hat{p} = \frac{n}{T}$$

A1

$$\hat{p} = \frac{40}{130} = \frac{4}{13} \approx 0.308$$

A1

$$P(X > 4) = (1 - p)^4 = \left(\frac{9}{13}\right)^4 \approx 0.230$$

A1

Answer: $\hat{p} = \frac{4}{13}$; 0.230

8 marks

24 (7 marks)

(a) $L(\theta) = \left(\frac{1}{\theta}\right)^n$ provided $\theta \geq \text{every } x_i$, and 0 otherwise

M1 A1

$\left(\frac{1}{\theta}\right)^n$ decreases as θ increases

M1

So θ should be as small as possible while still being at least the largest observation

E1

Answer: $L = \theta^{-n}$ for $\theta \geq \max x$

4 marks

(b) $\theta = \max x_i = 5.1$

B1 B1

The true θ is at least as big as the sample maximum, and usually bigger

E1

Answer: $\theta = 5.1$

3 marks

25 (6 marks)

$$E(X) = \frac{a+b}{2}, \text{Var}(X) = \frac{(b-a)^2}{12}$$

B1

$$\frac{a+b}{2} = 5.60, \text{ so } a+b = 11.20$$

M1

$$\frac{(b-a)^2}{12} = 1.282, \text{ so } b-a = \sqrt{12 \times 1.282} = 3.923$$

M1 A1

$$a = 3.64$$

A1

$$b = 7.56$$

A1

Answer: $a \approx 3.64$, $b \approx 7.56$

6 marks

26 (6 marks)

$$\text{Pooled estimate: } \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

M1

$$= \frac{11 \times 6.4 + 17 \times 9.2}{28}$$

A1

$$= \frac{70.4 + 156.4}{28} = 8.10$$

A1

The weights are $\frac{11}{28}$ and $\frac{17}{28}$, in proportion to the degrees of freedom ($n - 1$)

M1

The simple average gives equal weight 0.5 to each, ignoring that the larger sample carries more information

E1

Simple average = 7.8 compared with pooled 8.10

A1

Answer: 8.10

6 marks

27 (10 marks)

$$(a) L(\theta) = \theta^n \prod x_i^{\theta-1}$$

M1 A1

$$\ln L = n \ln \theta + (\theta - 1) \sum \ln x_i$$

M1

$$\frac{d \ln L}{d\theta} = \frac{n}{\theta} + \sum \ln x_i = 0$$

M1

$$\theta = -\frac{n}{\sum \ln x_i}$$

A1

Answer: $\theta = -n/\sum \ln x_i$

5 marks

$$(b) \sum \ln x_i = -2.630$$

M1 A1

$$\theta = \frac{6}{2.630} = 2.28$$

A1

$$\text{Method of moments: } \bar{x} = 0.6667, \frac{\theta}{\theta + 1} = \bar{x} \text{ gives } \theta = \frac{\bar{x}}{1 - \bar{x}} = 2.00$$

M1

The two estimates are close but not equal

A1

Answer: 2.28; 2.00

5 marks

28 (11 marks)

$$(a) \ln L = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum (x_i - \mu)^2 \quad \text{M1 A1}$$

$$\partial/\partial\mu: \frac{1}{\sigma^2} \sum (x_i - \mu) = 0, \text{ so } \hat{\mu} = \bar{x} \quad \text{M1 A1}$$

$$\partial/\partial\sigma^2: -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum (x_i - \mu)^2 = 0 \quad \text{M1 A1}$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum (x_i - \bar{x})^2 \quad \text{A1}$$

$$\text{Answer: } \hat{\mu} = \bar{x}, \hat{\sigma}^2 = \sum (x - \bar{x})^2 / n \quad \text{7 marks}$$

$$(b) \bar{x} = \frac{41.4}{8} = 5.175 \quad \text{B1}$$

$$\sum (x - \bar{x})^2 = 6.155, \hat{\sigma}^2 = 0.769 \quad \text{M1 A1}$$

$$s^2 = \frac{6.155}{7} = 0.879; \text{ the maximum likelihood estimator divides by } n \text{ and is biased} \quad \text{E1}$$

$$\text{Answer: } \hat{\mu} = 5.175, \hat{\sigma}^2 = 0.769, s^2 = 0.879 \quad \text{4 marks}$$

29 (9 marks)

$$(a) E(X) = \int_0^\theta \frac{2x(\theta - x)}{\theta^2} dx = \frac{2}{\theta^2} \left[\frac{\theta x^2}{2} - \frac{x^3}{3} \right]_0^\theta \quad \text{M1 A1}$$

$$= \frac{2}{\theta^2} \times \frac{\theta^3}{6} = \frac{\theta}{3} \quad \text{A1}$$

$$\theta = 3\bar{X} \quad \text{A1}$$

$$\text{Answer: } \theta = 3\bar{X} \quad \text{4 marks}$$

$$(b) \bar{x} = \frac{5.7}{5} = 1.14, \theta = 3.42 \quad \text{M1 A1}$$

$$\theta \geq \text{largest observation } 2.6, \text{ so it is consistent} \quad \text{E1}$$

$$P(X > 1) = \left(1 - \frac{1}{\theta}\right)^2 = \left(\frac{3.42 - 1}{3.42}\right)^2 \quad \text{M1}$$

$$= 0.501 \quad \text{A1}$$

$$\text{Answer: } 3.42; 0.501 \quad \text{5 marks}$$

30 (9 marks)

$$(a) \text{ Counts } N_i \sim \text{Po}(\lambda t_i); L = \prod \frac{e^{-\lambda t_i} (\lambda t_i)^{n_i}}{n_i!} \quad \text{M1 A1}$$

$$\ln L = -\lambda \sum t_i + (\sum n_i) \ln \lambda + \text{constant} \quad \text{M1}$$

$$\frac{d \ln L}{d\lambda} = -\sum t_i + \frac{\sum n_i}{\lambda} = 0 \quad \text{M1}$$

$$\hat{\lambda} = \frac{\sum n_i}{\sum t_i} \quad \text{A1}$$

$$= \frac{20}{6} = \frac{10}{3} \text{ per hour} \quad \text{A1}$$

$$\text{Answer: } \hat{\lambda} = \frac{10}{3} \quad \text{6 marks}$$

$$(b) \text{ Rate for 0.5 hours: } \frac{10}{6} = \frac{5}{3}; P(0) \approx e^{-5/3} = 0.189 \quad \text{M1 A1}$$

$$\text{Expected number in 8 hours} = 8 \times \frac{10}{3} = 26.7 \quad \text{A1}$$

$$\text{Answer: } 0.189; 26.7 \quad \text{3 marks}$$

Unbiased Estimators**1 (2 marks)**

$$T \text{ is an unbiased estimator of } \theta \text{ if } E(T) = \theta \quad \text{B1}$$

$$\text{for all possible values of } \theta \quad \text{B1}$$

$$\text{Answer: } E(T) = \theta \quad \text{2 marks}$$

2 (2 marks)

$$E(\bar{X}) = \frac{1}{n}(E(X_1) + \dots + E(X_n)) = \frac{1}{n} \times n\mu$$

M1

$$= \mu$$

A1

$$\text{Answer: } E(\bar{X}) = \mu$$

2 marks

3 (2 marks)

$$E(T) = 3\mu - 2\mu$$

M1

$$= \mu$$

A1

$$\text{Answer: } E(T) = \mu$$

2 marks

4 (3 marks)

$$E(U) = \frac{3\mu}{2}$$

M1

$$\text{Bias} = \frac{3\mu}{2} - \mu = \frac{\mu}{2}$$

A1

$$k = \frac{2}{3}$$

A1

$$\text{Answer: } E(U) = \frac{3\mu}{2}; \text{ bias } \frac{\mu}{2}; k = \frac{2}{3}$$

