



**General Certificate of Education (A-level)  
June 2012**

**Statistics**

**SS06**

**(Specification 6380)**

**Statistics 6**

***Mark Scheme***

---

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 events which all examiners participate in and is the scheme which was used by them in this examination. The standardisation process ensures that the mark scheme covers the students' responses to questions and that every examiner understands and applies it in the same correct way. As preparation for standardisation each examiner analyses a number of students' scripts: alternative answers not already covered by the mark scheme are discussed and legislated for. If, after the standardisation process, examiners encounter unusual answers which have not been raised 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 students' 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 from: [aqa.org.uk](http://aqa.org.uk)

Copyright © 2012 AQA and its licensors. All rights reserved.

**Copyright**

AQA retains the copyright on all its publications. However, registered schools/colleges 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 schools/colleges 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 abbreviations

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
CAO	correct answer only
CSO	correct solution only
AWFW	anything which falls within
AWRT	anything which rounds to
ACF	any correct form
AG	answer given
SC	special case
OE	or equivalent
A2,1	2 or 1 (or 0) accuracy marks
-x EE	deduct x marks for each error
NMS	no method shown
PI	possibly implied
SCA	substantially correct approach
c	candidate
sf	significant figure(s)
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.

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.**

<b>Q</b>	<b>Solution</b>	<b>Marks</b>	<b>Total</b>	<b>Comments</b>
<b>1(a)</b>	Apples not treated with chemical	E1	1	Need not chemical
<b>(b)</b>	To eliminate bias_/ reduce expt error – avoid ‘best’ apples being selected for chemical treatment for example	E1	1	So they would not have started to ripen thus chemical only factor_involved
<b>(c)</b>	2 factor analysis of variance	E1,1	2	E1 2 factor, E1 anova SC Latin Square B1
<b>(d)</b>	The variety of apple	E1	1	
	<b>Total</b>		<b>5</b>	

Q	Solution	Marks	Total	Comments
2(a)	Plan A Binomial $n = 30$ $p = 0.15$ $P(X \leq 1) = 0.048 < 0.05$	B1		For 0.048
	Plan B Binomial $n = 50$ $p = 0.15$ $P(X \leq 3) = 0.046 < 0.05$	B1		For 0.046
	Each plan meets requirement	E1	3	For comparison 0.05 or explanation meets requirement
(b)(i)	1% $p = 0.01$ Plan A $P(X \leq 1) = 0.964$	M1		$P(X \leq 1$ and 3) or 1% and 2%
	Plan B $P(X \leq 3) = 0.998$	A1		Any two
(ii)	2% $p = 0.02$ Plan A $P(X \leq 1) = 0.880$			
	Plan B $P(X \leq 3) = 0.982$	A1	3	All four
(c)(i)	Plots	M1 A1	2	Correct attempt (at least 6) Accurate, points joined and through (0, 1)
	(ii) Plan B has a higher probability of acceptance for (good quality) batches with both 1% and 2% non-conforming  Plan A better at rejecting high level non-conforming at 10% or both same / Plan B better at rejecting at 15%	E1  E1	2	Any 2 points made OE
(d)(i)	Plan C $P(\text{Acc}) = P(X = 0) + P(X = 1) \times P(X \leq 1)$ $= 0.0388 + (0.1368 \times 0.1756)$ $= 0.0628$	M1  m1  A1	3	Use of B(20, 0.15) seen 0.038/9, 0.136/7, 0.175/6  Attempt at 2 sample probs mult 0.062–0.063
	(ii) Advantage – sample size might be smaller most of time ( $n = 20$ ) than that for Plan A and will always be smaller than for Plan B	E1		Advantage On average / usually smaller sample size
	Disadvantage – Plan C prob of acceptance for $p = 0.15$ is above requirement or Plan C more complicated / time consuming	E1	2	Disadvantage Need to be labelled as adv or disadv
	<b>Total</b>		<b>15</b>	

