



## **General Certificate of Education**

# **Statistics 6380**

**SS03          Statistics Unit 3**

# **Mark Scheme**

*2007 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.

Further copies of this Mark Scheme are available to download from the AQA Website: [www.aqa.org.uk](http://www.aqa.org.uk)

Copyright © 2007 AQA and its licensors. All rights reserved.

#### COPYRIGHT

AQA retains the copyright on all its publications. However, registered centres for AQA are permitted to copy material from this booklet for their own internal use, with the following important exception: AQA cannot give permission to centres to photocopy any material that is acknowledged to a third party even for internal use within the centre.

Set and published by the Assessment and Qualifications Alliance.

## 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	$H_0 \eta = 6$ $H_1 \eta > 6$ 1 tail    5%	B1		or $H_0$ population median = 6 $H_1$ population median > 6 or fully in words																																						
	<table border="1"> <thead> <tr> <th rowspan="2">Difference</th> <th colspan="2">Rank</th> </tr> <tr> <th>+</th> <th>-</th> </tr> </thead> <tbody> <tr> <td><math>X - 6</math></td> <td></td> <td></td> </tr> <tr> <td>4</td> <td>4½</td> <td></td> </tr> <tr> <td>2</td> <td>2</td> <td></td> </tr> <tr> <td>-1</td> <td></td> <td>1</td> </tr> <tr> <td>-4</td> <td></td> <td>4½</td> </tr> <tr> <td>5</td> <td>6</td> <td></td> </tr> <tr> <td>.</td> <td></td> <td></td> </tr> <tr> <td>7</td> <td>8</td> <td></td> </tr> <tr> <td>3</td> <td>3</td> <td></td> </tr> <tr> <td>8</td> <td>9</td> <td></td> </tr> <tr> <td>6</td> <td>7</td> <td></td> </tr> </tbody> </table>	Difference	Rank		+	-	$X - 6$			4	4½		2	2		-1		1	-4		4½	5	6		.			7	8		3	3		8	9		6	7		M1		for differences $(X - 6$ or $6 - X)$
	Difference		Rank																																							
		+	-																																							
	$X - 6$																																									
	4	4½																																								
	2	2																																								
	-1		1																																							
	-4		4½																																							
	5	6																																								
.																																										
7	8																																									
3	3																																									
8	9																																									
6	7																																									
		m1		for ranks (1 = smallest  diff  ) m0 if 0 = rank 1																																						
	Rank totals $T_- = 5\frac{1}{2}$ $T_+ = 39\frac{1}{2}$	m1F		for total of +/- ranks ft if any ranks																																						
	Test stat $T = 5\frac{1}{2}$	A1																																								
	critical value = 8 $n = 9$	B1		for cv																																						
	$T < cv$	M1		for comparison ts/cv																																						
	Reject $H_0$																																									
	There is significant evidence to suggest that the median for 18-year-old females is greater than 6.	A1	8																																							
	<b>Total</b>		<b>8</b>																																							

SS03 (cont)

