

Question number	Scheme	Marks
<p>1.</p> <p>(a)</p> <p>(b)</p>	<p>Take a (simple) random sample from (mutually exclusive) groups of the population Sample sizes within strata in strict proportion to numbers in each strata in the population Advantage: More accurate estimate of variance of population mean Individual estimates for strata available Disadvantage: Difficult if strata are large Definition of strata problematic (may overlap)</p> <p>Non-random sampling from groups of the population Advantage: Representative sample can be achieved with small sample size Cheap (costs kept to a minimum) Administration relatively easy Disadvantage Not possible to estimate sampling errors due to lack of randomness Judgment of interviewer can affect choice of sample – bias OK Non-response not recorded Difficulties of defining controls e.g. social class</p>	<p>1g/1h B1 B1 Any one B1 Any one B1 B1 B1dep Any one (not quick) B1 Any one B1</p> <p>(4) (4)</p> <p style="text-align: center;">8</p>
<p>2.</p> <p>(a)</p> <p>(b)</p>	<p>$X \sim N(124, 20^2)$ or $\bar{X} \sim (124, \frac{20^2}{30})$ or assume σ^2 estimated by s^2 or CLT, vals. $\bar{x} \pm 2.5758 \frac{\sigma}{\sqrt{n}} = 124 \pm 2.5758 \frac{20}{\sqrt{30}}$ $= 124 \pm 9.405$ $= (115, 133)$</p> <p>140 is not in confidence interval Underweight apples chosen or Sample may not be representative/may be biased</p>	<p>B1,B1 B1M1A1 3 sf A1 M1 Any one A1f</p> <p>(6) (2)</p> <p style="text-align: center;">8</p>

Question number	Scheme	Marks																
3.																		
(a)	$E(X-Y)=20-10=10$	Require minus, 10 M1A1																
(b)	$\text{Var}(X-Y)=5+4=9$	Require plus, 9 M1A1																
(c)	$X-Y \sim N(10,9)$ $P(13 < X-Y < 16) = P(X-Y < 16) - P(X-Y < 13)$ $= P(Z < \frac{16-10}{3}) - P(Z < \frac{13-10}{3})$ $= P(Z < 2) - P(Z < 1)$ $= 0.9772 - 0.8413 = 0.1359$	Implied B1f Subtract M1 Standardise M1 2&1 A1 0.1359 A1																
		(2) (2) (5)																
		9																
4.	H_0 : Taking drug and catching a cold are independent (not associated) H_1 : Taking drug and catching a cold are not independent (associated) (not ditto)	B1 both B1																
	<table border="1"> <thead> <tr> <th></th> <th>Cold</th> <th>Not Cold</th> <th></th> </tr> </thead> <tbody> <tr> <td>Drug</td> <td>34 (39.5)</td> <td>66 (60.5)</td> <td>100</td> </tr> <tr> <td>Dummy</td> <td>45 (39.5)</td> <td>55 (60.5)</td> <td>100</td> </tr> <tr> <td></td> <td>79</td> <td>121</td> <td>200</td> </tr> </tbody> </table>		Cold	Not Cold		Drug	34 (39.5)	66 (60.5)	100	Dummy	45 (39.5)	55 (60.5)	100		79	121	200	All totals B1 $E = \frac{RT \times CT}{GT}$ M1A1A1
	Cold	Not Cold																
Drug	34 (39.5)	66 (60.5)	100															
Dummy	45 (39.5)	55 (60.5)	100															
	79	121	200															
	<table border="1"> <thead> <tr> <th>O</th> <th>E</th> <th>$\frac{(O - E)^2}{E}$</th> </tr> </thead> <tbody> <tr> <td>34</td> <td>39.5</td> <td>0.766</td> </tr> <tr> <td>66</td> <td>60.5</td> <td>0.5</td> </tr> <tr> <td>45</td> <td>39.5</td> <td>0.765</td> </tr> <tr> <td>55</td> <td>60.5</td> <td>0.5</td> </tr> </tbody> </table>	O	E	$\frac{(O - E)^2}{E}$	34	39.5	0.766	66	60.5	0.5	45	39.5	0.765	55	60.5	0.5		
O	E	$\frac{(O - E)^2}{E}$																
34	39.5	0.766																
66	60.5	0.5																
45	39.5	0.765																
55	60.5	0.5																
	$\sum \frac{(O - E)^2}{E} = 2.53$ (NB with Yates 2.09) attempt & add, awrt 0.766 & 0.5 twice, awrt 2.53	M1A1A1																
	$\nu = 1, \chi_1^2(5\%) = 3.841 > 2.53$	1, 3.841 B1,B1																
	No reason to believe that the chance of catching a cold is affected by taking the new drug	A1f																
		11																