Q	Solution	Marks	Total	Comments																								
3(a)(i)	$H_0$ pop mean diff $\mu_d = 5$ $H_1$ pop mean diff $\mu_d > 5$ 1 tail 10%  $d = \text{Brand} - \text{Sup Own}$ <table border="1" style="margin-left: 20px;"> <tr> <td></td> <td><b>1</b></td> <td><b>2</b></td> <td><b>3</b></td> <td><b>4</b></td> <td><b>5</b></td> </tr> <tr> <td><math>d</math></td> <td>+7</td> <td>+13</td> <td>-1</td> <td>+6</td> <td>-4</td> </tr> <tr> <td></td> <td><b>6</b></td> <td><b>7</b></td> <td><b>8</b></td> <td></td> <td></td> </tr> <tr> <td><math>d</math></td> <td>+16</td> <td>+1</td> <td>+5</td> <td></td> <td></td> </tr> </table>		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	$d$	+7	+13	-1	+6	-4		<b>6</b>	<b>7</b>	<b>8</b>			$d$	+16	+1	+5			B1		$H_0$ pop mean diff $\mu_d = 5$ $H_1$ pop mean diff $\mu_d < -5$ $H_1$ must be consistent with $d$
		<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>																						
	$d$	+7	+13	-1	+6	-4																						
		<b>6</b>	<b>7</b>	<b>8</b>																								
	$d$	+16	+1	+5																								
			M1		Differences (can be reversed)																							
		$\bar{d} = 5.375 \quad s = 6.781 \quad n = 8$	m1		Attempt to find $\bar{d}, s$ can be implied by correct ts																							
		$t = \frac{5.375 - 5}{\frac{6.781}{\sqrt{8}}} = 0.156$	M1		Use of $\frac{s}{\sqrt{8}}$ ft																							
			M1 m1 A1		Use of $\bar{x} - 5$ or $\bar{x} + 5$ Whole method for $t$ ( $\pm$ )0.156 (0.15–0.16) SC $t = -4.3275$ B3																							
		$df = 7 \quad  cv  = 1.415$	B1		For correct consistent cv CAO																							
	$0.156 < 1.415$ Accept $H_0$ Significant evidence to suggest that claim is correct.	E1	9	Clear correct conclusion in context of claim or jam contents																								
(ii)	Assumed that differences in jar jam contents are normally distributed and samples of jars are obtained at random.	E1		Diffs of contents / jam normal																								
		E1	2	Jars obtained at random																								
(b)(i)	Cannot assume scores are normally distributed as tasters are untrained or no guidance given for sweetness scores.	E1	1																									
(ii)	$H_0 \eta_{\text{difference}} = 0$ difference (supermarket – brand) $H_1 \eta_{\text{difference}} > 0$ 1 tail 5%  Signs + + + - + + + +  $7^+ / 1^-$ signs – test values  Binomial (8, 0.5) model $P(\geq 7^+) = P(\leq 1^-) = 0.035 < 0.05$ for one tail test  Reject $H_0$ . There is sufficient evidence, at the 5% level, to suggest that the claim is correct or supermarket jam is sweeter.	B1		Must be consistent																								
		M1																										
		A1		ts correct and identified or seen used																								
		M1		Correct probability compared with 0.05 or cr with probs																								
		E1	5	Correct conclusion in context  Allow W S-R B1, M1, A1, M1, E1 as sign differences: 22, 21, 9, -2, 11, 13, 8, 12 ranks + 8, 7, 3, 4, 6, 2, 5 $T_+ = 35$ - 1 $T_- = 1$ cv = 6 $T_- < 6$ reject $H_0$																								
	<b>Total</b>		<b>17</b>																									