3 marks

5 (2 marks)

$$E(\hat{p}) = \frac{E(X)}{n} = \frac{np}{n}$$

M1

$$= p$$

A1

$$\text{Answer: } E(\hat{p}) = p$$

2 marks

6 (2 marks)

$$E(T) = \frac{3\mu}{4}$$

M1

$$\text{Bias} = \frac{3\mu}{4} - \mu = -\frac{\mu}{4}$$

A1

$$\text{Answer: } -\frac{\mu}{4}$$

2 marks

7 (6 marks)

$$E(T_1) = \frac{2\mu}{2} = \mu$$

M1 A1

$$E(T_2) = \frac{3\mu}{3} = \mu$$

A1

$$\text{Var}(T_1) = \frac{\sigma^2 + \sigma^2}{4} = \frac{\sigma^2}{2}$$

M1

$$\text{Var}(T_2) = \frac{\sigma^2 + 4\sigma^2}{9} = \frac{5\sigma^2}{9}$$

A1

$$\frac{\sigma^2}{2} < \frac{5\sigma^2}{9}, \text{ so } T_1 \text{ is better (smaller variance)}$$

A1

$$\text{Answer: } T_1$$

6 marks

8 (5 marks)

$$E(T) = a\mu + (1 - a)\mu = \mu$$

M1 A1

$$\text{Var}(T) = a^2\sigma^2 + (1 - a)^2\sigma^2$$

M1

$$\frac{d}{da}: 2a - 2(1 - a) = 0$$

M1

$$a = \frac{1}{2}$$

A1

$$\text{Answer: } a = \frac{1}{2}$$

5 marks

9 (4 marks)

$$E(X) = \frac{\theta}{2}, \text{ so } E(2X) = 2 \times \frac{\theta}{2} = \theta$$

M1 A1

$$\text{Var}(X) = \frac{\theta^2}{12}, \text{ so } \text{Var}(2X) = 4 \times \frac{\theta^2}{12n}$$

M1

$$= \frac{\theta^2}{3n}$$

A1

$$\text{Answer: Var} = \frac{\theta^2}{3n}$$

4 marks

10 (3 marks)

$$E(\sum(X_i - \bar{X})^2) = 4 \times 10 = 40$$

M1

$$E(T) = \frac{40}{5} = 8$$

A1

$$\text{Bias} = 8 - 10 = -2$$

A1

$$\text{Answer: } E(T) = 8, \text{ bias } -2$$

3 marks

11 (5 marks)

$$E(T) = \frac{E(X) + 1}{12} = \frac{10p + 1}{12}$$

M1 A1

$$\text{Bias} = \frac{10p + 1}{12} - p = \frac{1 - 2p}{12}$$

M1 A1

$$\text{Unbiased when } 1 - 2p = 0, \text{ i.e. } p = \frac{1}{2}$$

A1

$$\text{Answer: } E(T) = \frac{10p + 1}{12}; \text{ unbiased only for } p = \frac{1}{2}$$

5 marks

12 (3 marks)

$$\text{Let } U = \frac{T - 3}{2}; E(U) = \frac{2\theta + 3 - 3}{2} = \theta$$

M1 A1

$$\text{Var}(U) = \frac{1}{4} \text{Var}(T)$$

A1

$$\text{Answer: } U = \frac{T - 3}{2}, \text{ Var}(U) = \frac{1}{4} \text{Var}(T)$$

3 marks

13 (5 marks)

$$E(A) = \frac{3\lambda}{3} = \lambda, E(B) = \frac{6\lambda}{6} = \lambda$$

M1 A1

$$\text{Var}(A) = \frac{3\lambda}{9} = \frac{\lambda}{3}$$

M1

$$\text{Var}(B) = \frac{(1 + 4 + 9)\lambda}{36} = \frac{7\lambda}{18}$$

A1

$$\frac{\lambda}{3} < \frac{7\lambda}{18}, \text{ so A has the smaller variance}$$

A1

$$\text{Answer: A has smaller variance}$$

5 marks

14 (6 marks)

$$E(X^2) = \text{Var}(X) + [E(X)]^2$$

M1

$$= \frac{\sigma^2}{n} + \mu^2$$

A1

$$\text{Bias} = \frac{\sigma^2}{n}$$

A1

$$E(s^2) = \sigma^2, \text{ so } E\left(\frac{s^2}{n}\right) = \frac{\sigma^2}{n}$$

M1

$$U = X^2 - \frac{s^2}{n} \text{ has } E(U) = \mu^2$$

M1 A1

$$\text{Answer: } X^2 - \frac{s^2}{n}$$

6 marks

15 (7 marks)

$$E(T) = a\mu + (1 - a)\mu = \mu \quad \text{M1 A1}$$

$$\text{Var}(T) = a^2 \frac{\sigma^2}{10} + (1 - a)^2 \frac{\sigma^2}{15} \quad \text{M1 A1}$$

$$\frac{d}{da} \left(\frac{a}{5} - \frac{2(1 - a)}{15} \right) = 0 \quad \text{M1}$$

$$3a = 2(1 - a), \text{ so } 5a = 2 \quad \text{A1}$$

$$a = \frac{2}{5} \text{ (weights proportional to the sample sizes)} \quad \text{A1}$$

$$\text{Answer: } a = \frac{2}{5} \quad 7 \text{ marks}$$

16 (5 marks)

$$\text{Bias} = \frac{n\theta}{n+1} - \theta = -\frac{\theta}{n+1} \quad \text{M1 A1}$$

$$\text{Unbiased estimator: } \frac{n+1}{n}M \quad \text{M1 A1}$$

$$\text{Estimate} = \frac{6}{5} \times 8.4 = 10.08 \quad \text{A1}$$

$$\text{Answer: } -\frac{\theta}{n+1}; \frac{n+1}{n}M; 10.08 \quad 5 \text{ marks}$$

17 (4 marks)

$$E((X_i - \mu)^2) = \text{Var}(X_i) = \sigma^2 \text{ for each } i \quad \text{M1 A1}$$

$$E(T) = \frac{1}{n} \times n\sigma^2 \quad \text{M1}$$

$$= \sigma^2 \quad \text{A1}$$

$$\text{Answer: } E(T) = \sigma^2 \quad 4 \text{ marks}$$

18 (5 marks)

$$E(T) = (a + b + c)\mu = \mu, \text{ so } a + b + c = 1 \quad \text{M1 A1}$$

$$c = 1 - 0.2 - 0.3 = 0.5 \quad \text{A1}$$

$$\text{Var}(T) = (0.04 + 0.09 + 0.25)\sigma^2 \quad \text{M1}$$

$$= 0.38\sigma^2 \quad \text{A1}$$

$$\text{Answer: } c = 0.5, \text{ Var}(T) = 0.38\sigma^2 \quad 5 \text{ marks}$$

19 (5 marks)

$$E(\bar{X}) = \lambda, \text{ Var}(\bar{X}) = \frac{\lambda}{n} \quad \text{B1}$$

$$E(\bar{X}^2) = \frac{\lambda}{n} + \lambda^2, \text{ so bias} = \frac{\lambda}{n} \quad \text{M1 A1}$$

$$E(\bar{X}) = \lambda, \text{ so } E(\bar{X}^2 - \frac{\bar{X}}{n}) = \frac{\lambda}{n} + \lambda^2 - \frac{\lambda}{n} \quad \text{M1}$$

$$U = \bar{X}^2 - \frac{\bar{X}}{n} \quad \text{A1}$$

$$\text{Answer: } U = \bar{X}^2 - \frac{\bar{X}}{n} \quad 5 \text{ marks}$$

20 (7 marks)

$$E(C) = a\theta + (1 - a)\theta = \theta$$

M1 A1

$$\text{Var}(C) = 4a^2 + 9(1 - a)^2$$

M1

$$\frac{d}{da}: 8a - 18(1 - a) = 0$$

M1

$$26a = 18, a = \frac{9}{13}$$

A1

$$\text{Var}(C) = 4 \times \frac{81}{169} + 9 \times \frac{16}{169} = \frac{324 + 144}{169}$$

M1

$$= \frac{36}{13} \approx 2.77, \text{ less than both 4 and 9}$$

A1

$$\text{Answer: } a = \frac{9}{13}, \text{ minimum variance } \frac{36}{13}$$

7 marks

21 (7 marks)

$$\Sigma(X_i - \bar{X})^2 = \Sigma X_i^2 - n\bar{X}^2$$

M1 A1

$$E(X_i^2) = \sigma^2 + \mu^2$$

B1

$$E(\bar{X}^2) = \frac{\sigma^2}{n} + \mu^2$$

B1

$$E(\Sigma(X_i - \bar{X})^2) = n(\sigma^2 + \mu^2) - n(\frac{\sigma^2}{n} + \mu^2)$$

