1. Faults occur in a roll of material at a rate of λ per m^2. To estimate λ, three pieces of material of sizes 3 m^2, 7 m^2 and 10 m^2 are selected and the number of faults X_1, X_2 and X_3 respectively are recorded.

The estimator $\hat{\lambda}$, where

$$\hat{\lambda} = k(X_1 + X_2 + X_3)$$

is an unbiased estimator of λ.

(a) Write down the distributions of X_1, X_2 and X_3 and find the value of k.

(4)

(b) Find Var($\hat{\lambda}$).

(3)

A random sample of n pieces of this material, each of size 4 m^2, was taken. The number of faults on each piece, Y, was recorded.

(c) Show that $\frac{1}{4}\overline{Y}$ is an unbiased estimator of λ.

(2)

(d) Find Var($\frac{1}{4}\overline{Y}$).

(3)

(e) Find the minimum value of n for which $\frac{1}{4}\overline{Y}$ becomes a better estimator of λ than $\hat{\lambda}$.

(2)

(Total 14 marks)

2. A continuous uniform distribution on the interval $[0, k]$ has mean $\frac{k}{2}$ and variance $\frac{k^2}{12}$. A random sample of three independent variables X_1, X_2 and X_3 is taken from this distribution.

(a) Show that $\frac{2}{3}X_1 + \frac{1}{2}X_2 + \frac{5}{6}X_3$ is an unbiased estimator for k.

(3)

An unbiased estimator for k is given by $\hat{k} = aX_1 + bX_2$ where a and b are constants.

(b) Show that $\text{Var}(\hat{k}) = (a^2 - 2a + 2)\frac{k^2}{6}$

(6)

(c) Hence determine the value of a and the value of b for which $\hat{k}$ has minimum variance, and calculate this minimum variance.

(6)
(Total 15 marks)

3. A random sample $X_1, X_2, ..., X_{10}$ is taken from a population with mean μ and variance σ^2.

(a) Determine the bias, if any, of each of the following estimators of μ.

$$\theta_1 = \frac{X_3 + X_4 + X_5}{3},$$

$$\theta_2 = \frac{X_{10} - X_1}{3},$$

$$\theta_3 = \frac{3X_1 + 2X_2 + X_{10}}{6}.$$

(4)

(b) Find the variance of each of these estimators.

(5)

(c) State, giving reasons, which of these three estimators for μ is

(i) the best estimator,

(ii) the worst estimator.

(4)
(Total 13 marks)

4. The value of orders, in £, made to a firm over the internet has distribution $N(\mu, \sigma^2)$. A random sample of n orders is taken and $\bar{X}$ denotes the sample mean.

(a) Write down the mean and variance of $\bar{X}$ in terms of μ and σ^2.

(2)

A second sample of m orders is taken and $\bar{Y}$ denotes the mean of this sample. An estimator of the population mean is given by

$$U = \frac{n\bar{X} + m\bar{Y}}{n+m}.$$

(b) Show that U is an unbiased estimator for μ.

(3)

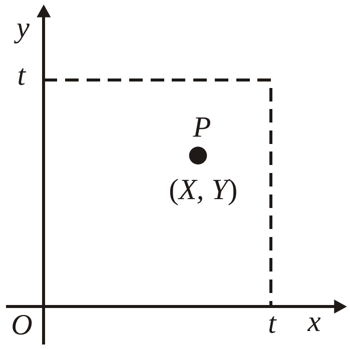
(c) Show that the variance of U is $\dfrac{\sigma^2}{n+m}$.

(4)

(d) State which of $\bar{X}$ or U is a better estimator for μ. Give a reason for your answer.

(2)
(Total 11 marks)

5.



The figure above shows a square of side t and area t^2 which lies in the first quadrant with one vertex at the origin. A point P with coordinates (X, Y) is selected at random inside the square and the coordinates are used to estimate t^2. It is assumed that X and Y are independent random variables each having a continuous uniform distribution over the interval $[0, t]$.

[You may assume that $E(X^nY^n) = E(X^n)E(Y^n)$, where n is a positive integer]

(a) Use integration to show that $E(X^n) = \dfrac{t^n}{n+1}$.

