1) (i)	$f(x) = \frac{1}{\theta} \quad 0 \le x \le \theta$ $E[X] = \frac{\theta}{2}$ $E[2\overline{X}] = 2E[\overline{X}] = 2E[X]$ $= \theta$ $\therefore$ unbiased	 B1 M1 A1 E1	 Write-down, or by symmetry, or by integration.	 4
(ii)	$\sum x = 2.3 \quad \therefore \bar{x} = \frac{2.3}{5} = 0.46 \quad \therefore 2\bar{x} = 0.92$ But we know $\theta \ge 1$ $\therefore$ estimator can give nonsense answers, i.e. essentially useless	B1 E1 E2	 (E1, E1)	 4
(iii)	$Y = \max\{X_i\}$, $g(y) = \frac{ny^{n-1}}{\theta^n} \quad 0 \le y \le \theta$ MSE $(kY) = E[(kY-\theta)^2] =$ $E[k^2Y^2 - 2k\theta Y + \theta^2] =$ $k^2E[Y^2] - 2k\theta E[Y] + \theta^2$ $\frac{d\text{MSE}}{dk} =$ $2kE[Y^2] - 2\theta E[Y] = 0$ for $k = \frac{\theta E[Y]}{E[Y^2]}$ $\frac{d^2\text{MSE}}{dk^2} = 2E[Y^2] > 0 \quad \therefore$ this is a minimum $E[Y] = \int_0^\theta \frac{ny^n}{\theta^n} dy = \frac{n}{\theta^n}\frac{\theta^{n+1}}{n+1} = \frac{n\theta}{n+1}$ $E[Y^2] = \int_0^\theta \frac{ny^{n+1}}{\theta^n} dy = \frac{n}{\theta^n}\frac{\theta^{n+2}}{n+2} = \frac{n\theta^2}{n+2}$ $\therefore$ minimising $k = \theta \frac{n\theta}{n+1}\frac{n+2}{n\theta^2} = \frac{n+2}{n+1}$	 M1 1 M1 M1 A1 M1 M1 A1 M1 A1 M1 A1	 BEWARE PRINTED ANSWER	 12
(iv)	With this k, kY is always greater than the sample maximum So it does not suffer from the disadvantage in part (ii)	E2 E2	(E1 E1) (E1 E1)	 4

2(i)	$G(t) = E[t^X] = \sum_{x=0}^{n} \binom{n}{x} (pt)^x (1-p)^{n-x}$ $= [(1-p)+pt]^n$ $= (q+pt)^n$	M1 2 1	Available as B2 for write-down or as 1+1 for algebra	4
(ii)	$\mu = G'(1) \quad G'(t) = np(q+pt)^{n-1}$ $G'(1) = np \times 1 = np$ $\sigma^2 = G''(1) + \mu - \mu^2$ $G''(t) = n(n-1)p^2(q+pt)^{n-2}$ $G''(1) = n(n-1)p^2$ $\therefore \sigma^2 = n^2p^2 - np^2 + np - n^2p^2$ $= -np^2 + np = npq$	1 1 1 1 M1 1		6
(iii)	$Z = \frac{X-\mu}{\sigma}$ Mean 0, Variance 1	B1	For <u>BOTH</u>	1
(iv)	$M(\theta) = G(e^\theta) = (q+pe^\theta)^n$ $Z = aX + b$ with: $a = \frac{1}{\sigma} = \frac{1}{\sqrt{npq}}$ and $b = -\frac{\mu}{\sigma} = -\sqrt{\frac{np}{q}}$ $M_Z(\theta) = e^{b\theta} M_X(a\theta)$ $\therefore M_Z(\theta) = e^{-\sqrt{\frac{np}{q}}\theta} \left(q + pe^{\frac{1}{\sqrt{npq}}\theta} \right)^n =$ $\left(qe^{-\frac{p\theta}{\sqrt{npq}}} + pe^{\frac{1-p}{\sqrt{npq}}\theta} \right)^n$	1 M1 1 1 1	BEWARE PRINTED ANSWER	5
(v)	$M_Z(\theta) = (q - \frac{qp\theta}{\sqrt{npq}} + \frac{qp^2\theta^2}{2npq} +$ terms in $n^{-3/2}$, n^{-2}, ………….. + $p + \frac{pq\theta}{\sqrt{npq}} + \frac{pq^2\theta^2}{2npq} + \cdots\cdots)^n =$ $(1 + \frac{\theta^2}{2n} + \cdots\cdots)^n \to$ $e^{\theta^2/2}$	M1 M1 1 1	For expansion of exponential terms For indication that these can be neglected as $n \to \infty$. Use of result given in question	4

(vi)	N(0,1) Because $e^{\theta^2/2}$ is the mgf of N(0,1) and the relationship between distributions and their mgfs is unique	1 E1 E1		3
(vii)	"Unstandardising", $\mathrm{N}(\mu,\sigma^2)$ ie $\mathrm{N}(np,npq)$	1	Parameters need to be given.	1

