

## Statistics S3 Mark scheme

Question	Scheme	Marks
<b>1(a)</b>	$\{w\} = 018$ or 18	B1
		<b>(1)</b>
<b>(b)</b>	$\{x\} = 18$	B1
		<b>(1)</b>
<b>(c)</b>	$\{\text{prob}\} = 0$	B1
		<b>(1)</b>
<b>(d)</b>	<p><b>Advantage:</b> Any one of:</p> <ul style="list-style-type: none"> <li>• <u>Simple</u> or <u>easy</u> to use also allow “quick” or “efficient” (o.e.)</li> <li>• It is suitable for large samples ( or populations)</li> <li>• Gives a good spread of the data</li> </ul> <p><b>Disadvantage:</b> Any one of:</p> <ul style="list-style-type: none"> <li>• The alphabetical list is (probably) <u>not random</u></li> <li>• <u>Biased</u> since the list is not (truly) random</li> <li>• <u>Some combinations</u> of names are <u>not possible</u></li> </ul>	B1
		B1
		<b>(2)</b>
<b>(5 marks)</b>		
<b>Notes:</b>		
<b>(d)</b>	If no labels are given treat the 1 <sup>st</sup> reason as an advantage and the 2 <sup>nd</sup> as a disadvantage	
<b>B1:</b>	For advantage	
<b>B1:</b>	For disadvantage – “it requires a sampling frame” is 2 <sup>nd</sup> B0 since the alphabetical list is given.	
	Note: Do not score both B1 marks for opposing advantages and disadvantages.	

Question	Scheme										Marks	
<b>2(a)</b>	<i>A</i>	<i>B</i>	<i>C</i>	<i>L</i>	<i>N</i>	<i>R</i>	<i>S</i>	<i>T</i>	<i>Y</i>		M1	
	Judge 1	6	3	4	9	2	8	1	5	7		
	Judge 2	8	4	5	7	3	9	1	2	6		
	or											
	<i>S</i>	<i>N</i>	<i>B</i>	<i>C</i>	<i>T</i>	<i>A</i>	<i>Y</i>	<i>R</i>	<i>L</i>			
	Judge 1	1	2	3	4	5	6	7	8	9		
	Judge 2	1	3	4	5	2	8	6	9	7		
	$\sum d^2 = 4 + 1 + 1 + 4 + 1 + 1 + 0 + 9 + 1$ or $0 + 1 + 1 + 1 + 9 + 4 + 1 + 1 + 4 = 22$										M1	
											$\sum d^2 = 22$	A1
												M1
	$r_s = 1 - \frac{6(22)}{9(80)}; = 0.8166666\dots$										$\frac{49}{60}$ or awrt <b>0.817</b>	A1
												<b>(5)</b>
<b>(b)</b>	$H_0 : \rho = 0, H_1 : \rho > 0$										B1	
	Critical Value = 0.7833 <u>or</u> CR: $r_s \geq 0.7833$										<b>0.7833</b>	B1
	Since $r_s = 0.8166\dots$ it lies in the CR, <u>or</u> reject $H_0$ (o.e.)											M1
	The two <u>judges</u> (or “they”) are in <u>agreement</u> <u>or</u> there is a <u>positive correlation</u> between the ranks of the two <u>judges</u> .											A1ft
												<b>(4)</b>
<b>(9 marks)</b>												
<b>Notes:</b>												
<b>(a)</b>												
<b>M1:</b> For an attempt to rank at least one row (at least 4 correct)												
<b>M1:</b> For an attempt at $d^2$ row (may be implied by sight of $\sum d^2 = 22$ or 221 for reverse ranks)												
<b>A1:</b> For $\sum d^2 = 22$ (or 221 if reverse ranking is used) Can be implied by correct answer.												
<b>M1:</b> For use of the correct formula with their $\sum d^2$ (if it is clearly stated)												
If the answer is not correct then a correct expression is required												
<b>False Ranking</b> - e.g. Alphabetic ranking: Gives												
Judge 1: 7 5 2 3 8 1 9 6 4												
Judge 2: 7 8 5 2 3 9 4 1 6 $\sum d^2 = 162$ and $r_s = -0.35$												

**Question 2 notes continued**

Scores: M0(for ranking), M1(for attempt at  $d^2$  row), A0, M1 (for use of their  $\sum d^2$ ), A0 i.e. 2 out of 5. Can follow through their  $r_s$  in (b)