Q	Solution	Marks	Total	Comments																					
4(a)(i)	Warning $2.00 \pm 1.96 \times \frac{0.02}{\sqrt{4}}$ (1.98, 2.02)	B1 M1  A1		For 1.96, 3.09 For $\frac{0.02}{\sqrt{4}}$  Warning correct both																					
	Action $2.00 \pm 3.09 \times \frac{0.02}{\sqrt{4}}$ (1.97, 2.03)	A1		Action correct both																					
	(ii) Ranges $0.02 \times 0.199 = 0.0040$ 0.004 $0.02 \times 0.595 = 0.0119$ 0.012 $0.02 \times 3.984 = 0.0797$ 0.080 $0.02 \times 5.309 = 0.1062$ 0.106	M1 A1	6	D values correct $\times 0.02$ AWRT or truncated 3dp																					
	(b)(i)																								
	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;"></td> <td style="width: 20%;">Sample 6</td> <td style="width: 20%;">Sample 7</td> <td style="width: 20%;">Sample 8</td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> <td style="width: 10%;"></td> </tr> <tr> <td><math>\bar{x}</math></td> <td>1</td> <td>-1</td> <td>-0.75</td> <td>A1</td> <td></td> <td>One mean and range correct or all 3 means / ranges</td> </tr> <tr> <td>Range</td> <td>4</td> <td>3</td> <td>7</td> <td>A1</td> <td>2</td> <td>All means and ranges correct</td> </tr> </table>		Sample 6	Sample 7	Sample 8				$\bar{x}$	1	-1	-0.75	A1		One mean and range correct or all 3 means / ranges	Range	4	3	7	A1	2	All means and ranges correct			
	Sample 6	Sample 7	Sample 8																						
$\bar{x}$	1	-1	-0.75	A1		One mean and range correct or all 3 means / ranges																			
Range	4	3	7	A1	2	All means and ranges correct																			
(ii)	Ranges all OK (0.01–0.07)	E1		Ranges considered OK																					
	Mean sample 10 = 2.0225 (or 2.25)	E1	2	Mean values considered Sample 10 a problem or identified																					
(iii)	Mean of sample 10 lies between upper warning and action lines. Take another sample immediately	E1 B1	2	Reason Take sample (B1 if no correct reason given) SC Beyond warning, take another sample B1																					
(c)(i)	$P(1.975 < X < 2.025) = (-0.75 < Z < 1.75)$ $= 0.733$	M1		-0.75 and 1.75 or 0.733																					
	Outside tolerances 0.267	A1	2	0.26–0.27																					
(ii)	Range of tolerances = $2.025 - 1.975$ $= 0.05$ $6 \times 0.02 = 0.12$	M1		Consideration of $6\sigma$ and range																					
	so range $< 6\sigma$ or [less than 6] $\times$ sd Process will not meet tolerances	E1	2	Not meet with correct reason SC comment on 0.267 out of tol B1																					
<b>Total</b>			<b>16</b>																						

Q	Solution	Marks	Total	Comments																
<b>5(a)(i)</b>	$A(20/29) \quad B(30/49) \quad C(50+)$ $T_A = 10.67 \quad T_B = 16.03 \quad T_C = 16.39$ $n_A = 5 \quad n_B = 6 \quad n_C = 6$ $T = 43.09$ $\sum \sum x_{ij}^2 = 111.138 \quad N = 17$ $\sum \frac{T_i^2}{n_i} = \frac{10.67^2}{5} + \frac{16.03^2}{6} + \frac{16.39^2}{6}$ $= 110.37$ $SS_{Ages} = 110.37 - \frac{43.09^2}{17}$ $= 1.148$ $SS_{Total} = 111.138 - \frac{43.09^2}{17}$ $= 1.917(5)$																			
	<table border="1"> <thead> <tr> <th></th> <th>SS</th> <th>df</th> <th>MS</th> </tr> </thead> <tbody> <tr> <td>Ages</td> <td>1.148</td> <td>2</td> <td>0.574</td> </tr> <tr> <td>Error</td> <td>0.769(5)</td> <td>14</td> <td>0.055</td> </tr> <tr> <td>Total</td> <td>1.9175</td> <td>16</td> <td></td> </tr> </tbody> </table>		SS	df	MS	Ages	1.148	2	0.574	Error	0.769(5)	14	0.055	Total	1.9175	16		M1		SS for ages
		SS	df	MS																
	Ages	1.148	2	0.574																
	Error	0.769(5)	14	0.055																
	Total	1.9175	16																	
			M1		SS for total (can be implied in table)															
			M1		Error SS ft (not -ve)															
			m1		Method for MS (dep error ss/df)															
		$F = \frac{0.574}{0.055} = 10.44$	m1 A1		Method for F (ft) 10.2–10.6 (or $p =$ )															
	$F_{14}^2 = 6.515 < 10.44$	B1 B1		df correct 2, 14 cv correct CAO allow $p = 0.00167$																
	$H_0 \mu_A = \mu_B = \mu_C$ $H_1$ at least 2 of the pop means differ	B1		hypotheses – subscripts identified OE																
	Reject $H_0$	A1	10																	
<b>(ii)</b>	<p>There is significant evidence of a difference in mean satisfaction scores for the 3 age groups so at least 2 groups differ.</p> <p>Ages 20/29 sig less satisfied than those aged 50+</p>	E1	1	In context																
<b>(iii)</b>	The normal populations of satisfaction scores have a common variance	E1	1	For either normally distributed satisfaction scores or populations of satisfaction scores have a common variance																



