



Mark Scheme (Results)

Summer 2015

Pearson Edexcel GCE
in Statistics 3 (6691/01)

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications are awarded by Pearson, the UK's largest awarding body. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information visit our qualifications websites at www.edexcel.com or www.btec.co.uk. Alternatively, you can get in touch with us using the details on our contact us page at www.edexcel.com/contactus.

Pearson: helping people progress, everywhere

Pearson aspires to be the world's leading learning company. Our aim is to help everyone progress in their lives through education. We believe in every kind of learning, for all kinds of people, wherever they are in the world. We've been involved in education for over 150 years, and by working across 70 countries, in 100 languages, we have built an international reputation for our commitment to high standards and raising achievement through innovation in education. Find out more about how we can help you and your students at: www.pearson.com/uk

Summer 2015

Publications Code UA042717

All the material in this publication is copyright

© Pearson Education Ltd 2015

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.
- Crossed out work should be marked UNLESS the candidate has replaced it with an alternative response.

PEARSON EDEXCEL GCE MATHEMATICS

General Instructions for Marking

1. The total number of marks for the paper is 75
2. The Edexcel Mathematics mark schemes use the following types of marks:
 - **M** marks: Method marks are awarded for 'knowing a method and attempting to apply it', unless otherwise indicated.
 - **A** marks: Accuracy marks can only be awarded if the relevant method (M) marks have been earned.
 - **B** marks are unconditional accuracy marks (independent of M marks)
 - Marks should not be subdivided.

3. Abbreviations

These are some of the traditional marking abbreviations that will appear in the mark schemes.

- bod – benefit of doubt
 - ft – follow through
 - the symbol \checkmark will be used for correct ft
 - cao – correct answer only
 - cso – correct solution only. There must be no errors in this part of the question to obtain this mark
 - isw – ignore subsequent working
 - awrt – answers which round to
 - SC: special case
 - oe – or equivalent (and appropriate)
 - d... or dep – dependent
 - indep – independent
 - dp decimal places
 - sf significant figures
 - * The answer is printed on the paper or ag- answer given
 - \square or d... The second mark is dependent on gaining the first mark
4. All A marks are 'correct answer only' (cao.), unless shown, for example, as A1 ft to indicate that previous wrong working is to be followed through. After a misread however, the subsequent A marks affected are treated as A ft, but manifestly absurd answers should never be awarded A marks.

5. For misreading which does not alter the character of a question or materially simplify it, deduct two from any A or B marks gained, in that part of the question affected.
6. If a candidate makes more than one attempt at any question:
 - If all but one attempt is crossed out, mark the attempt which is NOT crossed out.
 - If either all attempts are crossed out or none are crossed out, mark all the attempts and score the highest single attempt.
7. Ignore wrong working or incorrect statements following a correct answer.