**(b)**

- B1:** For both hypotheses stated correctly in terms of  $\rho$  (allow  $\rho_s$ )  $H_1$  must be compatible with ranking.
- B1:** For  $cv = 0.7833$  (independent of their  $H_1$  (no 2-tail value in tables) but compatible sign with their  $r_s$ ).
- M1:** For a correct statement (in words) relating their  $r_s$  with their critical value. E.g. “reject  $H_0$ ”, “in critical region”, “significant”, “positive correlation”. May be implied by a correct contextual comment.
- |cv|>1** - If their  $cv$  is  $|cv| > 1$  (often from using normal tables) award M0A0
- If  $|their r_s| > |their cv|$  then “significant” (o.e.) for M1 and “judges are in agreement” (o.e.) for A1ft
  - If  $|their r_s| < |their cv|$  then “not significant” (o.e.) for M1 and “judges don’t agree” (o.e.) for A1ft
- A1ft:** For a correct follow through conclusion in context. “Positive correlation” alone scores M1 A0. For reverse ranking should still say “judges are in agreement”

Question	Scheme		Marks																																																				
<b>3(a)</b>	$\hat{\lambda} = \frac{0(47) + 1(57) + 2(46) + 3(35) + 4(9) + 5(6)}{200} = \frac{320}{200} = 1.6$	Full exp' or at least 2 products and 320/200 seen	B1 *																																																				
			<b>(1)</b>																																																				
<b>(b)</b>	$r = 200 \times \frac{e^{-1.6}(1.6)^2}{2!} \{= 51.68550861...\}$	Using $r = 200 \times \frac{e^{-1.6}(1.6)^2}{2!}$	M1																																																				
	$s = 200 - (40.38 + 64.61 + \text{their } r + 27.57 + 11.03) \{= 4.72449139...\}$ <u>or</u> their $r + s = 56.41$		M1																																																				
	$r = 51.68550861...$ and $s = 4.72449139...$	$r = \text{awrt } \mathbf{51.69}$ and $s = \text{awrt } \mathbf{4.72}$	A1																																																				
			<b>(3)</b>																																																				
<b>(c)</b>	$H_0$ : Poisson (distribution) is a suitable/ sensible (model) $H_1$ : Poisson (distribution) is not a suitable/ sensible (model).		B1																																																				
	<table border="1"> <thead> <tr> <th>Number of accidents</th> <th>Observed</th> <th>Expected</th> <th>Combined Observed</th> <th>Combined Expected</th> <th><math>\frac{(O - E)^2}{E}</math></th> <th><math>\frac{O^2}{E}</math></th> </tr> </thead> <tbody> <tr> <td>0</td> <td>47</td> <td>40.38</td> <td>47</td> <td>40.38</td> <td>1.0853</td> <td>54.7053</td> </tr> <tr> <td>1</td> <td>57</td> <td>64.61</td> <td>57</td> <td>64.61</td> <td>0.8963</td> <td>50.2863</td> </tr> <tr> <td>2</td> <td>46</td> <td>51.69</td> <td>46</td> <td>51.69</td> <td>0.6264</td> <td>40.9364</td> </tr> <tr> <td>3</td> <td>35</td> <td>27.57</td> <td>35</td> <td>27.57</td> <td>2.0024</td> <td>44.4324</td> </tr> <tr> <td>4</td> <td>9</td> <td>11.03</td> <td rowspan="2">15</td> <td rowspan="2">15.75</td> <td rowspan="2">0.0357</td> <td rowspan="2">14.2857</td> </tr> <tr> <td><math>\geq 5</math></td> <td>6</td> <td>4.72</td> </tr> <tr> <td colspan="4">Totals</td> <td>4.6461</td> <td>204.6461</td> <td></td> </tr> </tbody> </table>		Number of accidents	Observed	Expected	Combined Observed	Combined Expected	$\frac{(O - E)^2}{E}$	$\frac{O^2}{E}$	0	47	40.38	47	40.38	1.0853	54.7053	1	57	64.61	57	64.61	0.8963	50.2863	2	46	51.69	46	51.69	0.6264	40.9364	3	35	27.57	35	27.57	2.0024	44.4324	4	9	11.03	15	15.75	0.0357	14.2857	$\geq 5$	6	4.72	Totals				4.6461	204.6461		M1
	Number of accidents	Observed	Expected	Combined Observed	Combined Expected	$\frac{(O - E)^2}{E}$	$\frac{O^2}{E}$																																																
	0	47	40.38	47	40.38	1.0853	54.7053																																																
	1	57	64.61	57	64.61	0.8963	50.2863																																																
	2	46	51.69	46	51.69	0.6264	40.9364																																																
	3	35	27.57	35	27.57	2.0024	44.4324																																																
	4	9	11.03	15	15.75	0.0357	14.2857																																																
	$\geq 5$	6	4.72																																																				
	Totals				4.6461	204.6461																																																	
	$X^2 = \sum \frac{(O - E)^2}{E}$ or $\sum \frac{O^2}{E} - 200 ; = 4.6461$			M1																																																			
			awrt <b>4.65</b>	A1																																																			
$v = 5 - 1 - 1 = 3$		<b>3</b>	B1 ft																																																				
$\chi^2_3(0.10) = 6.251 \Rightarrow \text{CR: } X^2 \geq 6.251$		<b>6.251</b>	B1 ft																																																				
[Since $X^2 = 4.6461$ does not lie in the CR, then there is insufficient evidence to reject $H_0$ ]																																																							
The number of <i>accidents</i> per day can be modelled by a Poisson distribution <u>or</u> the <i>supervisor's</i> belief is correct.			A1 ft																																																				
			<b>(7)</b>																																																				
			<b>(11 marks)</b>																																																				
<b>Notes:</b>																																																							
<b>(b)</b> <b>Note:</b> Allow A1 for $s = \text{awrt } 4.74$ (fou as a result of using expected values to full accuracy.)																																																							