(3)

The random variable $S = kXY$, where k is a constant, is an unbiased estimator for t^2.

(b) Find the value of k.

(3)

(c) Show that $\text{Var}(S) = \dfrac{7t^4}{9}$.

(3)

The random variable $U = q(X^2 + Y^2)$, where q is a constant, is also an unbiased estimator for t^2.

(d) Show that the value of $q = \dfrac{3}{2}$.

(3)

(e) Find Var(U).

(3)

(f) State, giving a reason, which of S and U is the better estimator of t^2.

(1)

The point (2, 3) is selected from inside the square.

(g) Use the estimator chosen in part (f) to find an estimate for the area of the square.

(1)
(Total 17 marks)

6. A diabetic patient records her blood glucose readings in mmol/l at random times of day over several days. Her readings are given below.

5.3 5.7 8.4 8.7 6.3 8.0 7.2

Assuming that the blood glucose readings are normally distributed calculate

(a) an unbiased estimate for the variance σ^2 of the blood glucose readings,

(2)

(b) a 90% confidence interval for the variance σ^2 of blood glucose readings.

(5)

(c) State whether or not the confidence interval supports the assertion that $\sigma = 0.9$.

Give a reason for your answer.

(1)
(Total 8 marks)

7. A population has mean μ and variance σ^2.
A random sample of size 3 is to be taken from this population and $\overline{X}$ denotes its sample mean.
A second random sample of size 4 is to be taken from this population and $\overline{Y}$ denotes its sample mean.

(a) Show that unbiased estimators for μ are given by

(i) $\hat{\mu}_1 = \frac{1}{3}\overline{X} + \frac{2}{3}\overline{Y},$

(2)

(ii) $\hat{\mu}_2 = \frac{5\overline{X} + 4\overline{Y}}{9}$

(2)

(b) Calculate Var ($\hat{\mu}_1$)

(2)

(c) Given that Var $(\hat{\mu}_2) = \frac{37}{243}\sigma^2$, state, giving a reason, which of these two estimators should be used.

(1)
(Total 7 marks)

8. A bag contains marbles of which an unknown proportion p is red. A random sample of n marbles is drawn, with replacement, from the bag. The number X of red marbles drawn is noted.

A second random sample of m marbles is drawn, with replacement. The number Y of red marbles drawn is noted.

Given that $p_1 = \frac{aX}{n} + \frac{bY}{m}$ is an unbiased estimator of p,

(a) show that $a + b = 1$.

(4)

Given that $p_2 = \dfrac{(X+Y)}{n+m}$,

(b) show that p_2 is an unbiased estimator for p.

(3)

(c) Show that the variance of p_1 is $p(1-p)\left(\dfrac{a^2}{n}+\dfrac{b^2}{m}\right)$.

(3)

(d) Find the variance of p_2.

(3)

(e) Given that $a = 0.4$, $m = 10$ and $n = 20$, explain which estimator p_1 or p_2 you should use.

(4)
(Total 17 marks)

9. (a) Explain briefly what you understand by

(i) an unbiased estimator,

(ii) a consistent estimator.

of an unknown population parameter θ.

(3)

From a binomial population, in which the proportion of successes is p, 3 samples of size n are taken. The number of successes X_1, X_2, and X_3 are recorded and used to estimate p.

(b) Determine the bias, if any, of each of the following estimators of p.

$$\hat{p}_1 = \frac{X_1 + X_2 + X_3}{3n},$$

$$\hat{p}_2 = \frac{X_1 + 3X_2 + X_3}{6n},$$

$$\hat{p}_3 = \frac{2X_1 + 3X_2 + X_3}{6n}.$$

(4)

(c) Find the variance of each of these estimators.

(4)

(d) State, giving a reason, which of the three estimators for p is

(i) the best estimator,

(ii) the worst estimator.

(4)
(Total 15 marks)

10. A random sample of three independent variables X_1, X_2 and X_3 is taken from a distribution with mean μ and variance σ^2.

(a) Show that $\frac{2}{3}X_1 - \frac{1}{2}X_2 + \frac{5}{6}X_3$ is an unbiased estimator for μ.