M1

$$= n\sigma^2 - \sigma^2 = (n - 1)\sigma^2$$

A1

$$\text{Dividing by } n - 1 \text{ gives } E(s^2) = \sigma^2$$

A1

$$\text{Answer: } E(s^2) = \sigma^2$$

7 marks

22 (6 marks)

$$E(X^2) = \text{Var}(X) + [E(X)]^2$$

M1

$$= \lambda + \lambda^2$$

A1

$$E(T) = E(X^2) - E(X) = \lambda + \lambda^2 - \lambda$$

M1

$$= \lambda^2$$

A1

$$t = 25 - 5 = 20$$

M1

$$x^2 = 25, \text{ which overestimates on average (its bias is } \lambda)$$

A1

$$\text{Answer: } T = X^2 - X; \text{ estimate 20}$$

6 marks

23 (7 marks)

$$E(X^2) = np(1 - p) + n^2p^2$$

M1 A1

$$E(X(X - 1)) = E(X^2) - E(X) = np - np^2 + n^2p^2 - np$$

M1

$$= n(n - 1)p^2$$

A1

$$E(T) = p^2$$

A1

$$\text{For } n = 20, x = 8: t = \frac{56}{380} = 0.147$$

M1

$$(x/n)^2 = 0.16, \text{ which is larger; } (X/n)^2 \text{ has bias } \frac{p(1 - p)}{n} > 0$$

A1

$$\text{Answer: } t = 0.147$$

7 marks

24 (8 marks)

$$P(m > x) = (e^{-x/\theta})^n = e^{-nx/\theta}$$

M1 A1

$$\text{So } m \sim \text{Exp with mean } \frac{\theta}{n}$$

A1

$$E(nm) = n \times \frac{\theta}{n} = \theta$$

M1 A1

$$\text{Var}(m) = \frac{\theta^2}{n^2}, \text{ so } \text{Var}(nm) = n^2 \times \frac{\theta^2}{n^2} = \theta^2$$

M1

$$\text{Var}(\bar{X}) = \frac{\theta^2}{n}$$

A1

For $n > 1$, $\bar{X}$ has the smaller variance, so it is the better estimator

A1

$$\text{Answer: } \text{Var}(nm) = \theta^2 > \text{Var}(\bar{X}) = \theta^2/n$$

8 marks

25 (9 marks)

$$E(M) = \frac{n\theta}{n+1}$$

B1

$$E(M^2) = \int_0^\theta \frac{nm^{n+1}}{\theta^n} dm = \frac{n\theta^2}{n+2}$$

M1 A1

$$\text{Var}(M) = \frac{n\theta^2}{n+2} - \frac{n^2\theta^2}{(n+1)^2}$$

M1

$$= \frac{n\theta^2[(n+1)^2 - n(n+2)]}{(n+1)^2(n+2)} = \frac{n\theta^2}{(n+1)^2(n+2)}$$

A1

$$\text{Var}(T) = \left(\frac{n+1}{n}\right)^2 \text{Var}(M) = \frac{\theta^2}{n(n+2)}$$

M1 A1

$$n = 10: \text{Var}(T) = \frac{\theta^2}{120}, \text{Var}(2\bar{X}) = \frac{\theta^2}{30}$$

M1

T has variance four times smaller, so it is the better estimator

A1

$$\text{Answer: } \text{Var}(T) = \frac{\theta^2}{n(n+2)}$$

9 marks

26 (7 marks)

$$\text{For two values } a, b: s = \frac{|a-b|}{\sqrt{2}}$$

B1

$$|a-b| = 0 \text{ (3 samples), } 2 \text{ (4 samples), } 4 \text{ (2 samples)}$$

M1

$$s = 0, \sqrt{2}, 2\sqrt{2} \text{ with probabilities } \frac{3}{9}, \frac{4}{9}, \frac{2}{9}$$

A1

$$E(s) = \frac{4\sqrt{2} + 4\sqrt{2}}{9} = \frac{8\sqrt{2}}{9}$$

M1 A1

$$\approx 1.257$$

A1

$$\sigma = \sqrt{8/3} \approx 1.633 \neq E(s), \text{ so } s \text{ is biased (it underestimates } \sigma)$$

A1

$$\text{Answer: } E(s) \approx 1.257 < \sigma \approx 1.633$$

7 marks

27 (10 marks)

$$(a) E\left(\frac{1}{S}\right) = \int_0^\infty \frac{1}{s} \times \frac{\lambda^n s^{n-1} e^{-\lambda s}}{(n-1)!} ds$$

M1

$$= \frac{\lambda^n}{(n-1)!} \int_0^\infty s^{n-2} e^{-\lambda s} ds$$

M1

$$\text{The integral is } \frac{(n-2)!}{\lambda^{n-1}} \text{ (gamma integral)}$$

A1

$$= \frac{\lambda^n (n-2)!}{(n-1)! \lambda^{n-1}} = \frac{\lambda}{n-1}$$

A1

$$\text{Answer: } \frac{\lambda}{n-1}$$

4 marks

$$(b) \frac{1}{\bar{X}} = \frac{n}{S}, \text{ so } E\left(\frac{1}{\bar{X}}\right) = \frac{n\lambda}{n-1}$$

M1 A1

$$\text{Bias} = \frac{n\lambda}{n-1} - \lambda = \frac{\lambda}{n-1}$$

A1

Unbiased estimator: $\frac{n-1}{n\bar{X}}$

A1

Answer: $\frac{n-1}{n\bar{X}}$

4 marks

(c) $\frac{1}{2.5} = 0.4$

B1

$\frac{7}{8 \times 2.5} = 0.35$

B1

Answer: 0.4 and 0.35

2 marks

28 (8 marks)

(a) $E(T) = \mu \sum a_i = \mu$ for all μ only if $\sum a_i = 1$

M1 A1

$\text{Var}(T) = \sigma^2 \sum a_i^2$

A1

Answer: $\sum a_i = 1$; $\text{Var} = \sigma^2 \sum a_i^2$

3 marks

(b) $\sum (a_i - \frac{1}{n})^2 = \sum a_i^2 - \frac{2}{n} \sum a_i + \frac{n}{n^2}$

M1 A1

Using $\sum a_i = 1$: $\sum a_i^2 - \frac{1}{n}$

A1

The left side is ≥ 0 , so $\sum a_i^2 \geq \frac{1}{n}$

M1

$\text{Var}(T) \geq \frac{\sigma^2}{n} = \text{Var}(\bar{X})$, with equality only when every $a_i = \frac{1}{n}$

A1

Answer: $\text{Var}(T) \geq \frac{\sigma^2}{n} = \text{Var}(\bar{X})$, with equality only for $T = \bar{X}$

5 marks

29 (9 marks)

(a) $\text{Bias}(M) = -\frac{\theta}{n+1}$

B1

$\text{MSE}(M) = \frac{n\theta^2}{(n+1)^2(n+2)} + \frac{\theta^2}{(n+1)^2}$

M1

$= \frac{\theta^2(n+n+2)}{(n+1)^2(n+2)}$

A1

$= \frac{2\theta^2}{(n+1)(n+2)}$

A1

Answer: $\text{MSE}(M) = \frac{2\theta^2}{(n+1)(n+2)}$

4 marks

(b) $\text{MSE}(M) = \frac{2}{132}\theta^2 = 0.0152\theta^2$

M1 A1

$\text{MSE}(\frac{11}{10}M) = \frac{1}{120}\theta^2 = 0.00833\theta^2$

A1

$\text{MSE}(2\bar{X}) = \frac{1}{30}\theta^2 = 0.0333\theta^2$

A1

The biased estimator M has a smaller error than $2\bar{X}$, but the unbiased scaled maximum is best of the three

E1

Answer: $0.0152\theta^2$, $0.00833\theta^2$, $0.0333\theta^2$

5 marks

30 (9 marks)

(a) $E(z^S) = \sum z^s \frac{e^{-n\lambda}(n\lambda)^s}{s!}$

M1

$= e^{-n\lambda} \sum \frac{(n\lambda z)^s}{s!} = e^{-n\lambda} e^{n\lambda z}$

M1 A1

$= e^{n\lambda(z-1)}$

A1

With $z = 1 - \frac{1}{n}$: $E(U) = e^{n\lambda(-1/n)} = e^{-\lambda}$

M1 A1

Answer: $E(U) = e^{-\lambda} = \theta$

6 marks

(b) $U = 0.9^{13} = 0.254$

B1

$e^{-X} = e^{-1.3} = 0.273$

B1

U is preferred as it is unbiased; e^{-X} is biased

E1

Answer: 0.254 and 0.273**3 marks**

Confidence Intervals

1 (2 marks)

$$\text{Sample mean} = \frac{12.3 + 14.7}{2} = 13.5$$

B1

$$\text{Width} = 14.7 - 12.3 = 2.4$$

B1**Answer: 13.5; 2.4****2 marks**

2 (3 marks)