**Question 3 notes** *continued*

(c)

**B1:** For both hypotheses and mentioning Poisson at least once. Allow Poisson is a “good fit/model” but not “good method”. Inclusion of 1.6 for mean in hypotheses is B0 but condone in conclusion.

**M1:** For an attempt to pool 4 accidents and  $\geq 5$  accidents or pool when  $E_i < 5$  No pooling is M0

**M1:** For an attempt at the test statistic, at least 2 correct expressions/values (to awrt 2 d.p.)

**A1:** For awrt 4.65 (score M1M1A1 if awrt 4.65 seen).

**No pooling:** If no pooling can allow 2<sup>nd</sup> M1 if  $X^2 = 5.33$  is seen

**B1ft:** For  $n - 1 - 1$  i.e. subtracting 2 from their  $n$ .

**B1ft:** For a correct ft for their  $\chi_k^2(0.10)$ , where  $k = n - 1 - 1$  from their  $n$ .

(B1B1 may be implied by 6.251 (if pooling) or 7.779 for no pooling)

**A1ft:** (*Dep. on the 2<sup>nd</sup> M1*) For correct comment in context based on their test statistic and their critical value that mentions **accidents** or **supervisor**. Condone mention of Po(1.6) in conclusion. Score A0 for inconsistencies e.g. “significant” followed by “supervisor’s belief is justified”

Note: Full accuracy gives a combined expected frequency of 15.76,  $\frac{(O - E)^2}{E} = 0.0366$ ,

$\frac{O^2}{E} = 14.2766$ ,  $X^2 = 4.64855\dots$  and p-value 0.199.

Question	Scheme		Marks
<b>4(a)</b>	Let $X$ = weight of a sack of potatoes, $X \sim N(25.6, 0.24^2)$		
	So $D = X_1 - X_2 \sim N(0, 2(0.24)^2)$ or $D \sim N(0, 0.1152)$	Attempt at $D$ and $D \sim N(0, ..)$	M1
		$(0.24)^2 + (0.24)^2$ ; <b>0.1152</b>	A1 A1
	$\{P( D  > 0.5)\} = 2P(D > 0.5)$	$2 \times P(D > 0.5)$ can be implied	dM1
	$= 2 \times P\left(Z > \frac{0.5}{\sqrt{0.1152}}\right)$		dM1
	$= 2 \times P(Z > 1.4731...)$ <u>or</u> $= 2(1 - 0.9292)$		
	$= 0.1416$	<b>awrt 0.141 or awrt 0.142</b>	A1
		<b>(6)</b>	
<b>(b)</b>	Let $Y$ = weight of an empty pallet, $Y \sim N(20.0, 0.32^2)$		
	So $T = X_1 + X_2 + \dots + X_{30} + Y$		
	$T \sim N(30(25.6) + 20, 30(0.24)^2 + 0.32^2)$	$30(25.6) + 20$ <u>or</u> <b>788</b>	B1
		$30(0.24)^2 + 0.32^2$	M1
	$T \sim N(788, 1.8304)$	N and 1.8304 or awrt <b>1.83</b>	A1
	$\{P(T > 785)\} = P\left(Z > \frac{785 - 788}{\sqrt{1.8304}}\right)$		M1
	$= P(Z > -2.2174...)$		
	$= 0.9868$	<b>awrt 0.987</b>	A1
		<b>(5)</b>	
<b>(Total 11)</b>			
<b>Notes:</b>			
<b>(a)</b>			
<b>M1:</b> For clear definition of $D$ and normal distribution with mean of 0 (Can be implied by 3 <sup>rd</sup> M1).			
<b>A1:</b> For correct use of $\text{Var}(X_1 - X_2)$ formula.			
<b>A1:</b> For 0.1152			
<b>dM1:</b> For realising need $2 \times P(D > 0.5)$ (Dependent on 1 <sup>st</sup> M1 i.e. must be using suitable $D$ ).			
<b>dM1:</b> Dep on 1 <sup>st</sup> M1 for standardising with 0.5, 0 and their s.d.( $\neq 0.24$ ) Must lead to $P(Z > +ve)$ (o.e.). $P(Z > 1.47)$ implies 1 <sup>st</sup> M1 1 <sup>st</sup> A1 2 <sup>nd</sup> A1 and 3 <sup>rd</sup> M1. Correct answer only will score 6 out of 6.			