June 2015
6691 S3
Mark Scheme

Question Number	Scheme	Marks																																												
1. (a)	Label all the books from 1 – 160 (o.e.) Use random numbers to select the 10 books	B1 B1 (2)																																												
(b)	<table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr> <th>Book</th> <th>A</th> <th>B</th> <th>C</th> <th>D</th> <th>E</th> <th>F</th> <th>G</th> <th>H</th> <th>I</th> <th>J</th> </tr> </thead> <tbody> <tr> <td>Borrow rank</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> <td>7</td> <td>8</td> <td>9</td> <td>10</td> </tr> <tr> <td>Page rank</td> <td>1</td> <td>6</td> <td>4</td> <td>2</td> <td>8</td> <td>3</td> <td>10</td> <td>7</td> <td>5</td> <td>9</td> </tr> <tr> <td>d^2</td> <td>0</td> <td>16</td> <td>1</td> <td>4</td> <td>9</td> <td>9</td> <td>9</td> <td>1</td> <td>16</td> <td>1</td> </tr> </tbody> </table> <p>$r_s = 1 - \frac{6 \times 66}{10(100-1)}, [= 1 - 0.4] = 0.6$ <u>0.6</u></p>	Book	A	B	C	D	E	F	G	H	I	J	Borrow rank	1	2	3	4	5	6	7	8	9	10	Page rank	1	6	4	2	8	3	10	7	5	9	d^2	0	16	1	4	9	9	9	1	16	1	M1 M1 M1,A1 (4)
Book	A	B	C	D	E	F	G	H	I	J																																				
Borrow rank	1	2	3	4	5	6	7	8	9	10																																				
Page rank	1	6	4	2	8	3	10	7	5	9																																				
d^2	0	16	1	4	9	9	9	1	16	1																																				
(c)	<p>$H_0: \rho = 0 \quad H_1: \rho > 0$ Critical value is 0.5636 $0.6 > cv$ so significant result and sufficient evidence to reject H_0 There is support for the librarian's belief <u>or</u> there is evidence of a correlation between the number of pages in a book and the number of times it is borrowed.</p>	B1 B1 B1ft (3)																																												
Notes		Total 9																																												
(a)	<p>1st B1 for labelling\numbering\listing\using sampling frame of all 160 books 2nd B1 for use of random numbers\selection and mentioning the number 10</p>																																													
(b)	<p>1st M1 for an attempt to rank the number of pages (at least 4 correct) Allow reverse ranks 2nd M1 for attempt at d^2 row (may be implied by sight of $\sum d^2 = 66$ or 264 for reverse ranks) 3rd M1 for use of the correct formula, follow through their $\sum d^2$ if clearly stated If answer is not correct, a correct expression is required. A1 for 0.6 (or -0.6 for reverse ranks)</p>																																													
(c)	<p>1st B1 for both hypotheses in terms of ρ, one tail H_1 (compatible with ranks) Allow use of ρ_s Hypotheses just in words e.g. “no correlation” score B0. 2nd B1 for cv of 0.5636 [If they have a two tail H_1 then allow 0.6485] Allow \pm for reverse ranking but must be same sign as r_s If hypotheses are the wrong way around this must be B0 but 3rd B1 is possible. 3rd B1ft for a correct contextualised comment. Must mention “librarian” (or he) <u>or</u> “number of pages” and “borrowing” Follow through their r_s and their cv (provided it is $cv < 1$) Don't insist on the word “positive” or “negative” for a one-tailed test Use of “association” is B0 Independent of 1st B1 so if $r_s > cv$ must say there is sufficient evidence of(o.e.) and if $r_s < cv$ must say insufficient evidence of ... (o.e.) regardless of their hypotheses</p>																																													

Question Number	Scheme	Marks
<p>2. (a)</p>	<p>$H_0 : \mu_g - \mu_s = 1.5$ [$g =$ in a group, $s =$ on their own]</p> <p>$H_1 : \mu_g - \mu_s > 1.5$</p> $\text{s.e.} = \sqrt{\frac{2.1^2}{80} + \frac{1.4^2}{65}} = [\sqrt{0.08527\dots}] = [0.292]$ $z = \frac{8.7 - 6.6 - 1.5}{\sqrt{\frac{2.1^2}{80} + \frac{1.4^2}{65}}}$ <p style="text-align: center;">$= 2.0546\dots$ awrt 2.05(5)</p> <p>cv 1% one tailed = 2.3263</p> <p style="text-align: center;">Not significant, accept H_0</p> <p>Insufficient evidence that using plan as part of a group leads to weight loss of more than 1.5 kg than using plan on one's own or researcher's belief not supported</p>	<p>B1</p> <p>B1</p> <p>M1</p> <p>dM1</p> <p>A1</p> <p>B1</p> <p>dM1</p> <p>A1ft</p> <p style="text-align: right;">(8)</p>
<p>(b)</p>	<p>Since sample is large Central Limit Theorem (CLT) applies No need to <u>assume</u> normal distribution</p>	<p>B1</p> <p>dB1</p> <p style="text-align: right;">(2)</p>
Notes		
<p>(a)</p>	<p>1st & 2nd B1 for hypotheses. Accept μ_1, μ_2 or μ_A, μ_B etc if there is some indication of which is which e.g. $G \sim N(\mu_g, 8.7)$</p> <p>1st M1 for an attempt at se with 3 out of 4 values correct. Condone switching 2.1 and 1.4</p> $\sqrt{\frac{2.1^2 \text{ or } 1.4^2}{80} + \frac{1.4^2 \text{ or } 2.1^2}{65}}$ <p>2nd dM1 dependent on 1st M1 for a correct numerator (must have -1.5) and ft their se.</p> <p>1st A1 for awrt 2.05</p> <p>3rd B1 for ± 2.3263 or better seen or probability of awrt 0.02</p> <p>3rd dM1 dep. on 1st M1 for a correct statement based on their normal cv and their test statistic</p> <p>2nd A1ft for correct comment in context. Must mention "plan" and "group or individual" and "1.5" <u>or</u> "researcher" and "belief or claim"</p> <p>NB Use of cv for difference in means D will have $D = 1.5 + 2.3263 \times \text{s.e.} =$ awrt 2.18 and requires sight of $d = 2.1$ with a comment for the 3rd M1</p>	
<p>(b)</p>	<p>1st B1 for mentioning "large samples" and "CLT"</p> <p>2nd dB1 dependent on 1st B1 for stating no need to assume normality (since CLT assures it)</p>	

