



**General Certificate of Education
June 2010**

Statistics

SS06

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 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

Copyright © 2010 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.

SS06

Q	Solution	Marks	Total	Comments
1(a)	Randomised block.	B1	1	
(b)	Each document type is scanned by each scanner model in Design 3. In Design 2, document types are unevenly spread amongst scanner models eg Scanner X does not scan a Document 3.	E1	1	
(c)	Each scanner model only scans one type of document so the overall performance of the scanner cannot be fairly judged.	E1	1	or Document variation might be confused with scanner variation
(d)	2 factor ANOVA	B1	1	Or equivalent
	Total		4	

SS06 (cont)

Q	Solution	Marks	Total	Comments
2(a)	$300 \pm 1.96 \times \frac{1.6}{\sqrt{5}}$ (298.6 , 301.4) W	B1 M1	4	For 1.96 and 3.09 For $300 \pm z \times \text{sd}$
	$300 \pm 3.09 \times \frac{1.6}{\sqrt{5}}$ (297.8 , 302.2) A	M1 A1		For sd correct $\frac{1.6}{\sqrt{5}}$ Correct to 1dp
(b)	$5.484 \times 1.6 = 8.8$ (8.77)	M1 A1	2	D $\times 1.6$ ($n = 5$) All correct to 1 dp
	$4.197 \times 1.6 = 6.7$ (6.72)			
	$0.850 \times 1.6 = 1.4$ (1.36)			
	$0.367 \times 1.6 = 0.6$ (0.59)			
(c)(i)	$\bar{X} = 297.5$ range = 2.9 Range OK but mean below lower action limit. Stop production	B1 E1	4	For mean and range attempted ft if mean and range found
	(ii) $\bar{X} = 300.8$ range = 6.8 Mean OK but range between upper warning and action limits. Take another sample immediately.	B1 E1		For mean and range attempted ft if mean and range found SC in (c) B1 if only mean or only range found)
Total			10	

SS06 (cont)

Q	Solution	Marks	Total	Comments																								
3(a)	H ₀ pop mean diff $\mu_d = 0$ H ₁ pop mean diff $\mu_d \neq 0$ 2 tail 5%	B1		Pop mean used																								
	$d = A - B$	M1		differences																								
	<table border="1"> <tr> <td>pers</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> </tr> <tr> <td>d</td> <td>88</td> <td>51</td> <td>0</td> <td>40</td> <td>303</td> </tr> <tr> <td>pers</td> <td>6</td> <td>7</td> <td>8</td> <td>9</td> <td>10</td> </tr> <tr> <td>d</td> <td>-107</td> <td>24</td> <td>31</td> <td>315</td> <td>112</td> </tr> </table>	pers	1	2	3	4	5	d	88	51	0	40	303	pers	6	7	8	9	10	d	-107	24	31	315	112			
	pers	1	2	3	4	5																						
	d	88	51	0	40	303																						
	pers	6	7	8	9	10																						
	d	-107	24	31	315	112																						
	$\bar{d} = 85.7$ $s = 131.3$ $n = 10$	m1			Use of $\frac{s}{\sqrt{n}}$																							
	$t = \frac{85.7 - 0}{131.3 / \sqrt{10}} = 2.06$ (2.0646)	m1			Method for t																							
	df = 9 cv = ± 2.262	A1			± 2.065 (2.060 – 2.067)																							
2.06 < 2.262 OE	B1			or p = 0.0690																								
Accept H ₀	B1			9 df																								
No significant evidence to suggest a difference between the mean amounts of pizza eaten for types A and B.	B1			for correct cv																								
	A1			correct conclusion																								
	E1		9	context ft																								
(b)	Difference between amounts of pizza eaten are normally distributed.	E1	1																									
(c)	Sensible idea as amounts of pizza eaten seem erratic (person 3 ate none). Range of amounts 422 but sd 131.6. Quite a large sd for that range if normally dist.	E1	1	Sensible attempt at justification.																								
(d)	H ₀ pop median diff $\eta_d = 0$ H ₁ pop median diff $\eta_d \neq 0$ 2 tail 5%	B1		Pop median used																								
	$8^+ / 1^-$	M1		for signs or signed differences																								
	$P(\leq 1^-) = 0.0195 < 0.025$	M1		comparison with correct bin prob																								
	Significant evidence to reject H ₀ . Conclude that there is a difference in the average amounts of pizza eaten. (Pizza type A preferred)	A1		correct conclusion or critical region found																								
		E1		context																								
			5																									
(e)	Normal assumption may not be valid so sign test more appropriate and conclude that there is a difference in the popularity of the pizza types (Type A preferred).	E1		or refer to more powerful t test using all numerical data so conclusion in part (a) more appropriate. No difference in popularity detected.																								
		E1	2	[Not conclusions repeated]																								
Total			18																									

