



# General Certificate of Education

## Statistics 6380

### *SS03 Statistics 3*

# Mark Scheme

## *2006 examination - January series*

Mark schemes are prepared by the Principal Examiner and considered, together with the relevant questions, by a panel of subject teachers. This mark scheme includes any amendments made at the standardisation meeting attended by all examiners and is the scheme which was used by them in this examination. The standardisation meeting ensures that the mark scheme covers the candidates' responses to questions and that every examiner understands and applies it in the same correct way. As preparation for the standardisation meeting each examiner analyses a number of candidates' scripts: alternative answers not already covered by the mark scheme are discussed at the meeting and legislated for. If, after this meeting, examiners encounter unusual answers which have not been discussed at the meeting they are required to refer these to the Principal Examiner.

It must be stressed that a mark scheme is a working document, in many cases further developed and expanded on the basis of candidates' reactions to a particular paper. Assumptions about future mark schemes on the basis of one year's document should be avoided; whilst the guiding principles of assessment remain constant, details will change, depending on the content of a particular examination paper.

## Key To Mark Scheme And Abbreviations Used In Marking

M	mark is for method		
m or dM	mark is dependent on one or more M marks and is for method		
A	mark is dependent on M or m marks and is for accuracy		
B	mark is independent of M or m marks and is for method and accuracy		
E	mark is for explanation		
✓ or ft or F	follow through from previous incorrect result	MC	mis-copy
CAO	correct answer only	MR	mis-read
CSO	correct solution only	RA	required accuracy
AWFW	anything which falls within	FW	further work
AWRT	anything which rounds to	ISW	ignore subsequent work
ACF	any correct form	FIW	from incorrect work
AG	answer given	BOD	given benefit of doubt
SC	special case	WR	work replaced by candidate
OE	or equivalent	FB	formulae book
A2,1	2 or 1 (or 0) accuracy marks	NOS	not on scheme
-x EE	deduct x marks for each error	G	graph
NMS	no method shown	c	candidate
PI	possibly implied	sf	significant figure(s)
SCA	substantially correct approach	dp	decimal place(s)

### No Method Shown

Where the question specifically requires a particular method to be used, we must usually see evidence of use of this method for any marks to be awarded. However, there are situations in some units where part marks would be appropriate, particularly when similar techniques are involved. Your Principal Examiner will alert you to these and details will be provided on the mark scheme.

Where the answer can be reasonably obtained without showing working and it is very unlikely that the correct answer can be obtained by using an incorrect method, we must award **full marks**. However, the obvious penalty to candidates showing no working is that incorrect answers, however close, earn **no marks**.

Where a question asks the candidate to state or write down a result, no method need be shown for full marks.

Where the permitted calculator has functions which reasonably allow the solution of the question directly, the correct answer without working earns **full marks**, unless it is given to less than the degree of accuracy accepted in the mark scheme, when it gains **no marks**.

**Otherwise we require evidence of a correct method for any marks to be awarded.**

## SS03

Q	Solution	Marks	Total	Comments
1(a)	$r = -0.708$ (3 sf from calc)	B3	3	Alternative $n = 11$ $\sum y = 348.1$ $\sum x = 115$ $\sum y^2 = 11046.75$ $\sum x^2 = 2007$ $\sum xy = 3527.4$ B1
	or $r = \frac{3527.4 - (\frac{115 \times 348.1}{11})}{\sqrt{804.727} \times \sqrt{30.96}}$	or B1 M1		
	$= \frac{-111.827}{\sqrt{804.727} \times \sqrt{30.96}}$ $= -0.708$	A1		
(b)	$H_0: \rho = 0$	B1	5	– 0.71 SC B1M1A0 0.708 SC B1M1A0  for cv  for comparison ts/cv not +cv / – pmcc
	$H_1: \rho < 0$ 1 tail 1 % sig level	B1		
	test stat $r = -0.708$ cv = – 0.6851 since $ts < -0.6851$	M1		
	Reject $H_0$ . Significant evidence at 1% level to suggest a negative linear association between the age at which a baby first learns to crawl and the average daily temperature during the sixth month of its life.	A1 E1		
(c)	A Type I error occurs when the Null Hypothesis is incorrectly rejected: in this case, when the conclusion made is that there is a negative association between temperature and age but, in fact, a negative association does not exist.	E1 E1	2	in context EO if $x/y$ used  in context Condone $x/y$  Allow 2 tail conclusion
	<b>Total</b>		<b>10</b>	