90%: 1.645

B1

95%: 1.96

B1

99%: 2.576

B1**Answer: 1.645, 1.96, 2.576****3 marks**

3 (3 marks)

$$\text{Standard error} = \frac{4}{\sqrt{16}} = 1$$

B1

$$50 \pm 1.96 \times 1$$

M1

$$= (48.04, 51.96)$$

A1**Answer: (48.04, 51.96)****3 marks**

4 (2 marks)

(a) The critical value increases, so the interval is wider**B1****Answer: Wider****1 mark****(b)** The standard error decreases, so the interval is narrower**B1****Answer: Narrower****1 mark**

5 (3 marks)

$$\text{Standard error} = \frac{3}{6} = 0.5$$

B1

$$20 \pm 2.576 \times 0.5$$

M1

$$= (18.712, 21.288)$$

A1**Answer: (18.712, 21.288)****3 marks**

6 (2 marks)

$$\text{Mean} = \frac{8.4 + 9.6}{2} = 9$$

B1

$$\text{Margin of error} = 9.6 - 9 = 0.6$$

B1**Answer: 9; 0.6****2 marks**

7 (3 marks)

If the sampling procedure were repeated many times and an interval calculated each time approximately 95% of the intervals would contain the true population mean

B1**B1**

For any one interval the mean either is or is not inside it

E1**Answer: 95% of such intervals contain μ** **3 marks**

8 (4 marks)

$$\text{Sample mean} = \frac{1198}{8} = 149.75$$

B1

$$\text{Standard error} = \frac{8}{\sqrt{8}} = 2.828$$

B1

$$149.75 \pm 2.326 \times 2.828$$

M1

$$= (143.17, 156.33)$$

A1**Answer: (143.17, 156.33)****4 marks**

9 (4 marks)

$$\text{Margin of error} = \frac{37.8 - 34.2}{2} = 1.8$$

M1

$$1.96 \times \frac{\sigma}{\sqrt{25}} = 1.8$$

M1 A1

$$\sigma = \frac{1.8 \times 5}{1.96} = 4.59$$

A1

Answer: $\sigma \approx 4.59$

4 marks

10 (3 marks)

$$1.96 \times \frac{4}{\sqrt{n}} \leq 0.5$$

M1

$$\sqrt{n} \geq \frac{1.96 \times 4}{0.5} = 15.68$$

A1

$$n \geq 245.9, \text{ so } n = 246$$

A1

Answer: $n = 246$

3 marks

11 (3 marks)

$$\text{Standard error} = \frac{3}{\sqrt{12}} = 0.8660$$

B1

$$48.5 \pm 2.326 \times 0.8660$$

M1

$$= (46.49, 50.51)$$

A1

Answer: (46.49, 50.51)

3 marks

12 (4 marks)

$$\text{Margin} = 1.96, \text{ standard error} = \frac{10}{\sqrt{100}} = 1$$

M1 A1

$$z = 1.96$$

A1

Confidence level 95%

A1

Answer: 95%

4 marks

13 (3 marks)

$$\text{Standard error} \approx \frac{12}{10} = 1.2$$

B1

$$65 \pm 1.96 \times 1.2$$

M1

$$= (62.65, 67.35)$$

A1

Answer: (62.65, 67.35)

3 marks

14 (3 marks)

The population mean is a fixed (unknown) value, not a random variable

E1

The interval is the random quantity: it varies from sample to sample

E1

Correct: 95% of such intervals constructed from repeated samples contain the true mean

E1

Answer: Interval, not μ , is random

3 marks

15 (5 marks)

Sample mean = 12, margin of error 1.6

B1

$$1.96 \times \text{se} = 1.6, \text{ so } \text{se} = 0.8163$$

M1 A1

$$99\% \text{ margin} = 2.576 \times 0.8163 = 2.103$$

M1

Interval: (9.90, 14.10)

A1

Answer: (9.90, 14.10)

5 marks

16 (5 marks)

$$\text{Standard error} = \frac{90}{6} = 15$$

B1

$$1165 \pm 1.96 \times 15$$

M1

$$= (1135.6, 1194.4)$$

A1

1200 lies outside the interval

M1

There is evidence that the mean is lower than the claimed 1200 hours

A1

Answer: (1135.6, 1194.4)

5 marks

17 (4 marks)

$$\text{Width} = 2 \times 2.576 \times \frac{3}{\sqrt{n}} \leq 1$$

M1 A1

$$\sqrt{n} \geq 15.456$$

M1

$$n \geq 238.9, \text{ so } n = 239$$

A1

Answer: n = 239

4 marks

18 (6 marks)

Difference of sample means = 1.5

B1

$$\text{Standard error} = \sqrt{\frac{25}{40} + \frac{16}{50}} = 0.9721$$

M1 A1

$$1.5 \pm 1.96 \times 0.9721$$

M1

$$= (-0.405, 3.405)$$

A1

The interval contains 0, so there is no clear evidence that the means differ

E1

Answer: (-0.405, 3.405)

6 marks

19 (6 marks)

$$\bar{x} = \frac{372}{60} = 6.2$$

B1

$$s^2 = \frac{1}{59}(2500 - 60 \times 6.2^2) = 3.281$$

M1 A1

$$\text{Standard error} = \frac{1.811}{\sqrt{60}} = 0.2339$$

M1

$$6.2 \pm 1.96 \times 0.2339$$

M1

$$= (5.74, 6.66)$$

A1

Answer: (5.74, 6.66)

6 marks

20 (5 marks)

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}), \text{ so } Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

B1

$$P(-1.96 < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < 1.96) = 0.95$$

M1

$$P(-1.96 \frac{\sigma}{\sqrt{n}} < \bar{X} - \mu < 1.96 \frac{\sigma}{\sqrt{n}}) = 0.95$$

M1

$$P(\bar{X} - 1.96 \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 1.96 \frac{\sigma}{\sqrt{n}}) = 0.95$$

M1

$$\text{So the interval is } \bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

A1

Answer: $\bar{x} \pm 1.96\sigma/\sqrt{n}$

5 marks

21 (5 marks)Standard deviation = mean ≈ 2.9

B1

$$\text{Standard error} \approx \frac{2.9}{\sqrt{100}} = 0.29$$

M1 A1

$$2.9 \pm 1.96 \times 0.29$$

M1

$$= (2.33, 3.47) \text{ minutes}$$

A1

Answer: (2.33, 3.47)

5 marks

22 (6 marks)

$$n = 20: \text{se} = \frac{6}{\sqrt{20}} = 1.342$$

B1

$$52 \pm 1.96 \times 1.342 = (49.37, 54.63)$$

M1 A1

$$n = 80: \text{se} = \frac{6}{\sqrt{80}} = 0.6708$$

B1

$$52 \pm 1.96 \times 0.6708 = (50.69, 53.31)$$

A1

Width for n = 80 is half that for n = 20 (sample size multiplied by 4)

E1

Answer: (49.37, 54.63) and (50.69, 53.31)

6 marks

23 (7 marks)

(a) Width $\propto \frac{1}{\sqrt{n}}$ M1

Half the width needs $\sqrt{n}$ doubled, so n is multiplied by 4 M1

$n = 200$ A1

Answer: $n = 200$ 3 marks

(b) Margin = $0.3 = z \frac{\sigma}{\sqrt{50}}$; from the original: $0.6 = 1.96 \frac{\sigma}{\sqrt{50}}$ M1

Dividing: $z = \frac{0.3}{0.6} \times 1.96 = 0.98$ M1 A1

Level = $2\Phi(0.98) - 1 = 2(0.8365) - 1 = 0.673$ i.e. about 67% A1

Answer: about 67% 4 marks

24 (6 marks)

Width = $2 \times 1.645 \times \frac{8}{\sqrt{n}} = 3.29$ M1 A1

$\sqrt{n} = \frac{2 \times 1.645 \times 8}{3.29} = 8$, so $n = 64$ A1

Standard error = $\frac{8}{8} = 1$ B1

$40 \pm 2.576 \times 1$ M1

$= (37.42, 42.58)$ A1

Answer: $n = 64$; (37.42, 42.58) 6 marks

25 (6 marks)