**Question 4 notes** *continued*

**(b)**

**B1:** For a mean of  $30(25.6) + 20$ . Can be implied by 788.

**M1:** For  $30(0.24)^2 + 0.32^2$ . Can be implied by 1.8304 or awrt 1.83

Allow M1 for swapping error i.e.  $30 \times 0.32^2 + 0.24^2$  if the expression is seen

**A1:** For normal and correct variance of 1.8304 or awrt 1.83. Normality may be implied by standardisation

**M1:** For standardising with 785 with their mean and st. dev..( $\neq 0.24$ ) Must lead to  $P(Z > -ve)$  o.e.

**A1:** Awrt 0.987. Correct answer only will score 5 out of 5

Note: Calculator answers are (a) 0.14071... , (b) 0.98670...

Question	Scheme			Marks																																
<b>5</b>	$H_0$ : Grades and gender are independent (or not associated) $H_1$ : Grades and gender are dependent (or associated)			“grades” and “gender” mentioned at least once.	B1																															
	<table border="1"> <thead> <tr> <th>Observed</th> <th>Male</th> <th>Female</th> </tr> </thead> <tbody> <tr> <td>Distinction</td> <td>37</td> <td>44</td> </tr> <tr> <td>Merit</td> <td>127</td> <td>96</td> </tr> <tr> <td>Unsatisfactory</td> <td>36</td> <td>20</td> </tr> </tbody> </table>			Observed	Male	Female	Distinction	37	44	Merit	127	96	Unsatisfactory	36	20	An attempt to convert percentages to observed frequencies.	M1																			
	Observed	Male	Female																																	
	Distinction	37	44																																	
	Merit	127	96																																	
	Unsatisfactory	36	20																																	
				All observed frequencies are correct.	A1																															
	<table border="1"> <thead> <tr> <th>Expected</th> <th>Male</th> <th>Female</th> <th>Totals</th> </tr> </thead> <tbody> <tr> <td>Distinction</td> <td>45</td> <td>36</td> <td>81</td> </tr> <tr> <td>Merit</td> <td>123.889</td> <td>99.111</td> <td>223</td> </tr> <tr> <td>Unsatisfactory</td> <td>31.111</td> <td>24.889</td> <td>56</td> </tr> <tr> <td>Totals</td> <td>200</td> <td>160</td> <td>360</td> </tr> </tbody> </table>			Expected	Male	Female	Totals	Distinction	45	36	81	Merit	123.889	99.111	223	Unsatisfactory	31.111	24.889	56	Totals	200	160	360	Some attempt at $\frac{(\text{Row Total})(\text{Column Total})}{(\text{Grand Total})}$ Can be implied by a correct $E_i$	M1											
	Expected	Male	Female	Totals																																
	Distinction	45	36	81																																
Merit	123.889	99.111	223																																	
Unsatisfactory	31.111	24.889	56																																	
Totals	200	160	360																																	
			All expected frequencies are correct to nearest integer.	A1																																
<table border="1"> <thead> <tr> <th>Observed</th> <th>Expected</th> <th><math>\frac{(O - E)^2}{E}</math></th> <th><math>\frac{O^2}{E}</math></th> </tr> </thead> <tbody> <tr> <td>37</td> <td>45</td> <td>1.422</td> <td>30.422</td> </tr> <tr> <td>44</td> <td>36</td> <td>1.778</td> <td>53.778</td> </tr> <tr> <td>127</td> <td>123.889</td> <td>0.078</td> <td>130.189</td> </tr> <tr> <td>96</td> <td>99.111</td> <td>0.098</td> <td>92.987</td> </tr> <tr> <td>36</td> <td>31.111</td> <td>0.768</td> <td>41.657</td> </tr> <tr> <td>20</td> <td>24.889</td> <td>0.960</td> <td>16.071</td> </tr> <tr> <td colspan="2">Totals</td> <td>5.104</td> <td>365.104</td> </tr> </tbody> </table>			Observed	Expected	$\frac{(O - E)^2}{E}$	$\frac{O^2}{E}$	37	45	1.422	30.422	44	36	1.778	53.778	127	123.889	0.078	130.189	96	99.111	0.098	92.987	36	31.111	0.768	41.657	20	24.889	0.960	16.071	Totals		5.104	365.104	At least 2 correct terms for $\frac{(O - E)^2}{E}$ or $\frac{O^2}{E}$ or correct expressions with their $E_i$ . Accept 2 sf accuracy for the M1 mark.	M1
Observed	Expected	$\frac{(O - E)^2}{E}$	$\frac{O^2}{E}$																																	
37	45	1.422	30.422																																	
44	36	1.778	53.778																																	
127	123.889	0.078	130.189																																	
96	99.111	0.098	92.987																																	
36	31.111	0.768	41.657																																	
20	24.889	0.960	16.071																																	
Totals		5.104	365.104																																	
			All correct $\frac{(O - E)^2}{E}$ or $\frac{O^2}{E}$ terms to either 2 dp or better. Allow truncation. ( $\Rightarrow$ by awrt 5.1 if 3 <sup>rd</sup> M1 seen)	A1																																
$X^2 = \sum \frac{(O - E)^2}{E} \text{ or } \sum \frac{O^2}{E} - 360 ; = \text{awrt } 5.1$			awrt <b>5.1</b>	A1																																
$\nu = (3 - 1)(2 - 1) = 2$			( $\nu =$ ) <b>2</b> (Can be implied by 5.991)	B1																																
$\chi_2^2(0.05) = 5.991 \Rightarrow \text{CR: } X^2 \geq 5.991$			For <b>5.991</b> only	B1																																
Since $X^2 = 5.1$ does not lie in the CR, then there is insufficient evidence to reject $H_0$				M1																																

