

GCE

Edexcel GCE

Mathematics

Statistics 3 S3 (6691)

June 2008

advancing learning, changing lives

Mark Scheme

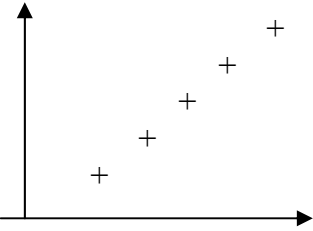
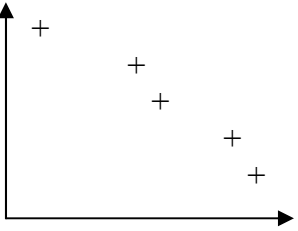
General Marking Guidance

- All candidates must receive the same treatment. Examiners must mark the first candidate in exactly the same way as they mark the last.
- Mark schemes should be applied positively. Candidates must be rewarded for what they have shown they can do rather than penalised for omissions.
- Examiners should mark according to the mark scheme not according to their perception of where the grade boundaries may lie.
- There is no ceiling on achievement. All marks on the mark scheme should be used appropriately.
- All the marks on the mark scheme are designed to be awarded. Examiners should always award full marks if deserved, i.e. if the answer matches the mark scheme. Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme.
- Where some judgement is required, mark schemes will provide the principles by which marks will be awarded and exemplification may be limited.
- When examiners are in doubt regarding the application of the mark scheme to a candidate's response, the team leader must be consulted.
- Crossed out work should be marked UNLESS the candidate has replaced it with an alternative response.

**June 2008
6691 Statistics S3
Mark Scheme**

Question number	Scheme	Marks	
1. (a)	$\bar{x} = \left(\frac{6046}{36} \right) = 167.94\dots$ $s^2 = \frac{1016338 - 36 \times \bar{x}^2}{35}$ $= 27.0253\dots$	<p style="text-align: right;">awrt 168</p> <p style="text-align: right;">awrt 27.0 (Accept 27)</p>	<p>B1</p> <p>M1</p> <p>A1 (3)</p>
(b)	<p>99% Confidence Interval is: $\bar{x} \pm 2.5758 \times \frac{5.1}{\sqrt{36}}$</p> $= (165.755\dots, 170.133\dots)$	<p style="text-align: right;">2.5758</p> <p style="text-align: right;">awrt (166,170)</p>	<p>M1A1ft</p> <p>B1</p> <p>A1 A1 (5)</p> <p style="text-align: right;">8 marks</p>
(a)	<p>M1 for a correct expression for s^2, follow through their mean, beware it is very “sensitive”</p> $167.94 \rightarrow \frac{999.63\dots}{35} \rightarrow 28.56\dots$ $167.9 \rightarrow \frac{1483.24\dots}{35} \rightarrow 42.37\dots$ $168 \rightarrow \frac{274}{35} \rightarrow 7.82$ <p>Use of 36 as the divisor (= 26.3...) is M0A0</p>	<div style="border: 1px solid black; padding: 5px; width: fit-content; margin: auto;"> <p>These would all score M1A0</p> </div>	
(b)	<p>M1 for substituting their values in $\bar{x} \pm z \times \frac{5.1 \text{ or } s}{\sqrt{36}}$ where z is a recognizable value from tables</p> <p>1st A1 follow through their mean and their z (to 2dp) in $\bar{x} \pm z \times \frac{5.1}{\sqrt{36}}$</p> <p>Beware: $167.94 \pm 2.5758 \times \frac{5.1^2}{36} \rightarrow (166.07\dots, 169.8\dots)$ but scores B1M0A0A0A0</p> <p>Correct answer only in (b) scores 0/5</p> <p>2nd & 3rd A marks depend upon 2.5758 and M mark.</p>		