Combined mean = $\frac{30 \times 12.4 + 50 \times 13.1}{80} = 12.84$ M1 A1

Standard error = $\frac{5}{\sqrt{80}} = 0.5590$ B1

Interval: $12.84 \pm 1.96 \times 0.5590 = (11.74, 13.93)$ M1 A1

Width $2 \times 1.96 \times 0.559 = 2.19$ compared with $2 \times 1.96 \times 0.707 = 2.77$: narrower, since more data is used E1

Answer: (11.74, 13.93) 6 marks

26 (6 marks)

Correct: $\sigma = 3$, $se = \frac{3}{5} = 0.6$ B1

$31.2 \pm 1.96 \times 0.6 = (30.02, 32.38)$ M1 A1

Incorrect: $\sigma = 9$, $se = \frac{9}{5} = 1.8$ B1

$31.2 \pm 1.96 \times 1.8 = (27.67, 34.73)$ A1

The second interval is three times too wide, because $N(\mu, 9)$ specifies the variance 9 and standard deviation 3 E1

Answer: (30.02, 32.38) 6 marks

27 (8 marks)

Each interval contains μ with probability 0.95, independently B1

$Y \sim B(10, 0.95)$ M1

$P(Y \geq 9) = 10(0.95)^9(0.05) + (0.95)^{10}$ M1 A1

$= 0.914$ A1

Expected number not containing $\mu = 10 \times 0.05 = 0.5$ A1

$P(\text{all 10 contain } \mu) = 0.95^{10} = 0.599$ M1 A1

Answer: 0.914; 0.5; 0.599 8 marks

28 (8 marks)

$$\bar{X} \sim N(\mu, \frac{64}{n}); \text{ the interval is } \bar{X} \pm 1.96 \frac{6}{\sqrt{n}} \quad \text{B1}$$

$$\text{It contains } \mu \text{ if } |\bar{X} - \mu| < 1.96 \times \frac{6}{\sqrt{n}} \quad \text{M1}$$

$$P = P(|Z| < \frac{1.96 \times 6}{\sqrt{n}} \div \frac{8}{\sqrt{n}}) = P(|Z| < 1.47) \quad \text{M1 A1}$$

$$= 2\Phi(1.47) - 1 = 0.858 \quad \text{A1}$$

The nominal 95% interval is really only about 86% reliable, since σ is underestimated E1

$$\text{For true } \sigma = 10: 2\Phi(\frac{1.96 \times 6}{10}) - 1 = 2\Phi(1.176) - 1 \quad \text{M1}$$

$$= 0.760 \quad \text{A1}$$

Answer: 0.858; 0.760 8 marks

29 (10 marks)

$$\text{(a) } U = 47 + 1.645 \times \frac{4}{6} = 48.10 \quad \text{M1 A1}$$

All the 5% error is placed in one tail, so the critical value is z with $P(Z < z) = 0.95$, i.e. 1.645 E1

Answer: 48.10 3 marks

$$\text{(b) } 47 + 1.645 \times \frac{4}{\sqrt{n}} < 50 \quad \text{M1}$$

$$\frac{6.58}{\sqrt{n}} < 3 \quad \text{M1}$$

$$\sqrt{n} > 2.193, n > 4.81 \quad \text{A1}$$

$$n = 5 \quad \text{A1}$$

Answer: n = 5 4 marks

$$\text{(c) } 47 + 1.645 \times \frac{4}{\sqrt{n}} < 49, \text{ so } \frac{6.58}{\sqrt{n}} < 2 \quad \text{M1}$$

$$\sqrt{n} > 3.29 \quad \text{A1}$$

$$n > 10.8, \text{ so } n = 11 \quad \text{A1}$$

Answer: n = 11 3 marks

30 (10 marks)

$$\text{(a) Width} = 2 \times 1.96 \frac{\sigma}{\sqrt{n}}, \text{ which involves only } \sigma \text{ and } n \quad \text{B1}$$

$$\text{The interval contains } \mu \text{ exactly when } |\bar{X} - \mu| < 1.96 \frac{\sigma}{\sqrt{n}} \quad \text{M1}$$

$$P(|Z| < 1.96) \text{ with } Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1) \quad \text{M1}$$

$$= 0.95 \quad \text{A1}$$

$$\text{Answer: Width} = \frac{2 \times 1.96\sigma}{\sqrt{n}}, P(\text{contains } \mu) = 0.95 \quad \text{4 marks}$$

$$\text{(b) Both: } 0.95^2 = 0.9025 \quad \text{B1}$$

$$\text{Exactly one: } 2 \times 0.95 \times 0.05 \quad \text{M1}$$

$$= 0.095 \quad \text{A1}$$

Answer: 0.9025; 0.095 3 marks

$$\text{(c) } P(\text{all contain}) = 0.95^k < 0.5 \quad \text{M1}$$

$$k > \frac{\ln 0.5}{\ln 0.95} = 13.5 \quad \text{A1}$$

$$k = 14 \quad \text{A1}$$

Answer: k = 14 3 marks

Estimation of Means and Proportions

1 (3 marks)

(a) $v = 9$, $t = 2.262$

Answer: 2.262

B1

1 mark

(b) $v = 15$, $t = 2.131$

Answer: 2.131

B1

1 mark

(c) $v = 24$, $t = 2.064$

Answer: 2.064

B1

1 mark

2 (3 marks)

$$\hat{p} = 0.3, \text{ standard error} = \sqrt{\frac{0.3 \times 0.7}{200}} = 0.03240$$

B1

$$0.3 \pm 1.96 \times 0.03240$$

M1

$$= (0.2365, 0.3635)$$

A1

Answer: (0.2365, 0.3635)

3 marks

3 (2 marks)

The population standard deviation is unknown and the sample is small

B1

The population is assumed to be normally distributed

B1

Answer: σ unknown, n small, normal population

2 marks

4 (3 marks)

$$\text{Standard error} = \frac{4}{\sqrt{10}} = 1.265, t_9 = 2.262$$

B1

$$25 \pm 2.262 \times 1.265$$

M1

$$= (22.14, 27.86)$$

A1

Answer: (22.14, 27.86)

3 marks

5 (3 marks)

$$\hat{p} = 0.4$$

B1

$$\text{Standard error} = \sqrt{\frac{0.4 \times 0.6}{100}}$$

M1

$$= 0.0490$$

A1

Answer: 0.4; 0.0490

3 marks

6 (3 marks)

$$\text{Standard error} = \sqrt{\frac{0.35 \times 0.65}{80}} = 0.05333$$

B1

$$0.35 \pm 1.645 \times 0.05333$$

M1

$$= (0.262, 0.438)$$

A1

Answer: (0.262, 0.438)

3 marks

7 (5 marks)

$$\Sigma x = 100.5, \bar{x} = 12.56$$

B1

$$\Sigma(x - \bar{x})^2 = 2.399, s = 0.5854$$

M1 A1

$$t_7 = 2.365, \text{ standard error} = 0.2070$$

B1

$$\bar{x} \pm 2.365 \times 0.2070 = (12.07, 13.05)$$

A1

Answer: (12.07, 13.05)

5 marks

8 (5 marks)

$$\bar{y} = \frac{30}{12} = 2.5, \text{ so } \bar{x} = 52.5$$

B1

$$s^2 = \frac{1}{11}(210 - 12 \times 2.5^2) = 12.27$$

M1 A1

$$\text{Standard error} = 1.011 \text{ and } t_{11} = 1.796$$

B1

$$52.5 \pm 1.796 \times 1.011 = (50.68, 54.32)$$

A1

Answer: (50.68, 54.32)

5 marks

9 (5 marks)

$$\hat{p} = \frac{215}{500} = 0.43$$

B1

$$\text{Standard error} = \sqrt{\frac{0.43 \times 0.57}{500}} = 0.02214$$

M1

$$0.43 \pm 1.96 \times 0.02214$$

M1

$$= (0.3866, 0.4734)$$

A1

The whole interval is below 0.5, so A is unlikely to have a majority

E1

Answer: (0.3866, 0.4734)