SS06 (cont)

Q	Solution	Marks	Total	Comments
4(a)(i)	$z = \frac{93-91}{\frac{3.1}{\sqrt{10}}} = 2.04$	M1		$\frac{3.1}{\sqrt{10}} = 0.980$
	$P(z > 2.04) = 0.0207$	m1		$z = \frac{93-\mu}{0.980}$ once
(ii)	$z = \frac{93-94}{\frac{3.1}{\sqrt{10}}} = -1.02$	m1	4	
	$P(z > -1.02) = 0.846$	A1		
(b)	$P(\bar{X} < x) < 0.10$		4	
	$\frac{x-94}{\frac{3.1}{\sqrt{8}}} < -1.2816$	B1		-1.2816
	$x < 94 - 1.2816 \times \frac{3.1}{\sqrt{8}}$	M1 M1		$\frac{3.1}{\sqrt{8}}$ correct method
	$x < 92.595$	A1		$x = (92.5 - 92.6)$
(c)	Type I error is to reject a batch as being below the quality required when the batch is of acceptable quality. (good batch rejected)	B1	2	Type I error correct.
		E1		'Producer's Risk' in context.
Total			10	

SS06 (cont)

Q	Solution	Marks	Total	Comments								
5.(a)	$P(\leq 2) \quad n = 50$	M1		Allow 2dp								
	<table border="1"> <tr> <td>p</td> <td>0.01</td> <td>0.05</td> <td>0.10</td> <td>0.15</td> </tr> <tr> <td>P(acc)</td> <td>0.986</td> <td>0.541</td> <td>0.112</td> <td>0.014</td> </tr> </table>	p	0.01		0.05	0.10	0.15	P(acc)	0.986	0.541	0.112	0.014
p	0.01	0.05	0.10	0.15								
P(acc)	0.986	0.541	0.112	0.014								
(b)	Smooth curve through the points (0, 1), (0.01, 0.986), (0.05, 0.541), (0.10, 0.112), (0.15, 0.014)	M1A1	2	Plot (must go through (0,1))								
(c)	$P(\text{acc}) = P(0) + P(1) \times P(0,1)$ $= 0.0424 + 0.1413 \times 0.1837$ $= 0.068$	M1	3									
		M1										
		A1										
(d)	Smooth curve through the points (0, 1), (0.01, 0.956), (0.05, 0.402), (0.1, 0.068), (0.15, 0.010)	B1	1	plot								
(e)(i)	A – simpler or -higher chance of accepting good quality batches	E1										
	B – better at rejecting batches with higher % non-conforming but still has 96% probability of accepting good quality batches. - smaller average sample size likely.	E1										
(ii)	Cost	B1	4	Any two								
	Ease of use Ability to reject poor quality batches/ accept good quality batches	B1										
Total			12									

SS06 (cont)