3(i)	$H_0: \mu_A = \mu_B$ $H_1: \mu_A \neq \mu_B$ Where μ_A, μ_B are the population means Test statistic $\dfrac{26.4 - 25.38}{\sqrt{\dfrac{2.45}{7} + \dfrac{1.40}{5}}} =$ $\dfrac{1.02}{\sqrt{0.63} = 0.7937} = 1.285$ Refer to N(0,1) Double-tailed 5% point is 1.96 Not significant No evidence that the population means differ	1 1 M1 M1 M1 A1 1 1 1 1	Do NOT allow $\overline{X} = \overline{Y}$ or similar Accept absence of "population" if correct notation μ is used. Hypotheses stated verbally <u>must</u> include the word "population". Numerator Denominator two separate terms correct No FT if wrong No FT if wrong	10
(ii)	CI (for $\mu_A - \mu_B$) is $1.02 \pm$ $1.645 \times$ $0.7937 =$ $1.02 \pm 1.3056 =$ (– 0.2856, 2.3256)	M1 B1 M1 A1 cao	Zero out of 4 if not N(0,1)	4
(iii)	H_0 is accepted if –1.96< test statistic < 1.96 i.e. if $-1.96 < \dfrac{\bar{x} - \bar{y}}{0.7937} < 1.96$ i.e. if $-1.556 < \bar{x} - \bar{y} < 1.556$ In fact, $\overline{X} - \overline{Y} \sim \text{N}(2, 0.7937^2)$ So we want $\text{P}(-1.556 < \text{N}(2, 0.7937^2) < 1.556) =$ $\text{P}\left(\dfrac{-1.556-2}{0.7937} < \text{N}(0,1) < \dfrac{1.556-2}{0.7937}\right) =$ $\text{P}(-4.48 < \text{N}(0,1) < -0.5594) = 0.2879$	M1 M1 A1 M1 M1 M1 A1 cao	SC1 Same wrong test can get M1,M1,A0. SC2 Use of 1.645 gets 2 out of 3. BEWARE PRINTED ANSWER Standardising	7
(iv)	Wilcoxon would give protection if assumption of Normality is wrong. Wilcoxon could not really be applied if underlying variances are indeed different. Wilcoxon would be less powerful (worse Type II error behaviour) with such small samples if Normality is correct.	E1 E1 E1		3

<table>
<tr><td>4 (i)</td><td>There might be some consistent source of plot-to-plot variation that has inflated the residual and which the design has failed to cater for.</td><td>E2</td><td>E1 – Some reference to extra variation.
E1 – Some indication of a reason.</td><td>2</td></tr>
<tr><td>(ii)</td><td>Variation between the fertilisers should be compared with experimental error.

If the residual is inflated so that it measures more than experimental error, the comparison of between - fertilisers variation with it is less likely to reach significance.</td><td>E1

E2</td><td>

(E1, E1)</td><td>3</td></tr>
<tr><td>(iii)</td><td>Randomised blocks
<table><tr><td>C
B
A
D
E</td><td>.
.
.</td><td>.
.
.</td></tr></table>
SPECIAL CASE: Latin Square $\frac{2}{4}$ (1, E1)</td><td>1

E1

E1

E1</td><td>

Blocks (strips) clearly correctly oriented w.r.t. fertiliser gradient.

All fertilisers appear in a block.

Different (random) arrangements in the blocks.</td><td>4</td></tr>
<tr><td>(iv)</td><td>Totals are: 95.0 123.2 86.8 130.2 67.4
(each from sample of size 4)
Grand total 502.6
"Correction factor" CF = $\frac{502.6^2}{20} = 12630.338$
Total SS = 13610.22 – CF = 979.882

Between fertilisers SS =
$\frac{95.0^2}{4} + \ldots + \frac{67.4^2}{4}$ – CF =
13308.07 – CF = 677.732

Residual SS (by subtraction) =
979.882 – 677.732=302.15
<table><tr><th>Source of variation</th><th>SS</th><th>df</th><th>MS</th><th>MS Ratio</th></tr><tr><td>Between fertiliser</td><td>677.732</td><td>4</td><td>169.433</td><td>8.41</td></tr><tr><td>Residual</td><td>302.15</td><td>15</td><td>20.143</td><td></td></tr><tr><td>Total</td><td>979.882</td><td>19</td><td></td><td></td></tr></table>Refer to $F_{4, 15}$
-upper 5% point is 3.06
Significant
- seems effects of fertilisers are not all the same</td><td>

M1

M1

A1

M1
M1
1,A1
1

1
1
1
1</td><td>

For correct method for any two

If each calculated SS is correct

No FT if wrong
No FT if wrong</td><td>12</td></tr>
<tr><td>(vii)</td><td>Independent N (0, σ^2 [constant])</td><td>1
1
1</td><td></td><td>3</td></tr>
</table>