(3)

An unbiased estimator for μ is given by $\hat{\mu} = aX_1 + bX_2$ where a and b are constants.

(b) Show that Var ($\hat{\mu}$) = $(2a^2 - 2a + 1)\sigma^2$.

(6)

(c) Hence determine the value of a and the value of b for which $\hat{\mu}$ has minimum variance.

(5)
(Total 14 marks)

1. (a) $X_1 \sim \text{Po}(3\lambda)$

$X_2 \sim \text{Po}(7\lambda)$ M1

$X_3 \sim \text{Po}(10\lambda)$

$\text{E}(\hat{\lambda}) = k\,[\text{E}(X_1) + \text{E}(X_2) + \text{E}(X_3)]$ M1

$= 20\lambda k$

$\hat{\lambda}$ unbiased therefore $20\lambda k = \lambda$ M1

$k = \dfrac{1}{20}$ A1 4

Note

M1 all 3 needed. Poisson and mean
M1 adding their means
M1 putting their $\text{E}(\hat{\lambda}) = \lambda$
A1 cao

(b) $\text{Var}(\hat{\lambda}) = \dfrac{1}{20^2}\text{Var}(X_1 + X_2 + X_3)$ M1

$= \dfrac{1}{20^2}(3\lambda + 7\lambda + 10\lambda)$ M1

$= \dfrac{\lambda}{20}$ A1ft 3

Note

M1 use of $k^2\text{Var}(X_1 + X_2 + X_3)$
M1 using their means from part(a) as Variances and adding together
A1 cao

(c) $Y \sim \text{Po}(4\lambda)$

$\text{E}\left(\dfrac{1}{4}\overline{Y}\right) = \dfrac{1}{4} \times 4\lambda = \lambda$ therefore unbiased M1 A1 2

Note

M1 use of 4λ
A1 cso plus conclusion. Accept working out bias to = 0

	(d)	$\text{Var}\left(\frac{1}{4}\bar{Y}\right) = \frac{1}{16}\times\frac{4\lambda}{n}$	M1 B1	
		$= \frac{\lambda}{4n}$	A1	3

Note

M1 $\frac{1}{16} \times \text{Var}\,\bar{Y}$

B1 for $\text{Var}\,\bar{Y} = \frac{4\lambda}{n}$

A1 cao

	(e)	$\frac{\lambda}{4n} < \frac{\lambda}{20}$	M1	
		$n > 5$ therefore $n = 6$	A1	2

Note

M1 for $\text{Var}\left(\frac{1}{4}\bar{Y}\right) > \text{Var}(\hat{\lambda})$

A1 n = 6

[14]

2.	(a)	$\text{E}\left(\tfrac{2}{3}X_1 + \tfrac{1}{2}X_2 + \tfrac{5}{6}X_3\right) = \tfrac{2}{3}\times\tfrac{k}{2} + \tfrac{1}{2}\times\tfrac{k}{2} + \tfrac{5}{6}\times\tfrac{k}{2} = k$	M1 A1	
		$\text{E}(X_1 + X_2 + X_3) = \text{k} \Rightarrow$ **unbiased**	B1	3
	(b)	$\text{E}(aX_1 + bX_2) = a\frac{k}{2} + b\frac{k}{2} = k$	M1	
		$a + b = 2$	A1	
		$\text{Var}(aX_1 + bX_2) = a^2\frac{k^2}{12} + b^2\frac{k^2}{12}$	M1A1	
		$= a^2\frac{k^2}{12} + (2-a)^2\frac{k^2}{12}$	M1	
		$= (2a^2 - 4a + 4)\frac{k^2}{12}$		
		$= (2a^2 - 2a + 4)\frac{k^2}{6}$ (*) since answer given	A1 cso	6

(c) Min value when $(2a-2)\frac{k^2}{6} = 0$

$\frac{\mathrm{d}}{\mathrm{da}}(\mathrm{Var}) = 0,$ all correct, condone missing $\frac{k^2}{6}$ M1A1

$\Rightarrow 2a–2=0$ A1A1

$A = 1$, $b = 1$.