5 marks

10 (4 marks)

$$\text{Margin} = 1.96 \sqrt{\frac{p(1-p)}{n}}, \text{ largest when } p = 0.5$$

B1

$$1.96 \times \frac{0.5}{\sqrt{n}} \leq 0.03$$

M1

$$\sqrt{n} \geq 32.67$$

A1

$$n \geq 1067.1, \text{ so } n = 1068$$

A1

Answer: n = 1068

4 marks

11 (5 marks)

$$\text{Difference} = 0.52 - 0.45 = 0.07$$

B1

$$\text{Standard error} = \sqrt{\frac{0.52 \times 0.48}{150} + \frac{0.45 \times 0.55}{200}}$$

M1

$$= \sqrt{0.001664 + 0.0012375} = 0.05387$$

A1

$$0.07 \pm 1.96 \times 0.05387$$

M1

$$= (-0.0356, 0.1756)$$

A1

Answer: (-0.0356, 0.1756)

5 marks

12 (5 marks)

$$\text{(a) Standard error} = \frac{3.1}{\sqrt{45}} = 0.4621$$

B1

$$t_{44} = 2.015, 12.2 \pm 0.9312 = (11.27, 13.13)$$

A1

Answer: (11.27, 13.13)

2 marks

$$\text{(b) } 12.2 \pm 1.96 \times 0.4621$$

M1

$$= (11.29, 13.11)$$

A1

Answer: (11.29, 13.11)

2 marks

(c) The intervals are very close since n is large; the t-interval is slightly wider

E1

Answer: Very close; the t-interval is slightly wider

1 mark

13 (4 marks)

$$\text{Sample mean} = \frac{18.4 + 21.6}{2} = 20$$

B1

$$\text{Margin} = 1.6 = t_{24} \times \frac{s}{\sqrt{25}} \text{ with } t_{24} = 2.064$$

M1

$$s = \frac{1.6 \times 5}{2.064}$$

M1

$$= 3.876$$

A1

Answer: $\bar{x} = 20$, $s = 3.876$

4 marks

14 (5 marks)

$$\text{Standard error} = \frac{12}{\sqrt{20}} = 2.683$$

B1

$$t_{19} = 2.093$$

B1

$$104.5 \pm 2.093 \times 2.683 = (98.88, 110.12)$$

M1 A1

100 lies inside the interval, so there is no evidence that the machine is incorrectly set

E1

Answer: (98.88, 110.12)

5 marks

15 (5 marks)

$$\text{Difference of means} = 1.7$$

B1

$$\text{Standard error} = \sqrt{\frac{36}{60} + \frac{49}{75}} = \sqrt{0.6 + 0.6533} = 1.120$$

M1 A1

$$1.7 \pm 1.96 \times 1.120$$

M1

$$= (-0.49, 3.89)$$

A1

Answer: (-0.49, 3.89)

5 marks

16 (5 marks)

$$\hat{p} = 0.55, \text{ so the interval is } 0.55 \pm 1.96\sqrt{\frac{0.55 \times 0.45}{n}}$$

B1

$$0.70 \text{ is outside the interval, so } 0.55 + 1.96\sqrt{\frac{0.55 \times 0.45}{n}} < 0.70$$

M1

$$\sqrt{\frac{0.2475}{n}} < \frac{0.15}{1.96}$$

M1

$$n > 0.2475 \times \left(\frac{1.96}{0.15}\right)^2 = 42.26$$

M1

$$\text{Least } n = 43$$

A1

Answer: n = 43

5 marks

17 (5 marks)

$$\hat{p} = \frac{232}{400} = 0.58$$

B1

$$\text{Standard error} = \sqrt{\frac{0.58 \times 0.42}{400}} = 0.02468$$

M1 A1

$$0.58 \pm 2.576 \times 0.02468 = (0.5164, 0.6436)$$

M1

0.5 lies outside the interval, so the claim is doubtful

A1

Answer: (0.5164, 0.6436)

5 marks

18 (5 marks)

$$E(\hat{P}) = p \text{ and } \text{Var}(\hat{P}) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}; \text{ the binomial is approximately normal for large } n$$

B1 M1

$$P(-1.96 < \frac{\hat{P} - p}{\sqrt{p(1-p)/n}} < 1.96) \approx 0.95$$

M1

$$\text{Rearranging gives } \hat{P} - 1.96\sqrt{p(1-p)/n} < p < \hat{P} + 1.96\sqrt{p(1-p)/n}$$

M1

For large n, replace p by the estimate $\hat{P}$ in the standard error

A1

$$\text{Answer: } \hat{P} \pm 1.96\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$$

5 marks

19 (5 marks)

$$\hat{p} = \frac{0.36 + 0.44}{2} = 0.4$$

B1

$$\text{Margin of error} = 0.04$$

B1

$$1.96\sqrt{\frac{0.4 \times 0.6}{n}} = 0.04$$

M1

$$n = \frac{1.96^2 \times 0.24}{0.04^2} = 576.2$$

A1

$$n \approx 576$$

A1

Answer: $\hat{p} = 0.4$, $n \approx 576$

5 marks

20 (5 marks)

Margin of error = 4

B1

$$\text{Standard error} = \frac{6}{\sqrt{10}} = 1.897$$

M1

$$t = \frac{4}{1.897} = 2.108$$

A1

$$1.833 < 2.108 < 2.262$$

M1

so the confidence level lies between 90% and 95% (about 93.6%)

A1

Answer: t = 2.108; between 90% and 95%

5 marks

21 (6 marks)

Differences: 4, 4, 1, 6, 3

B1

Mean = 3.6

B1

$$s^2 = \frac{1}{4}(16 + 16 + 1 + 36 + 9 - 5 \times 12.96) = 3.30, s = 1.817$$

M1 A1

$$t_4 = 2.776, 3.6 \pm 2.776 \times \frac{1.817}{\sqrt{5}}$$

M1

$$= (1.34, 5.86)$$

A1

Answer: (1.34, 5.86)

6 marks

22 (6 marks)

$$\bar{x} = \frac{150.6}{12} = 12.55$$

B1

$$s^2 = \frac{1}{11}(1905.4 - 12 \times 12.55^2)$$

M1

$$= 1.397, s = 1.182$$

A1

$$t_{11} = 2.201, \text{standard error} = 0.3412$$

B1

$$\bar{x} \pm 2.201 \times 0.3412$$

M1

$$= (11.80, 13.30)$$

A1

Answer: (11.80, 13.30)

6 marks

23 (6 marks)

$$\bar{x} = 8.850, s = 0.9247$$

B1 M1

$$\text{Standard error} = 0.3775, t_5 = 2.571$$

B1

$$t\text{-interval: } (7.88, 9.82)$$

A1

$$z\text{-interval: } (8.11, 9.59)$$

A1

The z-interval is too narrow for such a small sample, as it ignores the extra uncertainty from estimating σ

E1

Answer: (7.88, 9.82)

6 marks

24 (6 marks)

$$(a) \text{ Standard error} = \sqrt{0.0475/40} = 0.03446$$

M1

$$0.05 \pm 0.0675 = (-0.0175, 0.1175)$$

A1

A proportion cannot be negative

E1

The normal approximation is poor since $n\hat{p} = 2 < 5$

E1

Answer: (-0.0175, 0.1175)

4 marks

(b) Both $n\hat{p}$ and $n(1 - \hat{p})$ should be at least about 5 to 10

B1

$$n \geq \frac{10}{0.05} = 200$$

B1

Answer: n = 200

2 marks

25 (7 marks)

$$s_p^2 = \frac{7 \times 2.1^2 + 9 \times 2.6^2}{16}$$

M1

$$= \frac{30.87 + 60.84}{16} = 5.732$$

A1

$$\text{Standard error} = \sqrt{s_p^2 \left(\frac{1}{8} + \frac{1}{10} \right)} = 1.136$$

M1 A1

$$v = 16, t = 2.12$$

B1

$$2.7 \pm 2.12 \times 1.136$$

M1

$$= (0.29, 5.11)$$

A1

Answer: (0.29, 5.11)

7 marks

26 (6 marks)

$$\text{Difference} = 0.08$$

B1

$$\text{Standard error} = \sqrt{\frac{0.46 \times 0.54}{250} + \frac{0.38 \times 0.62}{300}}$$