**SS03 (cont)**

Q	Solution	Marks	Total	Comments																					
<b>2(a)</b>	H <sub>0</sub> pop median difference, $\eta_d = 0$ H <sub>1</sub> pop median difference, $\eta_d \neq 0$ 2 tail 5%	B1																							
	<table border="1"> <thead> <tr> <th>Tyre</th> <th>1</th> <th>2</th> <th>3</th> <th>4</th> <th>5</th> <th>6</th> </tr> </thead> <tbody> <tr> <td>diff</td> <td>10.2</td> <td>2.3</td> <td>-0.6</td> <td>4.3</td> <td>-0.8</td> <td>2.2</td> </tr> <tr> <td>rank</td> <td>12</td> <td>7½</td> <td>-1</td> <td>9</td> <td>-2</td> <td>6</td> </tr> </tbody> </table>	Tyre	1	2	3	4	5	6	diff	10.2	2.3	-0.6	4.3	-0.8	2.2	rank	12	7½	-1	9	-2	6	M1		For differences
	Tyre	1	2	3	4	5	6																		
	diff	10.2	2.3	-0.6	4.3	-0.8	2.2																		
	rank	12	7½	-1	9	-2	6																		
	<table border="1"> <thead> <tr> <th>7</th> <th>8</th> <th>9</th> <th>10</th> <th>11</th> <th>12</th> </tr> </thead> <tbody> <tr> <td>5</td> <td>-2.3</td> <td>-1.2</td> <td>8.4</td> <td>2.1</td> <td>2.0</td> </tr> <tr> <td>10</td> <td>-7½</td> <td>-3</td> <td>11</td> <td>5</td> <td>4</td> </tr> </tbody> </table>	7	8	9	10	11	12	5	-2.3	-1.2	8.4	2.1	2.0	10	-7½	-3	11	5	4	m1 M1		For ranks of differences For ties			
	7	8	9	10	11	12																			
	5	-2.3	-1.2	8.4	2.1	2.0																			
	10	-7½	-3	11	5	4																			
	T <sub>+</sub> = 12 + 7½ + ..... + 5 = 64½ T <sub>-</sub> = 1 + 2 + 7½ + 3 = 13½ Test stat T = 13½ cv = 14 T < 14 Reject H <sub>0</sub> Significant evidence at 5% level to suggest that there is a difference in average treadwear measurement for the two methods.	m1 A1 B1 M1 A1 E1		For totals For one correct total For cv Comparison cv/ts <b>10</b>																					
<b>(b)</b> In the original design, the same tyre is used each time which eliminates any individual differences between tyres and means that any difference due to measurement method is more likely to be detected, if one exists. If two separate tyres were used, even from the same car and both from the front of the car, there may well be individual differences between them.	E1 E1		<b>2</b>																						
<b>(c)</b> Max T = $\sum_{r=1}^{12} r = 1 + 2 + \dots + 12 = 78$	M1 A1		<b>2</b>																						
<b>Total</b>			<b>14</b>																						

## SS03 (cont)

Q	Solution	Marks	Total	Comments	
3(a)	The frequencies are very low in several categories (insufficient data) and so a lot of pooling might be necessary that could reduce the contingency table below the $2 \times 2$ minimum required to sensibly carry out such an analysis.	B1	4	All $E_i$ are below 5 and pooling will not solve this problem	
	or	E1		any two valid reasons with explanation of reason in context	
	The level of poultry in the meat hot dogs is variable – could be 0% or up to 25% - so conclusion would not be relevant to investigating link to actual amount of poultry and sodium levels.	B1			
	or	E1			
	The sodium level categories are not discrete so some hot dogs could have been ‘double counted’.				
	(b)	$H_0$ Samples are taken from identical populations	B1	11	Hypotheses referring to population averages also acceptable
		$H_1$ Samples are not taken from identical populations – population average calorie content is lower for poultry hot dogs sausages. 1 tail 5%	B1		1 tail / ok generous
		Ranks Beef 15 13 12 8 14 16 4 7 Poultry 6 9 2 3 1 10 5 11	M1M1		For ranks as one group – at least 10 correct Other alternative methods acceptable
		$T_B = 15 + 13 + \dots + 7 = 89$ $T_P = 6 + 9 + \dots + 11 = 47$	m1		For totals of ranks in each group
		$U_B = 89 - \frac{8 \times 9}{2} = 53$	m1		For U attempted
		$U_P = 47 - \frac{8 \times 9}{2} = 11$			
Test stat $U = 11$		A1		For U correct – either	
$Cv = 16$		B1		For consistent cv with U	
$U < 16$		M1		For comparison U/cv	
Reject $H_0$		A1			
Significant evidence at the 5% level to suggest that the population average calorie content for poultry hot dogs is lower than that for beef hot dogs.	E1		In context		
			<b>15</b>		