$\frac{\mathrm{d}^2(\mathrm{Var})}{\mathrm{da}^2} = \frac{2k^2}{6 > 0}$ since $k^2 > 0$ therefore it is a minimum M1

min variance $= (1-2+2)\frac{k^2}{6}$

$= \frac{k^2}{6}$ B1 6

<u>Alternative</u>

$\frac{k^2}{6}(a-1)^2 - \frac{k^2}{6} + \frac{2k^2}{6}$ M1 A1

$\frac{k^2}{6}(a-1)^2 + \frac{k^2}{6}$ M1

Min when $\frac{k^2}{6}(a-1)^2 = 0$ A1A1

$a = 1$ $b = 1$

min var $= k^2/6$ B1

[15]

3. (a) $E(\theta_1) = \dfrac{E(X_3) + E(X_4) + E(X_5)}{3}$

$= \dfrac{3\mu}{3}$

$= \mu$ Bias = 0 allow unbiased B1

$E(\theta_2) = \dfrac{E(X_{10}) - E(X_1)}{3}$

= 1/3

= 0 Bias = $-\mu$ allow $\pm\mu$ B1, B1

$E(\theta_3) = \dfrac{3E(X_1) + 2E(X_2) + E(X_{10})}{6}$

$= \dfrac{3\mu + 2\mu + \mu}{6}$

$= \mu$ Bias = 0 allow unbiased B1 4

(b) $\text{Var}(\theta_1) = \dfrac{1}{9}\{(\text{Var } X_2) + \text{Var}(X_3) + \text{Var}(X_4)\}$ M1

$= \dfrac{1}{9}\{\sigma^2 + \sigma^2 + \sigma^2\}$

$= \dfrac{1}{3}\sigma^2$ A1

$\text{Var}(\theta_2) = \dfrac{2}{9}\sigma^2$ B1

$\text{Var}(\theta_3) = \dfrac{1}{36}\{9\sigma^2 + 4\sigma^2 + \sigma^2\}$ M1

$= \dfrac{7}{18}\sigma^2$ A1 5

(c) (i) θ_1 is the better estimator. It has a lower var. and no bias B1depB1

(ii) θ_2 is the worst estimator. It is biased B1depB1 4

[13]

4. (a) $E(\overline{X}) = \mu$ B1

$\text{Var}(\overline{X}) = \text{Var}\left(\dfrac{X_1 + X_2 + X_3 + \ldots + X_n}{n}\right)$

$= \dfrac{\sigma^2}{n}$ B1 2

(b)	$\mathrm{E}(U) = \dfrac{1}{n+m}(n\mathrm{E}(\overline{X}) + m\mathrm{E}(\overline{Y}))$		M1	
	$= \dfrac{1}{n+m}(n\mu + m\mu)$		A1	
	$= \mu \Rightarrow$ <u>U is unbiased</u>	state unbiased	A1	3
(c)	$\mathrm{Var}(\overline{Y}) = \dfrac{\sigma^2}{m}$		B1	
	$\mathrm{Var}(U) = \dfrac{n^2\mathrm{Var}(\overline{X}) + m^2\mathrm{Var}(\overline{Y})}{(n+m)^2}$		M1	
	$= \dfrac{n^2\dfrac{\sigma^2}{n} + m^2\dfrac{\sigma^2}{m}}{(n+m)^2}$		A1	
	$= \dfrac{n\sigma^2 + m\sigma^2}{(n+m)^2}$			
	$= \dfrac{\sigma^2}{n+m}$ *	cso	A1	4
(d)	$\dfrac{n\overline{X} + m\overline{Y}}{n+m}$ is a better estimate since variance is smaller		B1B1	2

[11]