Question number	Scheme	Marks																																								
5	<p>μ_a and μ_b are mean weight of population after and before closure respectively.</p> <p>$H_0 : \mu_b = \mu_a$</p> <p>$H_1 : \mu_b > \mu_a$</p> $z = \frac{10 - 8}{\sqrt{\frac{2.64^2}{100} + \frac{1.94^2}{120}}}$ <p style="text-align: right;">Fraction, denom Ok alone</p> $z = \frac{2}{\sqrt{0.1011}} = 6.29$ <p style="text-align: right;">awrt 6.29</p> <p>Critical region is $z \geq 1.6449$, $6.29 > 1.6449$ or in critical region or Reject H_0 (or $P(Z \geq 6.29) = 0, 0 < 0.05$ or z in critical region or Reject H_0 B1M1)</p> <p>There is evidence that closing the factory has reduced mean river pollution</p>	<p>B1</p> <p>B1B1</p> <p>M1A1</p> <p>M1A1</p> <p>A1</p> <p>1.6449 B1, M1</p> <p>A1J</p> <p style="text-align: right;">(11)</p> <div style="border: 1px solid black; border-radius: 50%; width: 40px; height: 40px; display: flex; align-items: center; justify-content: center; margin: 0 auto;">11</div>																																								
6 (a)	<table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <tbody> <tr> <td>A</td><td>2</td><td>5</td><td>3</td><td>7</td><td>8</td><td>1</td><td>4</td><td>6</td><td></td></tr> <tr> <td>B</td><td>3</td><td>2</td><td>6</td><td>5</td><td>7</td><td>4</td><td>1</td><td>8</td><td></td></tr> <tr> <td> d </td><td>1</td><td>3</td><td>3</td><td>2</td><td>1</td><td>3</td><td>3</td><td>2</td><td></td></tr> <tr> <td>d²</td><td>1</td><td>9</td><td>9</td><td>4</td><td>1</td><td>9</td><td>9</td><td>4</td><td>46</td></tr> </tbody> </table> $r_s = 1 - \frac{6 \times 46}{8 \times 63}$ $r_s = 0.452$	A	2	5	3	7	8	1	4	6		B	3	2	6	5	7	4	1	8		d	1	3	3	2	1	3	3	2		d ²	1	9	9	4	1	9	9	4	46	<p>d M1</p> <p>$\sum d^2$ M1A1</p> <p>M1A1J</p> <p>0.452 A1</p> <p style="text-align: right;">(6)</p>
A	2	5	3	7	8	1	4	6																																		
B	3	2	6	5	7	4	1	8																																		
d	1	3	3	2	1	3	3	2																																		
d ²	1	9	9	4	1	9	9	4	46																																	
(b)	<p>$H_0 : \rho = 0, H_1 : \rho \neq 0 (\rho > 0)$</p> <p>critical values are ± 0.7381 (0.6429)</p> <p>$0.452 < 0.7381$ ($0.452 < 0.6429$) or not sig or Insufficient evidence to reject H_0</p> <p>No agreement between the two judges.</p>	<p>B1B1</p> <p>0.7381(0.6429) B1</p> <p>M1</p> <p>Context A1J</p> <p style="text-align: right;">(5)</p> <div style="border: 1px solid black; border-radius: 50%; width: 40px; height: 40px; display: flex; align-items: center; justify-content: center; margin: 0 auto;">11</div>																																								

Question number	Scheme	Marks														
7 (a)	$\mu = 0.3 \times 50 + 0.2 \times 10 + 0.5 \times 2 = 18$ $\sigma^2 = (0.3 \times 50^2 + 0.2 \times 10^2 + 0.5 \times 2^2) - 18^2 = 448$	M1A1 M1A1 (4)														
(b)	<p>(50,50) or (50,50) without ordered pairs</p> <p>(10,2) (10,2)</p> <p>(2,10) (10,10)</p> <p>(10,10) (50,10)</p> <p>(50,10) (2,2)</p> <p>(10,50) (50,2)</p> <p>(2,2)</p> <p>(50,2)</p> <p>(2,50)</p> <p style="text-align: right;">either, -1 each missing pair</p>	B2 (2)														
(c)	<table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td style="text-align: center;">\bar{x}</td> <td style="text-align: center;">2</td> <td style="text-align: center;">6</td> <td style="text-align: center;">10</td> <td style="text-align: center;">26</td> <td style="text-align: center;">30</td> <td style="text-align: center;">50</td> </tr> <tr> <td style="text-align: center;">$P(\bar{X} = \bar{x})$</td> <td style="text-align: center;">0.25</td> <td style="text-align: center;">0.2</td> <td style="text-align: center;">0.04</td> <td style="text-align: center;">0.3</td> <td style="text-align: center;">0.12</td> <td style="text-align: center;">0.09</td> </tr> </table> <p style="text-align: right;">All means, probabs multiplied, -1 each error</p>	\bar{x}	2	6	10	26	30	50	$P(\bar{X} = \bar{x})$	0.25	0.2	0.04	0.3	0.12	0.09	B1 M1 A2 (4)
\bar{x}	2	6	10	26	30	50										
$P(\bar{X} = \bar{x})$	0.25	0.2	0.04	0.3	0.12	0.09										
(d)	$P(2 \leq \bar{X} < 7) = 0.25 + 0.2 = 0.45$	Probabilities of 2 and 6 added, 0.45 M1 A1J (2)														
(e)	$E(\bar{X}) = 2 \times 0.25 + 6 \times 0.2 + \dots = 18$ $\text{Var}(\bar{X}) = 2^2 \times 0.25 + 6^2 \times 0.2 + \dots - 18^2 = 224$ $\sum x^2 P(X = x) - (\text{theirs})^2, 224$ M1A1 <p>So $E(\bar{X}) = 18 = \mu$ and $\text{Var}(\bar{X}) = 224 = \frac{1}{2} \sigma^2$ as required.</p>	$\sum xP(X = x)$ from table, 18 M1 A1 A1 (5)														