Question	Scheme	Marks
<b>5</b> <i>continued</i>	Business Studies <u>grades</u> and <u>gender</u> are independent <u>or</u> There is no association between Business Studies <u>grades</u> and <u>gender</u> <u>or</u> <u>Head of department</u> 's (belief) is correct	A1ft
		<b>(4)</b>
<b>(12 marks)</b>		
<b>Notes:</b>		
Final M1:	For a correct statement linking their test statistic and their critical value ( $> 3.8$ ) Note: Contradictory statements score M0. E.g. “significant, do not reject $H_0$ ”.	
Final A1ft:	For a correct ft statement in context – must mention “grades” and “gender” or “sex” <u>or</u> “head of department” Condone “relationship” or “connection” here but <b>not</b> “correlation”. e.g. “There is no evidence of a relationship between grades and gender”	
<b>5.10 only</b>	Just seeing 5.10... only can imply 1 <sup>st</sup> 3 Ms but loses 1 <sup>st</sup> 3 As so can score 4 out of 7 (Qu says show..”)	
<b>Note: Full accuracy gives</b> $X^2 = 5.104356...$ and p-value 0.0779		

Question	Scheme			Marks																															
<b>5</b>	<b><u>Mark Scheme for candidates who use percentages instead of observed values.</u></b>																																		
	$H_0$ : Grades and gender are independent (or not associated) $H_1$ : Grades and gender are dependent (or associated)		“grades” and “gender” mentioned at least once.	B1																															
	<table border="1"> <thead> <tr> <th>Observed</th> <th>Male</th> <th>Female</th> </tr> </thead> <tbody> <tr> <td>Distinction</td> <td>18.5</td> <td>27.5</td> </tr> <tr> <td>Merit</td> <td>63.5</td> <td>60.0</td> </tr> <tr> <td>Unsatisfactory</td> <td>18.0</td> <td>12.5</td> </tr> </tbody> </table>		Observed	Male	Female	Distinction	18.5	27.5	Merit	63.5	60.0	Unsatisfactory	18.0	12.5	These marks cannot be obtained.	M0 A0																			
	Observed	Male	Female																																
	Distinction	18.5	27.5																																
	Merit	63.5	60.0																																
	Unsatisfactory	18.0	12.5																																
	<table border="1"> <thead> <tr> <th>Expected</th> <th>Male</th> <th>Female</th> <th>Totals</th> </tr> </thead> <tbody> <tr> <td>Distinction</td> <td>23</td> <td>23</td> <td>46</td> </tr> <tr> <td>Merit</td> <td>61.75</td> <td>61.75</td> <td>123.5</td> </tr> <tr> <td>Unsatisfactory</td> <td>15.25</td> <td>15.25</td> <td>30.5</td> </tr> <tr> <td>Totals</td> <td>100</td> <td>100</td> <td>200</td> </tr> </tbody> </table>		Expected	Male	Female	Totals	Distinction	23	23	46	Merit	61.75	61.75	123.5	Unsatisfactory	15.25	15.25	30.5	Totals	100	100	200	Some attempt at (Row Total)(Column Total) (Grand Total) Can be implied by one of these $E_i$ 's	M1											
	Expected	Male	Female	Totals																															