M1

$$= \sqrt{0.0009936 + 0.0007853} = 0.04218$$

A1

$$0.08 \pm 1.96 \times 0.04218$$

M1

$$= (-0.0027, 0.1627)$$

A1

The interval includes 0 (just), so the evidence of an improvement is weak

E1

Answer: (-0.0027, 0.1627)

6 marks

27 (6 marks)

$$\text{Margin} = t_{n-1} \frac{s}{\sqrt{n}} \leq 0.5s, \text{ so } \frac{t_{n-1}}{\sqrt{n}} \leq 0.5$$

M1 A1

$$n = 17: \frac{2.12}{\sqrt{17}} = 0.514 > 0.5$$

M1 A1

$$n = 18: \frac{2.11}{\sqrt{18}} = 0.497 \leq 0.5$$

M1

$$n = 18$$

A1

Answer: n = 18

6 marks

28 (9 marks)

$$\text{(a) Standard error} = \frac{6.2}{\sqrt{15}} = 1.601$$

B1

$$t_{14} = 2.145$$

B1

$$40.8 \pm 2.145 \times 1.601$$

M1

$$= (37.37, 44.23)$$

A1

Answer: (37.37, 44.23)

4 marks

$$\text{(b) } 1.96 \times \frac{6.2}{\sqrt{n}} \leq 1$$

M1

$$\sqrt{n} \geq 12.15$$

A1

$$n \geq 147.7, \text{ so } n = 148$$

A1

Answer: n = 148

3 marks

(c) s from a small pilot is itself uncertain

E1

If σ is underestimated the interval will be wider than 2 units

E1

Answer: s is only an estimate

2 marks

29 (10 marks)

$$\text{(a) Standard error} = \sqrt{\frac{0.46 \times 0.54}{400}} = 0.02492$$

M1 A1

$$0.46 \pm 1.96 \times 0.02492 = (0.4112, 0.5088)$$

A1

Answer: (0.4112, 0.5088)

3 marks

(b) Total in favour: $184 + 312 + 490 = 986$ out of 2000

M1 A1

$$\hat{p} = 0.493$$

A1

$$\text{Standard error} = \sqrt{\frac{0.493 \times 0.507}{2000}} = 0.01118$$

M1

$$0.493 \pm 1.96 \times 0.01118 = (0.4711, 0.5149)$$

Answer: (0.4711, 0.5149)

A1

5 marks

(c) Combined width 0.0438 is less than the first width 0.0977

B1

The combined interval contains 0.5, so a majority is not established

E1

Answer: Narrower; majority not established

2 marks

30 (12 marks)

(a) Differences: 1.3, 1.6, 1.1, 0.8, 1.7, 0.7; mean = 1.200

B1 M1

$$s_d = 0.4099$$

M1

$$t_5 = 2.571, \text{ standard error} = 0.1673$$

B1

Interval: (0.770, 1.630)

A1

Answer: (0.770, 1.630)

5 marks

(b) Pooled $s_p^2 = \frac{5(3.131)^2 + 5(2.866)^2}{10}$ so $s_p = 3.002$

M1 A1

$$\text{Standard error} = s_p \sqrt{\frac{1}{6} + \frac{1}{6}} = 1.733$$

M1

$$t_{10} = 2.228$$

B1

$$\text{Interval: } 1.200 \pm 2.228 \times 1.733 = (-2.661, 5.061)$$

A1

Answer: (-2.661, 5.061)

5 marks

(c) The paired interval (width 0.860) is much narrower than the unpaired one (width 7.72)

B1

The plants differ greatly in height; pairing removes this between-plant variation and isolates the effect of the treatment

E1

Answer: Paired is narrower

2 marks

Estimation of Variance

1 (2 marks)

It has a chi-squared distribution

B1

with $n - 1$ degrees of freedom: χ^2_{n-1}

B1

Answer: χ^2_{n-1}

2 marks

2 (2 marks)

$$E(W) = \nu = 5$$

B1

$$\text{Var}(W) = 2\nu = 10$$

B1

Answer: 5 and 10

2 marks

3 (3 marks)

$$\nu = 9$$

B1

Upper critical value (2.5% in the upper tail): 19.023

B1

Lower critical value (2.5% in the lower tail): 2.7

B1

Answer: 2.7 and 19.023

3 marks

4 (2 marks)

$$n - 1 = 8$$

B1

$$8 \times 12.5 = 100$$

B1

Answer: 100

2 marks

5 (3 marks)

$$\bar{x} = 7$$

B1

$$\Sigma(x - \bar{x})^2 = 312 - 6 \times 49 = 18$$

M1

$$s^2 = \frac{18}{5} = 3.6$$

A1

Answer: 3.6

3 marks

6 (1 mark)

The population is normally distributed

B1

Answer: Normal population

1 mark

7 (4 marks)

$$(n - 1)s^2 = 9 \times 8.4 = 75.6$$

B1

 χ^2_9 : upper 2.5% point 19.023, lower 2.5% point 2.7

B1

$$\text{Interval: } \left(\frac{75.6}{19.023}, \frac{75.6}{2.7} \right)$$

M1

$$= (3.97, 28.00)$$

A1

Answer: (3.97, 28.00)

4 marks

8 (2 marks)

Take square roots of the limits

M1

$$(\sqrt{3.97}, \sqrt{28.0}) = (1.99, 5.29)$$

A1

Answer: (1.99, 5.29)

2 marks

9 (4 marks)

$$\nu = 15, (n - 1)s^2 = 375$$

B1

5% points: upper 24.996, lower 7.261

B1

$$\text{Interval: } \left(\frac{375}{24.996}, \frac{375}{7.261} \right)$$

M1

$$= (15.00, 51.65)$$

A1

Answer: (15.00, 51.65)

4 marks

10 (4 marks)

$$\frac{(n - 1)S^2}{\sigma^2} \sim \chi^2_{10}$$

B1

$$P(S^2 > 6.4) = P(\chi^2_{10} > \frac{10 \times 6.4}{4})$$

M1

$$= P(\chi^2_{10} > 16)$$

A1

$$\approx 0.0996$$

A1

Answer: 0.0996

4 marks

11 (6 marks)

$$\bar{x} = 15.17$$

B1

$$\Sigma(x - \bar{x})^2 = 5.395, s^2 = 0.7707$$

M1 A1

 $\nu = 7$: upper 16.013, lower 1.69

B1

$$\text{Interval: } \left(\frac{5.395}{16.013}, \frac{5.395}{1.69} \right)$$

M1

$$= (0.337, 3.192)$$

A1

Answer: (0.337, 3.192)

6 marks

12 (5 marks)

$$\nu = 20, (n - 1)s^2 = 150$$

B1

Points: upper 34.17, lower 9.591

B1

$$\text{Interval: } (4.39, 15.64)$$

M1 A1

9 lies inside the interval, so the data are consistent with the claim

E1

Answer: (4.39, 15.64)

5 marks

13 (4 marks)

$$\nu = 7, (n - 1)s^2 = 7 \times 6.25 = 43.75$$

B1

Points: upper 16.013, lower 1.69

M1

$$\text{Interval for } \sigma^2: (2.73, 25.89)$$

A1

$$\text{Interval for } \sigma: (1.65, 5.09)$$

A1

Answer: (1.65, 5.09)

4 marks

14 (4 marks)

$$P(\chi^2_{11} > 4.575) = 0.95$$

B1

$$\frac{11 \times 6}{\sigma^2} > 4.575 \text{ gives } \sigma^2 < \frac{66}{4.575}$$

M1 A1

$$U = 14.43, \text{ interval } (0, 14.43)$$

A1

Answer: (0, 14.43)

4 marks

15 (5 marks)

$$\frac{20S^2}{\sigma^2} \sim \chi^2_{20}$$

B1

$$P(0.6\sigma^2 < S^2 < 1.5\sigma^2) = P(12 < \chi^2_{20} < 30)$$

M1 A1

$$= 0.9301 - 0.08392$$

M1

$$\approx 0.846$$

A1

Answer: 0.846

5 marks

16 (5 marks)

$$\nu = 19, (n - 1)s^2 = 19 \times 0.1764 = 3.3516$$

B1

Points: upper 32.852, lower 8.907

B1

Interval for σ^2 : (0.1020, 0.3763)

M1

Interval for σ : (0.319, 0.613)

A1

The upper limit exceeds 0.5, so the requirement cannot be guaranteed

E1

Answer: (0.319, 0.613)