Q	Solution	Marks	Total	Comments																														
6(a)	$T_{Calm} = 6.97$ $T_{Windy} = 5.95$ $T_{VWindy} = 4.01$ $n_{Calm} = 7$ $n_{Windy} = 9$ $n_{VWindy} = 7$																																	
	$\sum \sum x_{ij}^2 = 14.6067$ $N = 23$																																	
	$\sum \frac{T_i^2}{n_i} = \frac{6.97^2}{7} + \frac{5.95^2}{9} + \frac{4.01^2}{7} = 13.171$																																	
	$SS_{Winds} = 13.171 - \frac{16.93^2}{23} = 0.70894$	M1		Winds SS																														
	$SS_{Total} = 14.6067 - \frac{16.93^2}{23} = 2.1447(5)$	M1		Total SS																														
	<table border="1"> <thead> <tr> <th></th> <th>SS</th> <th>df</th> <th>ms</th> </tr> </thead> <tbody> <tr> <td>Winds</td> <td>0.70894</td> <td>2</td> <td>0.35447</td> </tr> <tr> <td>Error</td> <td>1.43581</td> <td>20</td> <td>0.07179</td> </tr> <tr> <td>Total</td> <td>2.1447(5)</td> <td>22</td> <td></td> </tr> </tbody> </table>		SS	df	ms	Winds	0.70894	2	0.35447	Error	1.43581	20	0.07179	Total	2.1447(5)	22																		
		SS	df	ms																														
	Winds	0.70894	2	0.35447																														
	Error	1.43581	20	0.07179																														
	Total	2.1447(5)	22																															
$F = \frac{0.35447}{0.07179} = 4.94$	M1		Error SS ft																															
$H_0 \mu_c = \mu_w = \mu_{vw}$	m1		Method for MS → dep on df correct																															
H_1 at least 2 of the means differ	A1		Method for F 4.50 – 5.00 (or p = 0.081)																															
$F \frac{2}{20} = 3.493$ $4.94 > 3.493$	B1		For hypotheses																															
Reject H_0 . There is significant evidence of a difference in mean iron pollution for the 3 average wind speeds. (greater pollution when calm than when very windy)	B1 B1 A1		df correct for cv correct																															
(b)(i)	total $T_1 = 2.32$ $T = 7.16$ total $T_2 = 2.31$ total $T_3 = 2.53$																																	
	$\sum \frac{T_i^2}{n_i} = \frac{2.32^2}{3} + \frac{2.31^2}{3} + \frac{2.53^2}{3} = 5.7065$	M1		Totals attempted for temps																														
	$SS_{temps} = 5.7065 - \frac{7.16^2}{9} = 0.01029$	A1		SS attempted 0.006 – 0.014																														
	<table border="1"> <thead> <tr> <th></th> <th>SS</th> <th>df</th> <th>ms</th> <th>F</th> </tr> </thead> <tbody> <tr> <td>Winds</td> <td>0.34682</td> <td>2</td> <td>0.17341</td> <td>22.59</td> </tr> <tr> <td>Humidity</td> <td>0.01056</td> <td>2</td> <td>0.00528</td> <td>0.69</td> </tr> <tr> <td>Temps</td> <td>0.01029</td> <td>2</td> <td>0.00514</td> <td>0.67</td> </tr> <tr> <td>Error</td> <td>0.01535</td> <td>2</td> <td>0.00765</td> <td></td> </tr> <tr> <td>Total</td> <td>0.38302</td> <td>8</td> <td></td> <td></td> </tr> </tbody> </table>		SS	df	ms	F	Winds	0.34682	2	0.17341	22.59	Humidity	0.01056	2	0.00528	0.69	Temps	0.01029	2	0.00514	0.67	Error	0.01535	2	0.00765		Total	0.38302	8			m1		Error SS ft
		SS	df	ms	F																													
	Winds	0.34682	2	0.17341	22.59																													
Humidity	0.01056	2	0.00528	0.69																														
Temps	0.01029	2	0.00514	0.67																														
Error	0.01535	2	0.00765																															
Total	0.38302	8																																
	B1		df correct for Humidity, Temps & Error																															
	A1		any F correct (22.0 – 23.0) (0.62 – 0.72)																															

SS06 (cont)

Q	Solution	Marks	Total	Comments
6(b)(ii)	$H_0 \mu_c = \mu_w = \mu_{vw}$ H_1 at least 2 of the means differ $F \frac{2}{2} = 19 \quad 22.59 > 19$ Reject H_0 . There is significant evidence of a difference in mean iron pollution for the 3 average wind speeds.	B1 M1 A1	3	for cv correct F ts comparison in context
6(b)(iii)	Test stats $F_{\text{Humidity}} = 0.69$ $F_{\text{Temps}} = 0.67$ $F \frac{2}{2} = 19 \quad cv$ Both ts < 19 so not effective blocking factors	B1 E1	2	Identification of relevant ts and cv – can be implied Not effective as not significant
	Total		21	
	TOTAL		75	