	Distinction	23	23	46																															
Merit	61.75	61.75	123.5																																
Unsatisfactory	15.25	15.25	30.5																																
Totals	100	100	200																																
		Expected frequencies are not correct.	A0																																
<table border="1"> <thead> <tr> <th>Observed</th> <th>Expected</th> <th><math>\frac{(O - E)^2}{E}</math></th> <th><math>\frac{O^2}{E}</math></th> </tr> </thead> <tbody> <tr> <td>18.5</td> <td>23</td> <td>0.8804</td> <td>14.8804</td> </tr> <tr> <td>27.5</td> <td>23</td> <td>0.8804</td> <td>32.8804</td> </tr> <tr> <td>63.5</td> <td>61.75</td> <td>0.0496</td> <td>65.2996</td> </tr> <tr> <td>60.0</td> <td>61.75</td> <td>0.0496</td> <td>58.2996</td> </tr> <tr> <td>18.0</td> <td>15.25</td> <td>0.4959</td> <td>21.2459</td> </tr> <tr> <td>12.5</td> <td>15.25</td> <td>0.4959</td> <td>10.2459</td> </tr> <tr> <td colspan="2" style="text-align: center;">Totals</td> <td>2.8518</td> <td>202.8518</td> </tr> </tbody> </table>		Observed	Expected	$\frac{(O - E)^2}{E}$	$\frac{O^2}{E}$	18.5	23	0.8804	14.8804	27.5	23	0.8804	32.8804	63.5	61.75	0.0496	65.2996	60.0	61.75	0.0496	58.2996	18.0	15.25	0.4959	21.2459	12.5	15.25	0.4959	10.2459	Totals		2.8518	202.8518	At least 2 “correct” terms for $\frac{(O - E)^2}{E}$ or $\frac{O^2}{E}$ or correct expressions with their $E_i$ . Accept 2 sf accuracy for the M1 mark.	M1
Observed	Expected	$\frac{(O - E)^2}{E}$	$\frac{O^2}{E}$																																
18.5	23	0.8804	14.8804																																
27.5	23	0.8804	32.8804																																
63.5	61.75	0.0496	65.2996																																
60.0	61.75	0.0496	58.2996																																
18.0	15.25	0.4959	21.2459																																
12.5	15.25	0.4959	10.2459																																
Totals		2.8518	202.8518																																
		This mark cannot be obtained.	A0																																
$X^2 = \sum \frac{(O - E)^2}{E} \text{ or } \sum \frac{O^2}{E} - 360 ; = 2.8518$		This mark cannot be obtained.	A0																																
$\nu = (3 - 1)(2 - 1) = 2$		( $\nu =$ ) <b>2</b> (Can be implied by 5.991)	B1																																
$\chi^2_2(0.05) = 5.991 \Rightarrow \text{CR: } X^2 \geq 5.991$		For <b>5.991</b> only	B1																																

Question	Scheme		Marks
<b>5</b> <i>continued</i>	Since $X^2 = 2.86$ does not lie in the CR, then there is insufficient evidence to reject $H_0$		M1
		Not available since comes from incorrect	A0
			<b>(12)</b>
<b>(12 marks)</b>			
<b>Notes:</b>			
If a candidate uses percentages rather than observed values then they can obtain a maximum of <b>6 marks</b> . They can get B1 M0A0 M1A0 M1A0A0 B1B1M1A0.			