Q	Solution	Marks	Total	Comments																				
5(b)	<p><math>H_0 \mu_{Run} = \mu_{Cycle} = \mu_{Swim}</math>  <math>H_1</math> at least 2 of the means differ</p> <p><math>T_{Male} = 26.0</math>      <math>T_{Female} = 23.6</math>  <math>n_{Male} = 3</math>          <math>n_{Female} = 3</math></p> <p><math>T_{Run} = 14.7</math>    <math>T_{Cycle} = 17.5</math>    <math>T_{Swim} = 17.4</math>  <math>n_{Run} = 2</math>        <math>n_{Cycle} = 2</math>        <math>n_{Swim} = 2</math></p> <p><math>T = 49.6</math>    <math>\sum \sum x_{ij}^2 = 413.78</math>    <math>N = 6</math></p> <p>Total SS  <math>413.78 - \frac{49.6^2}{6} = 3.753</math></p> <p>Sex SS  <math>\frac{26.0^2}{3} + \frac{23.6^2}{3} - \frac{49.6^2}{6} = 0.96</math></p> <p>Sport SS  <math>\frac{14.7^2}{2} + \frac{17.5^2}{2} + \frac{17.4^2}{2} - \frac{49.6^2}{6} = 2.523</math></p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>SS</th> <th>df</th> <th>MS</th> </tr> </thead> <tbody> <tr> <td>Sex</td> <td>0.96</td> <td>1</td> <td>0.96</td> </tr> <tr> <td>Sport</td> <td>2.523</td> <td>2</td> <td>1.26</td> </tr> <tr> <td>Error</td> <td>0.27</td> <td>2</td> <td>0.135</td> </tr> <tr> <td>Total</td> <td>3.753</td> <td>5</td> <td></td> </tr> </tbody> </table> <p><math>F = \frac{1.26}{0.135} = 9.35</math>    <math>F_2^2 = 19.0</math></p> <p><math>9.35 &lt; 19</math>    Accept <math>H_0</math></p> <p>There is no significant evidence of a difference in mean training times for the 3 sports.</p>		SS	df	MS	Sex	0.96	1	0.96	Sport	2.523	2	1.26	Error	0.27	2	0.135	Total	3.753	5		<p>M1</p> <p>M1</p> <p>M1</p> <p>M1</p> <p>B1</p> <p>m1</p> <p>m1</p> <p>A1</p> <p>B1</p> <p>E1</p>	<p>22</p> <p>75</p> <p>10</p>	<p>Total SS method (can be implied in table)</p> <p>Sex SS method</p> <p>Sport SS method</p> <p>Error SS ft (not -ve)</p> <p>Error df correct <math>v = 2</math></p> <p>Method for MS ft (dep SSe)</p> <p>Method (dep prev Ms) for F (sex/error or sport/error) Not -ve</p> <p>Sports F correct 9.2–9.5</p> <p>cv correct CAO or <math>p = 0.117</math></p> <p>Correct conclusion in context, ft small arithmetic error in F</p>
	SS	df	MS																					
Sex	0.96	1	0.96																					
Sport	2.523	2	1.26																					
Error	0.27	2	0.135																					
Total	3.753	5																						
	<b>Total</b>		<b>22</b>																					
	<b>TOTAL</b>		<b>75</b>																					