Question Number	Scheme	Marks
3. (a)	Label staff (from 1 – 16) and children (from 1 – 40) Use random numbers to select 4 staff and 10 children	B1 B1 B1 (3)
(b)	$\bar{x} = \hat{\mu} = 31.2142\dots$ awrt <u>31.2</u> $s^2 = \frac{26983 - 14 \times "31.2\dots"}{13}$ $= 1026.33\dots$ awrt <u>1030</u>	B1 M1 A1ft A1 (4)
(c)	" $\frac{\sqrt{1026.33\dots}}{\sqrt{14}}$ ", = 8.562.. awrt <u>8.56</u>	M1, A1 (2)
(d)	The variation within each stratum is quite small (o.e.) The difference in the means will be quite large , (so variations from the overall mean will be large giving a larger overall s.e.)	B1 B1 (2)
Notes		Total 11
(a)	1 st B1 for labelling\numbering\listing staff <u>and</u> children 2 nd B1 for use of random numbers or “randomly select” in <u>each group</u> (may be implied) 3 rd B1 for selecting the correct number of staff <u>and</u> children e.g. randomly select 4 staff and 10 children scores 2 nd and 3 rd B marks since randomly selecting and the “each group” is implied,	
(b)	B1 for awrt 31.2 M1 for a correct expression ft their \bar{x} and allow transcription error in $\sum x^2$ e.g. 29683 1 st A1ft for a fully correct expression ft their \bar{x} only 2 nd A1 for awrt 1030	
(c)	M1 for attempting $\frac{\text{"their } s\text{"}}{\sqrt{14}}$ (must have 14) A1 for awrt 8.56	
(d)	1 st B1 for a suitable comment about variation (se) suggesting that variation (se) within strata is less than that overall 2 nd B1 for a suitable reason about means , pointing out that the individuals’ weights will vary a lot from the overall mean and so overall s.e. will be higher.	

Question Number	Scheme	Marks
<p>4. (a)</p> <p>(b)</p>	$H_0 : \mu = 0.5 \quad H_1 : \mu \neq 0.5$ <p>(Significance level =)10%</p> <p>(0.5 is in the interval so not significant, accept H_0, can accept) $\mu = 0.5$</p> $1.6449 \times \frac{\sigma}{\sqrt{100}} = 0.0247$ $\sigma = 0.15016 \quad \text{or} \quad \frac{10 \times 0.0247}{1.6449} \quad (\text{awrt } 0.15)$ $0.479 \pm 1.96 \times \frac{\sigma}{\sqrt{150}}$ <p style="text-align: right;">awrt <u>(0.455, 0.503)</u></p>	<p>B1 dB1 B1 (3)</p> <p>M1 B1 A1 M1 B1 A1 (6)</p> <p>Total 9</p>
Notes		
	<p>(a) 1st B1 for both hypotheses in terms of μ. 2nd dB1 for 10% but accept 5% if they have a one-tail test as H_1 3rd B1 for a correct comment leading to accepting H_0 Ignore any ‘further calculations’.</p> <p>(b) 1st M1 for $z \frac{\sigma}{\sqrt{100}} = k$, using $n = 100$ and where $z > 1.5$ and $0.02 < k < 0.03$ 1st B1 for 1.6449 or better in an attempt (could be $1.6449\sigma = k$ or even $1.6449 \sigma^2 = k$) 1st A1 for a correct expression for σ e.g. awrt 0.15 2nd M1 for $\bar{x} \pm z \times \frac{\sigma}{\sqrt{150}}$ for any $z (> 1)$ and ft their σ and allow $\bar{x} \in (0.4633, 0.5127)$ Allow use of letter σ without a value. 2nd B1 for 1.96 or better in an attempt (could be 1.96σ or even $1.96 \sigma^2$) 2nd A1 for awrt 0.455 <u>and</u> awrt 0.503</p>	