Question	Scheme		Marks
<b>6(a)</b>	$\left\{ \hat{\mu} = \frac{\sum x}{n} = \frac{1570}{50} = \right\} \bar{x} = 31.4$	$\bar{x} = \mathbf{31.4}$	B1 cao
	$\left\{ \hat{\sigma}^2 = \frac{\sum x^2 - n\bar{x}^2}{n-1} = \right\} s_x^2 = \frac{49467.58 - 50(31.4)^2}{50 - 1}$		M1 A1ft
	$= 3.460816\dots$	awrt <b>3.46</b>	A1
			<b>(4)</b>
<b>(b)</b>	[Let $Y =$ time taken to complete obstacle course in the afternoon.]		
	$H_0: \mu_x = \mu_y, H_1: \mu_x > \mu_y$		B1
	$(z =) \frac{"31.4" - 30.9}{\sqrt{\frac{"3.46"}{50} + \frac{3.03}{50}}}$		M1 A1ft
	$= 1.38781\dots$	awrt <b>1.39</b>	A1
	CR: $Z \geq 1.6449$ or probability = awrt 0.082 or awrt 0.083	<b>1.6449</b> or better	B1
	Since $z = 1.38781\dots$ does not lie in the CR, then there is insufficient evidence to reject $H_0$		M1
	Conclude that the <u>mean time</u> to complete the obstacle course is the same for the early <u>morning</u> and late <u>afternoon</u> .		A1
		<b>(7)</b>	
<b>(c)</b>	$\bar{X}$ and $\bar{Y}$ are both approx. <u>normally distributed</u> or $\bar{X} - \bar{Y}$ normal (Condone $\bar{x}$ and $\bar{y}$ )		B1
			<b>(1)</b>
<b>(d)</b>	Have assumed $s^2 \approx \sigma^2$ or variance of sample = variance of population		B1
			<b>(1)</b>
<b>(13 marks)</b>			
<b>Notes:</b>			
<b>(a)</b>			
<b>B1:</b> 31.4 cao. Allow 31 minutes, 24 seconds.			
<b>M1:</b> A correct expression for either $s$ or $s^2$ (ignore label)			
<b>A1ft:</b> A correct expression for $s^2$ with their ft $\bar{x}$ .			
<b>A1:</b> Awrt 3.46 (Correct answer scores 3 out of 3)			
<b>(b)</b>			
<b>B1:</b> Both hypotheses stated correctly, with some indication of which $\mu$ is which. E.g: $\mu_M, \mu_A$			

**Question 6 notes continued**

**M1:** For an attempt at  $\frac{a-b}{\sqrt{\frac{c}{50} + \frac{d}{50}}}$  with at least 3 of  $a, b, c$  or  $d$  correct. Allow  $\pm$

**A1ft:** For  $\pm \frac{\text{their } 31.4 - 30.9}{\sqrt{\frac{\text{their } 3.46}{50} + \frac{3.03}{50}}}$

$$\text{Allow } D = \bar{x} - \bar{y} \quad 1.64 \sim 1.65 = \frac{D - 0}{\sqrt{\frac{3.46}{50} + \frac{3.03}{50}}} \text{ [SE = 0.360277..]}$$

**A1:** For awrt 1.39 (possibly  $\pm$ ) (Allow for CV  $D =$  awrt 0.593) (NB  $d = 0.5$ )

Correct answer scores M1A1ftA1 but  $0 - (31.4 - 30.9) \rightarrow -1.39$  loses this 2<sup>nd</sup> A mark

**B1:** Critical value of 1.6449 or better (seen). Allow for probability = awrt 0.082 or awrt 0.083.

Note: p-values are 0.0823 (tables) and 0.0826 (calculator).

**M1:** For a correct statement linking their test statistic and their critical value.

Note: Contradictory statements score M0. E.g. “significant, do not reject  $H_0$ ”.

**A1:** For a correct statement in context that accepts  $H_0$  (no ft) Condone “no difference in mean times”. Must mention “mean time”, “morning” and “afternoon” or “both times of day”

**(c)**

**B1:** E.g.  $\bar{X} \sim N(\dots)$  need both. Allow in words e.g “sample means are normally distributed”.

**(d)**

**B1:** Condone only mentioning “ $x$ ” or “ $y$ ” but watch out for  $s_x = s_y$  or  $\sigma_x = \sigma_y$  which scores B0.

Question	Scheme	Marks
<b>7(a)</b>	Let $X =$ score on a die	
	$E(S) = 3.5, \text{Var}(S) = \frac{35}{12}$	$E(S) = 3.5$ B1
		$\text{Var}(S) = \frac{35}{12}$ or awrt <b>2.92</b> B1
		<b>(2)</b>
<b>(b)</b>	$\text{So, } \bar{S} \sim N\left(3.5, \frac{\left(\frac{35}{12}\right)}{40}\right) \text{ or } \bar{S} \sim N\left(3.5, \frac{7}{96}\right)$	B1ft
	$P(\bar{S} < 3) = P\left(Z < \frac{3 - 3.5}{\sqrt{\frac{7}{96}}}\right) \{= P(Z < -1.85164\dots)\}$	M1
	$\{= 1 - 0.9678\} = 0.0322$	<b>0.032 to 0.0322</b> A1
		<b>(3)</b>
<b>(5 marks)</b>		
<b>Notes:</b>		
<b>(a)</b>		
<b>B1:</b> (2 <sup>nd</sup> ) allow awrt 2.92		
<b>(b)</b>		
<b>B1ft:</b> For $\bar{S} \sim N\left(3.5, \frac{\left(\frac{35}{12}\right)}{40}\right)$ seen or implied. Follow through their $E(S)$ and their $\text{Var}(S)$		
N.B $\frac{7}{96} = 0.07291\dot{6}$ accept awrt 0.0729		
<b>M1:</b> For an attempt to standardise with 3, their mean ( $>3$ ) and $\sqrt{\frac{\text{their Var}(S)}{40}}$ . Must lead to $P(Z < -ve)$		
<b>A1:</b> For 0.032 ~ 0.0322		
<b>Alternative <math>\Sigma S</math></b>		
<b>B1ft:</b> For $\sum S \sim N\left(140, \frac{350}{3}\right)$ where 140 is $40 \times$ their $E(S)$ and variance is $40 \times$ their $\text{Var}(S)$ .		