Question number	Scheme	Marks												
2.	$\frac{115 \times 70}{217} = 37.0967... \quad \text{or} \quad \frac{1150}{31} \text{ (etc)} \quad \frac{1265}{31}, \frac{1020}{31}, \frac{1122}{31}$ <table border="1" data-bbox="245 376 1197 546"> <thead> <tr> <th>Expected (Obs)</th> <th>A</th> <th>S</th> <th>H</th> </tr> </thead> <tbody> <tr> <td>Boy</td> <td>37.1 (30)</td> <td>37.1 (50)</td> <td>40.8 (35)</td> </tr> <tr> <td>Girl</td> <td>32.9 (40)</td> <td>32.9 (20)</td> <td>36.2 (42)</td> </tr> </tbody> </table> <p> H_0 : There is no association between course and gender H_1 : There is some association between course and gender (both) </p> $\sum \frac{(O-E)^2}{E} = \frac{(37.1-30)^2}{37.1} + \frac{(32.9-40)^2}{32.9} + \dots + \frac{(36.2-42)^2}{36.2}$ <p> $= 1.358 + 4.485 + 0.824 + 1.532 + 5.058 + 0.929 = 14.189... \quad \text{awrt } \mathbf{14.2}$ </p> <p> $\nu = (3-1)(2-1) = 2, \quad \chi^2_2(1\%) \text{ critical value is } 9.210 \quad \text{(condone } 9.21)$ </p> <p> Significant result or reject null hypothesis There is evidence of an association between course taken and gender [Correct answers only score full marks] </p>	Expected (Obs)	A	S	H	Boy	37.1 (30)	37.1 (50)	40.8 (35)	Girl	32.9 (40)	32.9 (20)	36.2 (42)	M1 A1A1 B1 M1A1ft A1 B1, B1ft M1 A1ft (11) 11 marks
Expected (Obs)	A	S	H											
Boy	37.1 (30)	37.1 (50)	40.8 (35)											
Girl	32.9 (40)	32.9 (20)	36.2 (42)											
ALT	$\sum \frac{O^2}{E} - N = \frac{30^2}{37.1} + \frac{40^2}{32.9} + \dots + \frac{42^2}{36.2} - 217$	M1A1ft												
	<p>1st M1 for some use of the $\frac{\text{row total} \times \text{col total}}{\text{grand total}}$ formula</p> <p>1st A1 for one correct row or one correct column of expected frequencies to nearest integer</p> <p>2nd A1 for all expected frequencies correct to awrt 1 dp (Allow exact fractions)</p> <p>1st B1 for hypotheses. Independence is OK. Must mention courses and gender at least once. Use of ρ or “correlation” is B0 but allow ISW.</p> <p>2nd M1 for an attempt to calculate test statistic. At least one correct expression, ft expected freq.</p> <p>3rd A1 follow through expected frequencies for at least 3 expressions</p> <p>3rd M1 for a correct statement relating their test statistic and their cv (may be implied by comment)</p> <p>5th A1 for a contextualised comment relating their test statistic and their cv. Ignore their H_0 or H_1 or assume that they were correct. Must mention courses and gender</p>													

Question number	Scheme	Marks																																
3. (a)	<p>(i) </p> <p>(ii) </p> <p>(b)(i)</p> <table border="1" data-bbox="228 622 1120 786"> <thead> <tr> <th></th> <th>A</th> <th>B</th> <th>C</th> <th>D</th> <th>E</th> <th>F</th> <th>G</th> </tr> </thead> <tbody> <tr> <td>Rank (Judge 1)</td> <td>1</td> <td>4</td> <td>2</td> <td>3</td> <td>5</td> <td>6</td> <td>7</td> </tr> <tr> <td>Rank (Judge 2)</td> <td>1</td> <td>2</td> <td>4</td> <td>3</td> <td>5</td> <td>7</td> <td>6</td> </tr> <tr> <td>d^2</td> <td>0</td> <td>4</td> <td>4</td> <td>0</td> <td>0</td> <td>1</td> <td>1</td> </tr> </tbody> </table> <p style="text-align: right;">$\sum d^2 = 10$</p> <p>$r_s = 1 - \frac{6 \times 10}{7 \times (49 - 1)} = 1 - \frac{5}{28} = \frac{23}{28}$ or awrt 0.821</p> <p>(ii) $H_0 : \rho = 0$ $H_1 : \rho > 0$ (Allow ρ_s) ($H_1 : \rho \neq 0$ scores B0)</p> <p>r_s 5% one tail critical value is 0.7143</p> <p>Significant result or reject null hypothesis</p> <p>There is evidence of a (positive) correlation between the judges <u>or</u> the judges agree</p>		A	B	C	D	E	F	G	Rank (Judge 1)	1	4	2	3	5	6	7	Rank (Judge 2)	1	2	4	3	5	7	6	d^2	0	4	4	0	0	1	1	<p>(i) B1</p> <p>(ii) B1B1 (3)</p> <p>M1M1</p> <p>M1A1</p> <p>M1A1 (6)</p> <p>B1,B1</p> <p>B1</p> <p>M1</p> <p>A1ft (5)</p> <p style="text-align: right;">14 marks</p>
	A	B	C	D	E	F	G																											
Rank (Judge 1)	1	4	2	3	5	6	7																											
Rank (Judge 2)	1	2	4	3	5	7	6																											
d^2	0	4	4	0	0	1	1																											
	<p>(a) (i) 1st B1 for 5 or more points on a straight line of positive gradient</p> <p>(ii) 2nd B1 for 4 or more points satisfying $-1 < r < 0$</p> <p>3rd B1 for 5 or more points of decreasing ranks not on a straight line</p> <p>(b)(i) 1st M1 for attempting to rank one of the judges (at least 2 correct rankings)</p> <p>2nd M1 for ranking both (may be reversed) (at least 2 correct rankings)</p> <p>3rd M1 for attempting d^2.</p> <p>1st A1 for $\sum d^2 = 10$</p> <p>4th M1 for correct use of the r_s formula</p> <p>(ii) 3rd B1 for the correct critical value - depends upon their $H_1 : \rho > 0$ needs 0.7143, $\rho \neq 0$, 0.7857</p> <p>The H_1 may be in words so B0B1 is possible. If no H_1 award for 0.7143 only.</p> <p>5th M1 for a correct statement relating their r_s and their cv (may be implied by correct comment)</p> <p>3rd A1ft follow through their r_s and their cv. Comment in context. Must mention judges.</p> <p>Don't insist on "positive" and condone it if they are using $\rho \neq 0$.</p>																																	