SS03 (cont)

Q	Solution	Marks	Total	Comments																		
4(a)	H <sub>0</sub> Samples from identical populations	B1		or																		
	H <sub>1</sub> Samples not from identical populations 5% sig level	B1		H <sub>0</sub> $\eta_A = \eta_B = \eta_C$ H <sub>1</sub> at least two of $\eta_A, \eta_B, \eta_C$ do differ Allow $\eta_A \neq \eta_B \neq \eta_C$																		
	<b>Ranks</b>																					
	<table border="1"> <thead> <tr> <th>Fish Market A</th> <th>Fish Market B</th> <th>Fish Market C</th> </tr> </thead> <tbody> <tr> <td>3</td> <td>1</td> <td>10</td> </tr> <tr> <td>6</td> <td>2</td> <td>11</td> </tr> <tr> <td>8</td> <td>4</td> <td>13</td> </tr> <tr> <td>9</td> <td>5</td> <td>14</td> </tr> <tr> <td>12</td> <td>7</td> <td>15</td> </tr> </tbody> </table>	Fish Market A	Fish Market B	Fish Market C	3	1	10	6	2	11	8	4	13	9	5	14	12	7	15	M1		For ranks all as one group – can be reversed
	Fish Market A	Fish Market B	Fish Market C																			
	3	1	10																			
	6	2	11																			
	8	4	13																			
	9	5	14																			
	12	7	15																			
	M1		For at least 8 correct CAO																			
$T_A = 38$ $T_B = 19$ $T_C = 63$ $n_A = 5$ $n_B = 5$ $n_C = 5$	m1 A1		totals any one correct																			
$\sum_{i=1}^m \frac{T_i^2}{n_i} = \frac{38^2}{5} + \frac{19^2}{5} + \frac{63^2}{5} = 1154.8$	m1																					
$H = \frac{12}{15 \times 16} \times 1154.8 - (3 \times 16)$ $= 9.74$	m1 A1		9.60 – 9.80																			
Critical value from $\chi^2_2 = 5.99$ H > 5.99	B1 M1																					
Sig evidence to reject H <sub>0</sub> and conclude that samples are not from identical populations	A1																					
There is significant evidence that at least two of the average prices (from Fish Markets A, B or C) do differ.	E1		Difference in context Mention of ‘at least two’ ✓ E1, E0 if Accept H <sub>0</sub>																			
(b)		E1	<b>14</b>	Significant evidence to suggest that the mean price for C is certainly greater than the mean price for B																		
	Medians 227.3, 223.4, 249.6 It would appear that average prices at Fish Market C were significantly higher (as there is significant evidence of a difference detected in part (a)) and this would be the recommended Fish Market for Chinook salmon	B1																				
		E1	<b>2</b>	Identification of C with reason – generous																		
			<b>16</b>																			

SS03 (cont)