Question Number	Scheme	Marks
<p>5 (i)</p> <p>(ii)(a)</p> <p>(b)</p> <p>(c)</p> <p>(d)</p>	<p>Let $R = B_1 + B_2 + B_3 + B_4 + B_5 - 5H$ so $E(R) = -25$ (o.e.)</p> $\text{Var}(R) = 5 \times 6^2 + 5^2 \times 4^2$ $R \sim N(-25, \sqrt{580^2})$ <p>$P(R > 0) = P(Z > \frac{0-(-25)}{\sqrt{580}}) = P(Z > 1.04)$, = 0.149619...(calc) <u>or</u> 0.1492 (tables)</p> <p>$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{5}\right)$</p> $\text{Var}(D) = \sigma^2 + \frac{\sigma^2}{5} \left[= \frac{6\sigma^2}{5} \right], \text{ so } D \sim N\left(0, \frac{6\sigma^2}{5}\right)$ <p>$P(Y_1 > \bar{X} + \sigma) = P(D > \sigma) = P\left(Z > \frac{\sigma}{\sqrt{\frac{6}{5}\sigma^2}}\right)$</p> <p>= $P(Z > 0.912\dots) = 0.181(3 \text{ dp})$ (*)</p> <p>Since U_1 and \bar{U} are not independent (so variance formula cannot be used) Can be implied e.g. U_1 used to calculate \bar{U}, U_1 and \bar{U} from same sample o.e.</p> <p>Let $F = U_1 - \bar{U} = U_1 - \frac{(U_1 + U_2 + U_3 + U_4 + U_5)}{5} = \frac{4U_1 - (U_2 + U_3 + U_4 + U_5)}{5}$</p> $\text{Var}(F) = \frac{4^2\sigma^2 + 4\sigma^2}{5^2} = 0.8\sigma^2, \text{ so } F \sim N(0, 0.8\sigma^2)$ <p>$P(F > \sigma) = P\left(Z > \frac{\sigma}{\sigma\sqrt{0.8}}\right) = P(Z > 1.118\dots)$</p> <p>= 0.1314 (tables) or 0.131776...(calc) awrt 0.131-0.132</p>	<p>B1</p> <p>M1A1</p> <p>dM1 A1</p> <p>(5)</p> <p>B1</p> <p>M1, A1</p> <p>(3)</p> <p>M1</p> <p>A1cso</p> <p>(2)</p> <p>B1</p> <p>(1)</p> <p>M1, A1</p> <p>dM1, A1</p> <p>M1</p> <p>A1cso</p> <p>(6)</p> <p>Total 17</p>
Notes		
<p>(i)</p> <p>(ii)(a)</p> <p>(ii)(b)</p> <p>(c)</p> <p>(d)</p>	<p>1st B1 for $E(R) = -25$ (or 25 if their R is defined the other way around)</p> <p>1st M1 for an attempt at $\text{Var}(R) = 5\text{Var}(B) + 25\text{Var}(H)$. Condone swapping of 6^2 and 4^2</p> <p>1st A1 for normal and correct variance (ft their mean)</p> <p>2nd dM1 for attempting the correct probability and standardising with their mean and sd. This mark is dependent on 1st M1 so if R is not being used or M0 for variance score M0 If their method is not crystal clear then they must be attempting $P(Z > +ve \text{ value})$ o.e</p> <p>2nd A1 for answer in the range [0.149, 0.150]</p> <p>B1 for correct distribution of \bar{X} (may be implied for a correct answer for D)</p> <p>M1 for correct attempt at $\text{Var}(D)$ (ft their $\text{Var}(\bar{X})$) [A1 needs must be fully correct]</p> <p>M1 for expressing probability in terms of D and standardising</p> <p>A1cso for seeing $P(Z > 0.912\dots)$ or prob of $1 - 0.8186$ (tables) or $0.180655\dots$(calc)</p> <p>B1 correct statement that should mention U_1 and \bar{U}</p> <p>1st M1 for forming an expression in terms of $U_1\dots U_5$ only</p> <p>1st A1 for collecting U_1 terms and getting in a form where $\text{Var}(aX \pm bY)$ can be used.</p> <p>2nd dM1 for a correct expression for $\text{Var}(\text{their } F)$. Dependent on 1st M1.</p> <p>2nd A1 for a correct distribution for F</p> <p>3rd M1 attempting a correct prob and standardising using their $\text{Var}(F)$, σ must cancel</p> <p>3rd A1cso for awrt 0.131 or 0.132</p>	