Question number	Scheme	Marks
4. (a)	$X = M_1 + M_2 + M_3 + M_4 \sim N(336, 22^2)$ $\mu = 336$ $\sigma^2 = 22^2 \text{ or } 484$ $P(X < 350) = P\left(Z < \frac{350 - 336}{22}\right)$ $= P(Z < 0.64)$ $=$	B1 B1 M1 A1 A1 (5)
(b)	$M \sim N(84, 121) \text{ and } W \sim N(62, 100) \quad \text{Let } Y = M - 1.5W$ $E(Y) = 84 - 1.5 \times 62 = -9$ $\text{Var}(Y) = \text{Var}(M) + 1.5^2 \text{Var}(W)$ $= 11^2 + 1.5^2 \times 10^2 = 346$ $P(Y < 0), \quad = P(Z < 0.48\dots) =$	M1 A1 M1 A1 M1, A1 (6) 11 marks
(a)	2^{nd} B1 for $\sigma = 22$ or $\sigma^2 = 22^2$ or 484 M1 for standardising with their mean and standard deviation (ignore direction of inequality)	
(b)	1^{st} M1 for attempting to find Y . Need to see $\pm(M - 1.5W)$ or equiv. May be implied by $\text{Var}(Y)$. 1^{st} A1 for a correct value for their $E(Y)$ i.e. usually ± 9 . Do not give M1A1 for a “lucky” ± 9 . 2^{nd} M1 for attempting $\text{Var}(Y)$ e.g. $\dots + 1.5^2 \times 10^2$ or $11^2 + 1.5^2 \times \dots$ 3^{rd} M1 for attempt to calculate the correct probability. Must be attempting a probability > 0.5 . Must attempt to standardise with a relevant mean and standard deviation Using $\sigma_M^2 = 11$ or $\sigma_W^2 = 10$ is not a misread.	

Question number	Scheme	Marks
5. (a)	<p>Only cleaners - no managers i.e. not all <u>types</u>. OR Not a random sample 1st 50 may be in same shift/group/share <u>same views</u>. OR Not a random sample (Allow “not a representative sample” in place of “not a random sample”)</p> <p>(b)(i) Label employees (1-550) or obtain an ordered list Select <u>first</u> using <u>random numbers</u> (from 1 - 11) Then select every 11th person from the list</p> <p>(ii) Label managers (1-55) and cleaners (1-495) Use random numbers to select... ...5 managers and 45 cleaners</p> <p>(c) 390, 372 (They must be in this order)</p>	<p>B1g B1h (2)</p> <p>B1 B1 B1</p> <p>M1 M1 A1 (6)</p> <p>B1, B1 (2) 10 marks</p>
(a)	<p>After 1st B1, comments should be in context, i.e. mention cleaners, managers, types of worker etc</p> <p>1st B1g for one row 2nd B1h for both rows. “Not a random sample” only counts once. Score B1B0 or B1B1 or B0B0 on EPEN</p> <p>(b)(i) 1st B1 for idea of labelling or getting an ordered list. No need to see 1-550. 2nd B1 selecting first member of sample using random numbers (1-11 need not be mentioned) 3rd B1 selecting every nth where $n = 11$.</p> <p>(ii) 1st M1 for idea of <u>two</u> groups and labelling <u>both</u> groups. (Actual numbers used not required) 2nd M1 for use of random numbers within each strata. Don’t give for SRS from all 550. “Assign random numbers to managers and cleaners” scores M0M1 A1 for 5 managers <u>and</u> 45 cleaners. (This mark is dependent upon scoring at least one M)</p>	