5 marks

17 (6 marks)

$$\nu = 17, (n - 1)s^2 = 17 \times 27.04 = 459.68$$

B1

5% points: upper 27.587, lower 8.672

B1

$$\text{Interval for } \sigma^2: \left(\frac{459.68}{27.587}, \frac{459.68}{8.672} \right) = (16.66, 53.01)$$

M1 A1

Interval for σ : (4.08, 7.28)

A1

4 lies inside the interval, so $\sigma = 4$ is plausible

E1

Answer: (4.08, 7.28)

6 marks

18 (5 marks)Let a and b be the lower and upper 2.5% points of χ^2_{n-1} : $P(a < \frac{(n-1)S^2}{\sigma^2} < b) = 0.95$

B1 M1

$$\text{Invert each part: } \frac{1}{b} < \frac{\sigma^2}{(n-1)S^2} < \frac{1}{a}$$

M1

$$P\left(\frac{(n-1)S^2}{b} < \sigma^2 < \frac{(n-1)S^2}{a}\right) = 0.95$$

A1

Taking reciprocals reverses the inequalities, so the larger critical value b gives the lower limit

E1

Answer: $(n-1)s^2/b < \sigma^2 < (n-1)s^2/a$

5 marks

19 (5 marks)

$$\text{Lower limit} = \frac{14s^2}{26.119} = 5.360$$

M1

$$s^2 = \frac{5.360 \times 26.119}{14} = 10.0$$

M1 A1

$$\text{Upper limit} = \frac{14 \times 10.0}{5.629}$$

M1

$$= 24.87$$

A1

Answer: $s^2 = 10$; upper limit 24.87

5 marks

20 (5 marks)

The larger critical value must give the lower limit, so the limits have been swapped (the lower limit exceeds the upper)

E1

Correct interval: $(\frac{75.6}{19.023}, \frac{75.6}{2.700})$

M1

= (3.97, 28.0)

A1

The chi-squared distribution is not symmetrical (skewed to the right)

E1

so the interval is wider above s^2 than below it

E1

Answer: (3.97, 28.0)

5 marks

21 (6 marks)

$\bar{x} = 5.467$, $\Sigma(x - \bar{x})^2 = 1.133$, $s^2 = 0.2267$

M1 A1

$\nu = 5$, 5% points: upper 11.07, lower 1.145

B1

Interval for σ^2 : (0.102, 0.990)

M1 A1

Interval for σ : (0.320, 0.995)

A1

Answer: (0.320, 0.995)

6 marks

22 (6 marks)

$\Sigma(x - \bar{x})^2 = 2262.7 - \frac{182.4^2}{15} = 44.72$

M1 A1

$s^2 = \frac{44.72}{14} = 3.194$

A1

$\nu = 14$: upper 26.119, lower 5.629

B1

Interval: $(\frac{44.72}{26.119}, \frac{44.72}{5.629})$

M1

= (1.71, 7.94)

A1

Answer: (1.71, 7.94)

6 marks

23 (6 marks)

$n = 10$: $\nu = 9$, $9 \times 16 = 144$; points 19.023 and 2.7

B1 M1

Interval: (7.57, 53.33)

A1

$n = 40$: $\nu = 39$, $39 \times 16 = 624$; points 58.12 and 23.654

B1

Interval: (10.74, 26.38)

A1

The interval from the larger sample is much narrower: more observations give a more precise estimate

E1

Answer: (7.57, 53.33) and (10.74, 26.38)

6 marks

24 (7 marks)

$S^2 = \frac{\sigma^2}{n-1} \times W$ where $W \sim \chi^2_{n-1}$

M1

$\text{Var}(S^2) = \frac{\sigma^4}{(n-1)^2} \times 2(n-1) = \frac{2\sigma^4}{n-1}$

M1 A1

$n = 26$: $\text{Var}(S^2) = \frac{2 \times 81}{25} = 6.48$

A1

Standard deviation = 2.546

A1

$E(S^2) = 9$; $P(S^2 > 12) \approx P(Z > \frac{3}{2.546}) = P(Z > 1.178)$

M1

≈ 0.119

A1

Answer: 0.119

7 marks

25 (7 marks)

$$\text{Pooled estimate: } s_p^2 = \frac{9 \times 4.2 + 14 \times 5.1}{23}$$

M1

$$= \frac{37.8 + 71.4}{23} = 4.748$$

A1

$$\nu = 9 + 14 = 23$$

B1

Points: upper 38.076, lower 11.689

B1 B1

$$\text{Interval: } \left(\frac{23 \times 4.748}{38.076}, \frac{23 \times 4.748}{11.689} \right)$$

M1

$$= (2.87, 9.34)$$

A1

Answer: (2.87, 9.34)

7 marks

26 (6 marks)

$$\Sigma(x - 50)^2 = 17.66$$

M1 A1

Points for $\nu = 10$: upper 20.483, lower 3.247

B1

$$\text{Interval: } \left(\frac{17.66}{20.483}, \frac{17.66}{3.247} \right) = (0.86, 5.44)$$

M1 A1

No degree of freedom is lost because μ is known and not estimated by $\bar{x}$

E1

Answer: (0.86, 5.44)

6 marks

27 (7 marks)

$$\text{Interval contains } \sigma^2 \text{ if } \frac{9s^2}{19.023} < \sigma^2 < \frac{9s^2}{2.700}$$

M1

$$\text{i.e. } 2.700 < \frac{9s^2}{\sigma^2} < 19.023$$

A1

$$\text{With } W = \frac{9s^2}{\sigma^2} \sim \chi^2_9: P(2.700 < W < 19.023) = 0.95$$

A1

$$\text{Upper limit } < 2\sigma^2 \text{ means } \frac{9s^2}{2.700} < 2\sigma^2$$

M1

$$\text{i.e. } W < 5.4$$

A1

$$P(\chi^2_9 < 5.4) = 0.202$$

M1 A1

Answer: 0.202

7 marks

28 (9 marks)

$$\text{(a) } E(T_k) = \frac{(n-1)\sigma^2}{k}$$

B1

$$\text{Var}(T_k) = \frac{2(n-1)\sigma^4}{k^2}$$

M1 A1

$$\text{Bias} = \frac{(n-1-k)\sigma^2}{k}$$

M1

$$\text{MSE} = \frac{\sigma^4[2(n-1) + (n-1-k)^2]}{k^2}$$

A1

$$\text{Answer: } \text{MSE}(T_k) = \frac{\sigma^4[2(n-1) + (n-1-k)^2]}{k^2}$$

5 marks

$$\text{(b) } k = 9: \frac{18+0}{81} = 0.2222$$

M1 A1

$$k = 10: \frac{18+1}{100} = 0.19; k = 11: \frac{18+4}{121} = 0.1818$$

A1

The biased estimators with larger divisors have smaller mean squared error than the unbiased one

E1

Answer: 0.222, 0.190, 0.182

4 marks

29 (10 marks)

$$\text{(a) } \sqrt{(2\chi^2)} \approx \sqrt{199 \pm 1.96} = 14.107 \pm 1.96$$

M1 A1

$$\text{Upper: } \chi^2 \approx \frac{1}{2}(16.067)^2 = 129.1$$

M1

$$\text{Lower: } \chi^2 \approx \frac{1}{2}(12.147)^2 = 73.77$$

A1

Answer: 73.77 and 129.1

4 marks

(b) Interval for σ^2 : $(\frac{100 \times 12}{129.1}, \frac{100 \times 12}{73.77})$

M1 A1

= (9.30, 16.27)

A1

σ : (3.05, 4.03)

A1

Answer: (9.30, 16.27)

4 marks

(c) Approximations 129.1 and 73.77 are within about 1% of the exact values

B1

The approximation improves as ν increases

E1

Answer: Accurate to about 1%

2 marks

30 (8 marks)

Need $\frac{(n-1) \times 0.1764}{c} < 0.25$

M1 A1

$n = 20$: $\frac{19 \times 0.1764}{10.117} = 0.3313$, σ limit $0.576 > 0.5$

M1 A1

$n = 40$: $\frac{39 \times 0.1764}{25.695} = 0.2677$, σ limit $0.517 > 0.5$

M1 A1

$n = 60$: $\frac{59 \times 0.1764}{42.339} = 0.2458$, σ limit $0.496 < 0.5$

M1

Only $n = 60$ gives evidence that $\sigma < 0.5$

A1

Answer: Only $n = 60$

8 marks