Q	Solution	Marks	Total	Comments											
2(a)(i)	H <sub>0</sub> Response is independent of sex H <sub>1</sub> Response is not independent of sex 1 tail 1%	B1													
	<table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>Approve</th> <th>Dis-approve</th> <th>Don't care</th> </tr> </thead> <tbody> <tr> <th>Male</th> <td>83.3</td> <td>47.8</td> <td>221.9</td> </tr> <tr> <th>Female</th> <td>85.7</td> <td>49.2</td> <td>228.1</td> </tr> </tbody> </table>		Approve	Dis-approve	Don't care	Male	83.3	47.8	221.9	Female	85.7	49.2	228.1	M1 m1	
	Approve	Dis-approve	Don't care												
Male	83.3	47.8	221.9												
Female	85.7	49.2	228.1												
	$ts = \sum \frac{(O - E)^2}{E}$ $= \frac{12.3^2}{83.3} + \frac{12.8^2}{47.8} + \frac{25.1^2}{221.9} + \frac{12.3^2}{85.7} + \frac{12.8^2}{49.2} + \frac{25.1^2}{228.1}$ $= 16.0$	m1  A1 B1 m1		ts sum with correct denominators  for ts in range 15.7 - 16.2 for cv for comparison ts/cv											
	cv df = 2 1% cv = 9.210 ts > 9.210 Reject H <sub>0</sub> Sig evidence to suggest response is not independent of sex	A1F	8	ft if ts is very close and method is ok											
(a)(ii)	Males are much less likely than expected to disapprove of the royal wedding and females are much more likely than is expected to disapprove of the royal wedding Most noticeable differences in the sexes is that females were more likely to have an opinion of some sort about the wedding whereas males were more likely not to care	E1  E1	  2	explanation in context... ... with reference to expected/observed											
(a)(iii)	Most adults involved did not appear to care about the royal wedding	E1	1	not men/women											
(b)(i)	Data in Table 2 cannot be analysed to investigate whether an association exists because the raw frequencies are not supplied, only the percentages. This means that a $\chi^2$ test cannot be carried out.	E1	1	mention of % or not <b>actual</b> frequencies given											
(ii)	The total number of males and the total number of females in the sample is required.	E1		totals required for males and females											
	The raw frequencies in each category can then be found by evaluating the relevant percentage of the total eg 41% of the total number of males gives the raw frequency in the first cell (male / support monarchy)	E1	2	how total is used  SC 'convert into frequencies' B1											
	<b>Total</b>		<b>14</b>												

**SS03 (cont)**

Q	Solution	Marks	Total	Comments
<p><b>3(a)</b></p>	<p>(‘difference’:weekend – weekday)  <math>H_0 \eta_{\text{difference}} = 0</math>  <math>H_1 \eta_{\text{difference}} &gt; 0</math>      1 tail    10%</p> <p>Signs            + + + - + - + + + +  <math>8^+ / 2^-</math> signs - test values            Binomial ( 10, 0.5 ) model</p> <p><math>P(\geq 8^+) = P(\leq 2^-) = 0.0547 &lt; 0.10</math>            for one tail test            Reject <math>H_0</math>.            There is sufficient evidence, at the 10% level, to suggest that the median difference is greater than 0            Significant evidence that standardised mortality ratio is greater at the weekend</p>	<p>B1</p> <p>M1</p> <p>A1</p> <p>M1</p> <p>M1</p> <p>A1</p> <p>E1F</p>	<p>7</p> <p>2</p>	<p>direction            generous if fully worded and median / average</p> <p>signs            test stat correct and identified            binomial model used and probability attempted            comparison of Binomial probability with 0.10</p> <p>interpretation in context            ft conclusion</p> <p>concept of Type II correct</p> <p>in context</p>
	<b>Total</b>		<b>9</b>	

SS03 (cont)