5.	(a)	$\int x^n \dfrac{1}{t}$		M1	
		$\int_0^t dx$		M1	
		$\mathrm{E}(x^n) = \int_0^t x^n \dfrac{1}{t}dx = \left[\dfrac{x^{n+1}}{t(n+1)}\right]_0^t = \left(\dfrac{t^{n+1}}{t(n+1)} - 0\right)\dfrac{t^n}{n+1}$		A1c.s.o.	3
	(b)	$\left(\mathrm{E}(x) = \dfrac{t}{2}\right)$ <u>$\mathrm{E}(s) = k\mathrm{E}(x)\mathrm{E}(y),$</u> $k.\dfrac{t^2}{4}$		<u>M1</u>, A1	
		$\mathrm{E}(s) = t^2 \quad \Rightarrow k = 4$		A1	3

(c) $\text{Var}(xy) = \text{E}(x^2)\text{E}(y^2) - [\text{E}(xy)]^2$ M1

$= \frac{t^2}{3} \times \frac{t^2}{3} - \left(\frac{t^2}{4}\right)^2 = \left\{\frac{7t^4}{144}\right\}$ M1

$\text{Var}(s) = k^2\text{var}(xy) = 16 \times \frac{7t^4}{144} = \frac{7t^4}{9}$ A1 c.s.o. 3

(d) $\text{E}(u) = t^2 \quad \Rightarrow 2\text{E}(x^2)q = t^2, \quad \Rightarrow \frac{2t^2}{8}q = t^2, \quad \Rightarrow q = \frac{3}{2}$

M1, M1, A1 c.s.o. 3

(e) $\text{Var}(u) = q^2\left[\text{var}(x^2)+\text{var}(y^2)\right] = 2q^2\text{var}(x^2)$ M1

$\text{Var}(x^2) = \text{E}(x4) - \left[\text{E}(x^2)\right]^2 = \frac{t^4}{5} - \left(\frac{t^2}{3}\right)^2 = \left(\frac{4}{45}t^4\right)$ M1

$\text{Var}(u) = 2 \times \frac{9}{4} \times \frac{4}{45}t^4 = \frac{2}{5}t^4$ A1 3

(f) $\frac{2}{5} < \frac{7}{9} \quad \therefore\ u$ is better $\because$ smaller variance B1ft 1

(g) Using u estimate is $\frac{3}{2}\left(2^2 + 3^2\right) = \frac{3}{2} \times 13 = \frac{39}{2}$ or 19.5 B1ft 1

[17]

6. (a) $\Sigma x = 49.6$; $\Sigma x^2 = 362.36$

$s^2 = \frac{1}{6}(362.36 - \frac{49.6^2}{7}) = 1.8180952 \ldots$ awrt 1.82 M1 A1 2

(b) CI = $\left(\frac{6 \times 1.818..}{12.592}, \frac{6 \times 1.818..}{1.635}\right)$ M1

12.592, 1.635 B1 B1

= (0.866 ... , 6.67...) awrt (0866, 6.67) A1 A1 5

(c) $0.9^2 < 0.866$, interval does not support $\sigma = 0.9$ as out of range. B1 1

7. (a) (i) $E\left(\frac{1}{3}\overline{X} + \frac{2}{3}\overline{Y}\right) = \frac{1}{3}E\left(\frac{X_1 + X_2 + X_3}{3}\right) + \frac{2}{3}E\left(\frac{Y_1 + Y_2 + Y_3 + Y_4}{4}\right)$ M1

$= \frac{1}{3}\mu + \frac{2}{3}\mu$

$= \mu$ therefore unbiased estimator A1 2

(ii) $E\left(\frac{5\overline{X} + 4\overline{Y}}{9}\right) = \frac{1}{9}(5E(\overline{X}) + 4E(\overline{Y}))$

$= \frac{1}{9}(5\mu + 4\mu)$ M1 A1

$= \mu$ therefore unbiased estimator A1 2

(b) $\text{Var}(\overline{X}) = \frac{\sigma^2}{3}$, $\text{Var}(\overline{Y}) = \frac{\sigma^2}{4}$

$\text{Var}(\hat{\mu}_1) = \frac{1}{9}.\frac{\sigma^2}{3} + \frac{4}{9}.\frac{\sigma^2}{4} = \frac{4\sigma^2}{27}$ M1 A1 2

(c) $\frac{4}{27}\sigma^2 < \frac{37}{243}\sigma^2$ so use $\hat{\mu}_1$. B1 1

[7]