Question number	Scheme	Marks
6. (a)	$p = \frac{0 \times 11 + 1 \times 21 + \dots}{10 \times (11 + 21 + \dots) \text{ or } 10 \times 100} = \frac{223}{1000} = 0.223 \quad (*) \quad \left(\text{Accept } \frac{223}{1000}\right)$	M1, A1cso (2)
(b)	$r = (0.8)^{10} \times 100 = 10.7374$	awrt 10.74 M1A1
	$s = \binom{10}{2} (0.8)^8 \times (0.2)^2 \times 100 = 30.198\dots$	awrt 30.2 A1
	$t = 100 - [r + s + 26.84 + 20.13 + 8.81] =$	awrt 3.28 A1cao (4)
(c)	H_0 : Binomial ($[n=10], p=0.2$) is a suitable model for these data	B1
	H_1 : Binomial ($[n=10], p=0.2$) is NOT a suitable model for these data	B1 (2)
(d)	Since $t < 5$, the last two groups are combined	M1
	and $\nu = 4 = 5 - 1$	A1 (2)
(e)	Critical value $\chi_4^2(5\%) = 9.488$	B1
	Not significant or do not reject null hypothesis	M1
	The binomial distribution with $p = 0.2$ is a suitable model for the number of	
	cuttings that do not grow	A1 (3)
13 marks		
(a)	M1 Must show clearly how to get either 223 or 1000. As printed or better.	
	A1cso for showing how to get <u>both</u> 223 and 1000 and reaching $p = 0.223$	
(b)	M1 for any correct method (a correct expression) seen for r or s .	
	1 st A1 for correct value for r awrt 10.74	
	2 nd A1 for $s =$ awrt 30.2	
	3 rd A1 for $t = 3.28$ only	
(c)	B1 for each. The value of p must be mentioned at least once. Accept B(10, 0.2)	
	If hypotheses are correct but with no value of p then score B0B1	
	Minimum is $X \sim B(10, 0.2)$. If just B(10, 0.2) and not B(10, 0.2) award B1B0	
(d)	M1 for combining groups (must be stated or implied by a new table with combined cell seen)	
	A1 for the calculation $4 = 5 - 1$	
(e)	M1 for a correct statement based on 4.17 and their cv(context not required) (may be implied)	
	Use of 4.17 as a critical value scores B0M0A0	
	A1 for a correct interpretation in context and $p = 0.2$ and cuttings mentioned.	

Question number	Scheme	Marks
7. (a)	$H_0 : \mu_F = \mu_M \quad H_1 : \mu_F \neq \mu_M \quad (\text{Allow } \mu_1 \text{ and } \mu_2)$ $z = \frac{6.86 - 5.48}{\sqrt{\frac{4.51^2}{200} + \frac{3.62^2}{100}}}$ $= 2.860\dots \quad \text{awrt } (\pm) \mathbf{2.86}$ <p>2 tail 5% critical value $(\pm) 1.96$ (or probability awrt 0.0021~0.0022)</p> <p>Significant result or reject the null hypothesis (o.e.)</p> <p>There is evidence of a difference in the (mean) amount spent on junk food by male and female teenagers</p>	B1 M1 A1 A1 B1 M1 A1ft (7) B1 (1) 8 marks
(b)	CLT enables us to assume \bar{F} and \bar{M} are normally distributed	
(a)	1 st M1 for an attempt at $\frac{a-b}{\sqrt{\frac{c}{100 \text{ or } 200} + \frac{d}{100 \text{ or } 200}}}$ with 3 of a, b, c or d correct 1 st A1 for a fully correct expression 2 nd B1 for ± 1.96 <u>but</u> only if their H_1 is two-tail (it may be in words so B0B1 is OK) If H_1 is one-tail this is automatically B0 too. 2 nd M1 for a correct statement based on comparison of their z with their cv. May be implied 3 rd A1 for a correct conclusion in context based on their z and 1.96. Must mention <u>junk food</u> or <u>money</u> and <u>male vs female</u> .	
(b)	B1 for \bar{F} or \bar{M} mentioned. Allow “ <u>mean</u> (amount spent on junk food) is <u>normally distributed</u> ” Read the whole statement e.g. “original distribution is normal so mean is...” scores B0	