Q	Solution	Marks	Total	Comments																					
4(a)	<p><math>H_0</math> Samples from identical populations  <math>H_1</math> Samples not from identical populations                      5% sig level</p> <p>Ranks</p> <table border="1"> <thead> <tr> <th>Department A</th> <th>Department B</th> <th>Department C</th> </tr> </thead> <tbody> <tr><td>3</td><td>1</td><td>7</td></tr> <tr><td>9</td><td>2</td><td>8</td></tr> <tr><td>11</td><td>4</td><td>10</td></tr> <tr><td>14</td><td>5</td><td>12</td></tr> <tr><td>15</td><td>6</td><td>13</td></tr> <tr><td>17</td><td></td><td>16</td></tr> </tbody> </table> <p> <math>T_A = 69</math>      <math>T_B = 18</math>      <math>T_C = 66</math>  <math>n_A = 6</math>      <math>n_B = 5</math>      <math>n_C = 6</math>  <math>\sum_{i=1}^m \frac{T_i^2}{n_i} = \frac{69^2}{6} + \frac{18^2}{5} + \frac{66^2}{5} = 1584.3</math>  <math>H = \frac{12}{17 \times 18} \times 1584.3 - (3 \times 18) = 8.13</math>                      Critical value from <math>\chi_2^2 = 5.99</math>  <math>H &gt; 5.99</math>                      Sig evidence to reject <math>H_0</math> and conclude that samples are not from identical populations. At least 2 differ.                 </p>	Department A	Department B	Department C	3	1	7	9	2	8	11	4	10	14	5	12	15	6	13	17		16	<p>B1</p> <p>M1</p> <p>A2,1</p> <p>M1</p> <p>A1</p> <p>M1</p> <p>m1</p> <p>A1</p> <p>B1</p> <p>M1</p> <p>A1</p>	<p>12</p> <p>2</p>	<p>or <math>H_0 \quad \eta_A = \eta_B = \eta_C</math>  <math>H_1</math> at least two of <math>\eta_A, \eta_B, \eta_C</math> do differ</p> <p>some ranks</p> <p>A1 for 5 A2 for all</p> <p>totals of some ranks any one correct</p> $\frac{12}{N(N+1)} \sum_{i=1}^m \frac{T_i^2}{n_i} - 3(N+1)$ <p>test stat <math>H = 7.80 - 8.40</math></p> <p>identification of B</p> <p>explanation in context of reason B selected                      'median' or 'average' required or explanation all B scores low</p>
	Department A	Department B	Department C																						
	3	1	7																						
	9	2	8																						
	11	4	10																						
	14	5	12																						
	15	6	13																						
	17		16																						
	<b>Total</b>			<b>14</b>																					

**SS03 (cont)**

Q	Solution	Marks	Total	Comments
5	<p><math>H_0</math> Samples are taken from identical populations  <math>H_1</math> Samples are not taken from identical populations – population average time to become over-ripe is lower for ‘chilled’ bananas                      1 tail 5%</p> <p>Sum of ranks                      ‘Chilled’ <math>2+6+5+8+1 = 22 = T_C</math>                      ‘Stored at 10°C’  <math>3+4+9+10+7+11 = 44 = T_S</math></p> <p><math>U_C = 22 - \frac{5 \times 6}{2} = 7</math>  <math>U_S = 44 - \frac{6 \times 7}{2} = 23</math></p> <p>Test stat <math>U = 7</math>  <math>cv = 5</math>  <math>U &gt; 5</math>                      Accept <math>H_0</math>                      No significant evidence at the 5% level to suggest that the population average time to become over-ripe is lower for ‘chilled’ bananas</p>	<p>B1                      B1                        m1                        m1                        A1                      B1                      M1                      A1                      E1</p>	<p>9</p>	<p>hypotheses referring to population, averages also acceptable for direction/explanation [other alternative methods acceptable]</p> <p>for totals of ranks in each group</p> <p>for <math>U</math> attempted</p> <p>for <math>U</math> correct, either for consistent cv with <math>U</math> for comparison <math>U/cv</math> for any valid <math>U/cv</math></p> <p>in context</p>
<b>Total</b>			<b>9</b>	



SS03 (cont)