8. (a) $E(X) = np$, $E(Y) = mp$ B1

both; can be implied

$E(p_1) = \frac{aE(X)}{n} + \frac{bE(Y)}{m} = p; \Rightarrow \frac{anp}{n} + \frac{bmp}{m} = p$ M1 A1

$\Rightarrow$ $\underline{(a + b) = 1}$ A1 4

(b) $E(p_2) = \frac{1}{n + m}\{E(X) + E(Y)\}$ M1

$= \frac{1}{n + m}\{np + mp\}$ A1

$= \frac{1}{n + m}.p(n + m) = p \Rightarrow$ $\underline{p_2 \text{ is unbiased}}$ A1 3

	(c)	Var(X) = $np(1-p)$; Var (Y) = $mp(1-p)$ *both; can be implied*	B1	
		Var(p_1) = $\dfrac{a^2\text{Var}(X)}{n^2} + \dfrac{b^2\text{Var}(Y)}{m^2}$ *Use of Var(ax) = a^2Var(x)*	M1	
		$= \dfrac{a^2np(1-p)}{n^2} + \dfrac{b^2mp(1-p)}{m^2}$ $= \underline{p(1-p)\left\{\dfrac{a^2}{n} + \dfrac{b^2}{m}\right\}}$	A1	3
	(d)	Var(p_2) = $\dfrac{1}{(n+m)^2}\{np(1-p) + mp(1-p)\}$	M1 A1	
		$= \underline{\dfrac{p(1-p)}{n+m}}$	A1	3
	(e)	Var(p_1) = $0.044p(1-p)$; Var(p_2) = $0.033\dot{3}\,p(1-p)$	B1; B1	
		Use Var(p_2) < Var(p_1); p_2	depB1; B1	4

[17]

9.	(a)	(i) E ($\hat{\theta}$) = θ	B1	
		(ii) E ($\hat{\theta}$) = θ or E ($\hat{\theta}$) $\to \theta$	B1	
		and Var ($\hat{\theta}$) $\to 0$ as n $\to \infty$ where n is the sample size	B1	3
	(b)	$\text{E}(\hat{p}_1) = p,\ \therefore$ Bias = 0	B1	
		$\text{E}(\hat{p}_2) = \dfrac{5p}{6},\ \therefore$ Bias = $\pm\ \dfrac{1}{6}p$	B1 B1	
		$\text{E}(\hat{p}_3) = p,\ \therefore$ Bias = 0	B1	4
	(c)	$\text{Var}\ (\hat{p}_1) = \dfrac{1}{9n^2}\ \{npq + npq + npq\}$	M1	
		$= \underline{\dfrac{pq}{3n}}$	A1	
		$\text{Var}(\hat{p}_2) = \dfrac{11pq}{36n}$; $\text{Var}(\hat{p}_3) = \dfrac{7pq}{18n}$	A1 A1	4

(d)	(i) $\hat{p}_1$; unbiased and smallest variance	B1 dep; B1		
	(ii) $\hat{p}_2$; biased	B1 dep; B1	4	
				[15]

10.	(a)	$\mathrm{E}\left(\frac{2}{3}X_1 - \frac{1}{2}X_2 + \frac{5}{6}X_3\right) = \frac{2}{3}\mu - \frac{1}{2}\mu + \frac{5}{6}\mu = \mu$	M1 A1		
		$\mathrm{E}(Y) = \mu \Rightarrow$ unbiased	B1	3	
	(b)	$\mathrm{E}(aX_1 + bX_2) = a\mu + b\mu = \mu$	M1		
		$a + b = 1$	A1		
		$\mathrm{Var}(aX_1 + bX_2) = a^2\sigma^2 + b^2\sigma^2$	M1 A1		
		$= a^2\sigma^2 + (1 - a)^2\sigma^2$	M1		
		$= (2a^2 - 2a + 1)\sigma^2$ **given**	A1	6	
	(c)	Min value when $(4a - 2)\sigma^2 = 0$	M1 A1		
		$\frac{d}{da}$ *(Var) = 0, all correct*			
		$\Rightarrow 4a - 2 = 0$	A		
		$a = \frac{1}{2}$, $b = \frac{1}{2}$.	A1 A1∫	5	
					[14]

1. This question proved to be the most challenging question for many candidates. Few candidates wrote down the distributions of X_1, X_2 and X_3 in part (a) and those who tried were unable to do so accurately. The Normal and Binomial distributions were commonly seen. This aside candidates were then able to progress and gain at least two marks in part (a).