**Question 7 notes continued**

**M1:** For  $P\left(Z < \frac{120 - "140"}{\sqrt{\frac{350}{3}}}\right)$  or  $P\left(Z < \frac{119.5 - "140"}{\sqrt{\frac{350}{3}}}\right) \{= P(Z < -1.8979...)\}$

**A1:** for 0.032~0.0322 or (with continuity correction) 0.0287 (tables) or 0.0289 (calculator).

Question	Scheme		Marks
<b>8(a)</b>	$\left\{ \bar{x} = \frac{29.74 + 31.86}{2} \right\} \Rightarrow \bar{x} = 30.8$	$\bar{x} = 30.8$ This can be implied. See note.	B1
	$"1.96" \left( \frac{\sigma}{\sqrt{n}} \right) = 31.86 - 30.8$ or $2("1.96") \left( \frac{\sigma}{\sqrt{n}} \right) = 31.86 - 29.74$		M1
	$SE_{\bar{x}} = \frac{31.86 - 30.8}{1.96} = 0.540816... = 0.54$ (2dp)	awrt <b>0.54</b>	A1
			<b>(3)</b>
<b>(b)</b>	A 90% CI for $\mu$ is $\bar{x} \pm 1.6449 \left( \frac{\sigma}{\sqrt{n}} \right)$		B1
	$= 30.8 \pm 1.6449(0.54)$	(their $\bar{x}$ ) $\pm$ (their $z$ )(their $SE_{\bar{x}}$ from (a))	M1
	$= (29.91, 31.69)$	(awrt <b>29.9</b> , awrt <b>31.7</b> )	A1
			<b>(3)</b>
<b>(c)</b>	Let $X$ = number of confidence intervals containing $\mu$		
	or $Y$ = number of confidence intervals not containing $\mu$		
	So $X \sim \text{Bin}(4, 0.9)$ or $Y \sim \text{Bin}(4, 0.1)$		M1
	$P(X \geq 3)$ or $P(Y \leq 1) = {}^4C_3(0.9)^3(0.1) + (0.9)^4$	${}^4C_3(0.9)^3(0.1) + (0.9)^4$ oe	A1
	$= 0.2916 + 0.6561 = 0.9477$	<b>0.9477</b> or <b>0.948</b>	A1
			<b>(3)</b>
<b>(9 marks)</b>			
<b>Notes:</b>			
<b>(a)</b>			
<b>B1:</b> $\bar{x} = 30.8$ may be implied by $1.96 \left( \frac{\sigma}{\sqrt{n}} \right) = [31.86 - 30.8] = 1.06$ <u>or</u>			
$2(1.96) \left( \frac{\sigma}{\sqrt{n}} \right) = 31.86 - 29.74$			
<b>M1:</b> A correct equation for either a width or a half-width involving a $z$ -value $1.5 \leq z \leq 2$			
Eg: "their $z$ " $\left( \frac{\sigma}{\sqrt{n}} \right) = 31.86 - "30.8"$ ft their $\bar{x}$ <u>or</u> $2("their z") \left( \frac{\sigma}{\sqrt{n}} \right) = 31.86 - 29.74$			
or "their $z$ " $(SE_{\bar{x}}) = 31.86 - "30.8"$ <u>or</u> $2("their z")(SE_{\bar{x}}) = 31.86 - 29.74$ are fine for M1.			
<b>A1:</b> 0.54 or awrt 0.54 Must be seen as final answer to (a) NB $\frac{53}{98}$ as final answer is A0			
Condone $\bar{x} \pm 1.96\sigma = \dots$ for B1 and M1 but A0 even if they say " $\sigma$ = standard error = 0.54". Otherwise answer only of 0.54 scores 3 out of 3			