Q	Solution	Marks	Total	Comments																		
6(a)(i)	<table border="1"> <tr> <td>student</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td>micro rank</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td>macro rank</td> <td>1</td> <td>2</td> <td>5</td> <td>7</td> <td>4</td> </tr> </table>	student	1	2	3	4	5	micro rank	1	2	3	4	5	macro rank	1	2	5	7	4	M1 A1		( $r = 0.927$ )
	student	1	2	3	4	5																
micro rank	1	2	3	4	5																	
macro rank	1	2	5	7	4																	
<table border="1"> <tr> <td>student</td> <td>6</td> <td>7</td> <td>8</td> <td>9</td> <td>10</td> </tr> <tr> <td>micro rank</td> <td>6</td> <td>7</td> <td>8</td> <td>9</td> <td>10</td> </tr> <tr> <td>macro rank</td> <td>3</td> <td>6</td> <td>8</td> <td>9</td> <td>10</td> </tr> </table>	student	6	7	8	9	10	micro rank	6	7	8	9	10	macro rank	3	6	8	9	10			attempt at ranks	
student	6	7	8	9	10																	
micro rank	6	7	8	9	10																	
macro rank	3	6	8	9	10																	
	<p><math>r_s = 0.854(5)</math> (3 sf from calc)</p>	B3	5	<p><b>Alternative</b>  <math>d = 0, 0, 2, 3, 1, 3, 1, 0, 0, 0</math>  <math>\sum d^2 = 24</math>      <math>B1 r_s = 1 - \frac{6 \times 24}{10 \times 99}</math>  <math>= 0.854(5)</math> M1, A1                      (2 sf and no working SC4, A0)</p>																		
(a)(ii)	<p><math>H_0</math> Rank orders of micro-economics marks and macro-economics marks are independent.  <math>H_1</math> Rank orders of micro-economics marks and macro-economics marks are not independent. 2 tail 2%</p>	B1		<p><math>H_0</math> no association  <math>H_1</math> association</p>																		
	<p><math>cv = \pm 0.7333</math></p>	B1		for cv																		
	<p>test stat <math>r_s = 0.854(5)</math>  <math>r_s &gt; 0.7333</math> or <math> r_s  &gt;  cv </math></p>	M1		for comparison ts/cv $r_s = 0.854(5)$																		
	<p>Reject <math>H_0</math>                      Significant evidence at 2% level to suggest an association between rank orders of micro-economics marks and macro-economics marks.                      [Student with higher rank mark in micro-economics also has higher rank mark in macro-economics]</p>	A1  E1	5	allow A1 if $r$ 'close' and marks lost in (a)(i)  in context																		

**SS03 (cont)**

Q	Solution	Marks	Total	Comments																																	
6(b)	$H_0 \mu_{\text{difference}} = 0$ $H_1 \mu_{\text{difference}} \neq 0$ 2 tail    5%	B1		or $\eta$ or population average or words																																	
	<table border="1"> <thead> <tr> <th>Student</th> <th>Difference mac-mic</th> <th>Rank -      +</th> </tr> </thead> <tbody> <tr><td>1</td><td>-10</td><td>10</td></tr> <tr><td>2</td><td>4</td><td>4</td></tr> <tr><td>3</td><td>6</td><td>6½</td></tr> <tr><td>4</td><td>6</td><td>6½</td></tr> <tr><td>5</td><td>- 2</td><td>2</td></tr> <tr><td>6</td><td>- 9</td><td>8</td></tr> <tr><td>7</td><td>- 5</td><td>5</td></tr> <tr><td>8</td><td>- 3</td><td>3</td></tr> <tr><td>9</td><td>.</td><td>Discard</td></tr> <tr><td>10</td><td>- 1</td><td>1</td></tr> </tbody> </table>	Student	Difference mac-mic	Rank -      +	1	-10	10	2	4	4	3	6	6½	4	6	6½	5	- 2	2	6	- 9	8	7	- 5	5	8	- 3	3	9	.	Discard	10	- 1	1	M1		for differences
	Student	Difference mac-mic	Rank -      +																																		
	1	-10	10																																		
	2	4	4																																		
	3	6	6½																																		
	4	6	6½																																		
	5	- 2	2																																		
	6	- 9	8																																		
	7	- 5	5																																		
	8	- 3	3																																		
	9	.	Discard																																		
10	- 1	1																																			
Rank totals $T_- = 28$ $T_+ = 17$	m1F			for ranks (1 = smallest  diff  )																																	
Test stat $T = 17$	A1			ties																																	
critical value = 6 $n = 9$	B1			for total of +/- ranks																																	
$T > cv$	M1			for cv																																	
Accept $H_0$				for comparison ts/cv																																	
There is no significant evidence to suggest that there is a difference between the average marks for macro-economics and micro-economics	A1	9																																			
(c)	It appears, from (a), that students who do well in micro-economics also do well in macro-economics but there is no significant evidence that students perform better in one of the exams than in the other. Good students achieve higher marks in both exams with no pattern as to which they perform better in	B1																																			
		E1	2																																		
	<b>Total</b>		<b>21</b>																																		
	<b>TOTAL</b>		<b>75</b>																																		