Question Number	Scheme	Marks																									
<p>6. (a)</p>	<p>$H_0 : U[0, 10]$ is a suitable model $H_1 : U[0, 10]$ is not a suitable model</p> <table border="1" data-bbox="279 246 1029 459"> <thead> <tr> <th>D</th> <th>O_i</th> <th>E_i</th> <th>$\frac{(O_i - E_i)^2}{E_i}$</th> <th>$\frac{O_i^2}{E_i}$</th> </tr> </thead> <tbody> <tr> <td>0 – 4</td> <td>22</td> <td>40</td> <td>8.1</td> <td>12.1</td> </tr> <tr> <td>4 – 7</td> <td>39</td> <td>30</td> <td>2.7</td> <td>50.7</td> </tr> <tr> <td>7 – 9</td> <td>25</td> <td>20</td> <td>1.25</td> <td>31.25</td> </tr> <tr> <td>9 – 10</td> <td>14</td> <td>10</td> <td>1.6</td> <td>19.6</td> </tr> </tbody> </table> <p>Values of D Expected Freq 4th or 5th col $\chi^2 = 13.65$</p> <p>$\nu = 3, \chi_3^2(1\%) = 11.345$ [Reject H_0,] the uniform distribution over $[0, 10]$ is not a suitable model</p>	D	O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$	$\frac{O_i^2}{E_i}$	0 – 4	22	40	8.1	12.1	4 – 7	39	30	2.7	50.7	7 – 9	25	20	1.25	31.25	9 – 10	14	10	1.6	19.6	<p>B1 B1 M1A1 M1 A1 B1, B1 A1 (9) M1, A1 B1ft (3) M1, A1 (2) B1 (1) B1 (1) B1 (1) B1 B1 (2) Total 19</p>
D	O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$	$\frac{O_i^2}{E_i}$																							
0 – 4	22	40	8.1	12.1																							
4 – 7	39	30	2.7	50.7																							
7 – 9	25	20	1.25	31.25																							
9 – 10	14	10	1.6	19.6																							
Notes																											
<p>(a)</p>	<p>2nd B1 for the correct values for D (can be implied by 40, 30, 20, and 10.) 1st M1 for at least 2 expected frequencies or clear use of a correct formula e.g. $0.4N$ 1st A1 for all the correct E_i 2nd M1 for at least 2 correct calculations from 4th or 5th column 2nd A1 for a test statistic of 13.65 (accept 13.7 to 3 sf) Awrt 13.7 only scores 2nd B1M1A1M1A1 3rd A1 for a correct conclusion rejecting the uniform model. Award provided their test statistic > 11.345</p>																										
	<p>(b) M1 for some attempt to use πR^2 to find r</p>																										
	<p>(c) M1 for a correct statement that it is not significant A1 for correctly stating that Henry's model is suitable o.e.</p>																										
	<p>(d) B1 Independence or association mentioned at least once if ditto marks used. Allow connection but not correlation.</p>																										
	<p>(f) B1 for recognising there is an $E_i < 5$ and need for pooling/combining oe</p>																										
	<p>(g) 2ndB1 for correctly stating that Phoebe's belief is not supported by the data oe (depends on their cv being > 1.411)</p>																										