Q	Solution	Marks	Total	Comments																
5(a)(i)	<p><math>H_0</math> Violence level is independent of type of offence</p> <p><math>H_1</math> Violence level is not independent of type of offence      1 tail 5%</p>	B1																		
	<table border="1"> <thead> <tr> <th></th> <th>No violence</th> <th>Violence but no weapons</th> <th>Violence involving weapons</th> </tr> </thead> <tbody> <tr> <td>Non drug rel theft or damage</td> <td>55.47</td> <td>14.20</td> <td>5.33</td> </tr> <tr> <td>Drug rel theft or damage</td> <td>47.34</td> <td>12.12</td> <td>4.54</td> </tr> <tr> <td>Other</td> <td>22.19</td> <td>5.68</td> <td>2.13</td> </tr> </tbody> </table>		No violence	Violence but no weapons	Violence involving weapons	Non drug rel theft or damage	55.47	14.20	5.33	Drug rel theft or damage	47.34	12.12	4.54	Other	22.19	5.68	2.13	M1		M1 E method for 5 correct
		No violence	Violence but no weapons	Violence involving weapons																
	Non drug rel theft or damage	55.47	14.20	5.33																
	Drug rel theft or damage	47.34	12.12	4.54																
	Other	22.19	5.68	2.13																
	<p>Two <math>E_i</math> in the ‘violence involving weapons’ column are below 5 so pooling is required</p>	m1		For all E correct																
	<table border="1"> <thead> <tr> <th></th> <th>No Violence</th> <th>Violence</th> </tr> </thead> <tbody> <tr> <td>Non drug rel theft or damage</td> <td>55.47</td> <td>19.53</td> </tr> <tr> <td>Drug rel theft or damage</td> <td>47.34</td> <td>16.66</td> </tr> <tr> <td>Other</td> <td>22.19</td> <td>7.81</td> </tr> </tbody> </table>		No Violence	Violence	Non drug rel theft or damage	55.47	19.53	Drug rel theft or damage	47.34	16.66	Other	22.19	7.81	m1		For pooling				
		No Violence	Violence																	
	Non drug rel theft or damage	55.47	19.53																	
	Drug rel theft or damage	47.34	16.66																	
	Other	22.19	7.81																	
$ts = \sum \frac{(O - E)^2}{E}$ $= \frac{2.53^2}{55.47} + \frac{2.53^2}{19.53} + \frac{4.34^2}{47.34} + \frac{4.34^2}{16.66}$ $\frac{1.81^2}{22.19} + \frac{1.81^2}{7.81}$ $= 2.54$	A1		ts sum with correct denominators (condone no pooling)																	
<p>cv df = 2    5%    cv = 5.991</p>	B1		For ts in range 2.30 – 2.80 ( or 6.10 - 6.50 ft)																	
<p>ts &lt; 5.991 Accept <math>H_0</math>.</p>	m1		For cv (9.488 ft)																	
<p>No significant evidence to suggest use of violence is associated with type of offence</p>	A1		For comparison ts/cv ft																	
			sc If ts only sc (m1, 1, m1, A1) 3    6.1 – 6.5 sc (m1, 1, 1 A1, M1, A1) 4    2.3 – 2.8																	
			<b>10</b>																	

SS03 (cont)

Q	Solution	Marks	Total	Comments									
(b)(i)	H <sub>0</sub> Type of sentence is independent of whether firearms were used H <sub>1</sub> Type of sentence is not independent of whether firearms were used 1 tail 1%	B1											
	<table border="1"> <thead> <tr> <th></th> <th>Not used</th> <th>Used</th> </tr> </thead> <tbody> <tr> <th>Non custodial</th> <td>26</td> <td>8</td> </tr> <tr> <th>Custodial</th> <td>52</td> <td>16</td> </tr> </tbody> </table>		Not used	Used	Non custodial	26	8	Custodial	52	16	B1		For E values method
		Not used	Used										
	Non custodial	26	8										
	Custodial	52	16										
$ts = \sum \frac{( O - E  - 0.5)^2}{E} =$ $\frac{5.5^2}{26} + \frac{5.5^2}{8} + \frac{5.5^2}{52} + \frac{5.5^2}{16} = 7.42$	M1 M1		For ts for Yates' corr										
cv df= 1 1% cv = 6.635 ts > 6.635  Reject H <sub>0</sub> Significant evidence to suggest that type of sentence is not independent of whether firearms were used	A1	<b>8</b>	For ts For comparison ts/cv										
(b)(ii)	Offences where firearms are used are <b>much more likely</b> to result in a custodial sentence (and those where firearms are not used are <b>less likely</b> to result in a custodial sentence.)	B1		Correct association identified									
		E1	<b>2</b>	Explained in context									
			<b>20</b>										