In part (b) the main error was not to use their means from part (a). Even the candidates who correctly identified the Poison introduced a variety of variances including σ^2.

In part (c) and (d) the candidates who knew $\mathrm{E}(\bar{Y}) = \mu$ and $\mathrm{Var}(\bar{Y}) = \frac{\sigma^2}{n}$ gained full marks.

In part (e) the majority of candidates used $\mathrm{Var}\left(\frac{1}{4}\bar{Y}\right) < \mathrm{Var}(\hat{\lambda})$ although it was not always clear from their working that this was the case with many writing $\mathrm{Var}\left(\frac{1}{4}\bar{Y}\right) = \mathrm{Var}(\hat{\lambda})$ and simply solving the equation.

2. This question proved to be the most challenging for many candidates. In part (a) many candidates tried to prove it was equal to $\frac{k}{2}$ and few made the concluding statement that it was unbiased.

In part (b) few candidates were able to find $a + b = 2$ and hence made little progress. Those who did find this were able to gain full marks.

In part (c) a mix of both of the given methods on the mark scheme were used. If they chose the first method the majority of candidates did not prove that it was a minimum. If they chose the second they rarely completed the square correctly choosing to leave out the $\frac{k^2}{6}$.

3. This proved to be a good starter question and most candidates gave good solutions. A minority of candidates did not state the bias.

In part (b) many candidates did not know that the Variance of X_2, X_3 etc was σ^2 not $\frac{\sigma^2}{n}$.

4. This question was generally well answered. In part (b) a minority of candidates did not draw the conclusion that U is unbiased. Part (c) caused the most problems. Many did not use $\mathrm{Var}(n\bar{Y}) = n^2\,\mathrm{Var}(\bar{X})$ and although most had part (a) correct they did not use $\mathrm{Var}(\bar{Y}) = \frac{\sigma^2}{m}$. These candidates demonstrated poor algebraic skills in an endeavour to get the correct answer, even though the answer was given

5. The structure and given answers helped many candidates here and those who attempted it were often able to pick up marks in at least parts (b), (d) and (g). Perhaps surprisingly part (a) proved to be the most challenging. There were many unconvincing attempts based on integrating x^n and then for some reason dividing by x but those who simply applied their S2 knowledge and wrote $E(X^n) = \int_0^t x^n \frac{1}{t} dt$ were usually able to complete this part and often most of the question. A common error in part (c) was to assume Var(XY)=Var(X)Var(Y) and this was perhaps the most challenging part. In part (f) the reasoning was usually sound, although the values were sometimes incorrect, and candidates who persevered to the end were often able to use the point (2, 3) in their estimator to answer part (g).

6. Part (a) was well done, but many candidates started with a confidence interval in part (b) that required a significant amount of manipulation and inevitably gave rise to errors. The incorrect comparison of 0.9 with the interval, rather than 0.81 was not unusual in part (c).

7. A reasonable attempt at the proofs were usually seen here, but the use of the fractional variance in part (b) was missed by many.

8. Although the notation of some of the candidates could have been better the norm for this question was full marks. This aspect of the specification was well applied to this question.

9. The candidates found this question very much to their liking and only a few of them did not gain most of the marks. Common errors were to omit the condition that a consistent estimator needs to be unbiased and to ignore the mean and the variance of the binomial distribution in parts (b) and (c).

10. The proof in part (a) was done well, but some candidates struggled with parts (b) and (c). The common error was to miss that part (a) was useful in part (b), so only 2 marks were scored. Some good solutions were seen for part (c), but weaker candidates were unable to cope with what they perceived to be 2 variables. Many attempted to solve an arbitrary quadratic equation and concluded that there were multiple possible solutions, usually involving 0 and 1.