Mark Scheme (Results)

January 2015

Pearson Edexcel International A Level in Statistics 1
(WST01/01)
Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications come from Pearson, the world’s leading learning company. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information, please visit our website at www.edexcel.com.

Our website subject pages hold useful resources, support material and live feeds from our subject advisors giving you access to a portal of information. If you have any subject specific questions about this specification that require the help of a subject specialist, you may find our Ask The Expert email service helpful.

www.edexcel.com/contactus

Pearson: helping people progress, everywhere
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

January 2015
Publications Code IA040679
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.
EDEXCEL IAL 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 \( \sqrt{ } \) 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)
   - dep – dependent
   - indep – independent
   - dp decimal places
   - sf significant figures
   - \* The answer is printed on the paper
   - 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. Ignore wrong working or incorrect statements following a correct answer.
1. (a) $(d) = 1$

(b) $a = 0.26 - 0.1 \text{ or } b = 0.26 + 0.28 \text{ or } 'a' + 0.38 \text{ or } 0.76 - 'c'$

\[ a = 0.16 \quad b = 0.54 \quad c = 0.22 \]

M1

A2

(c) 0.24 (only)

B1

(d) $P(X \text{ is an odd number}) = 0.1 + 0.28 + 0.24 = 0.62$

$P(X_1 \text{ and } X_2 \text{ are both odd}) = 0.62^2$

\[ = 0.3844 \quad \text{awrt 0.384} \]

A1

(e) $P(X_1 + X_2 = 6 | \text{ both are odd})$

\[ = \frac{P(X_1 + X_2 = 6 \cap X_1 \text{ and } X_2 \text{ are odd})}{P(X_1 \text{ and } X_2 \text{ are odd})} \]

\[ = \frac{0.1 \times 0.24 + 0.28 \times 0.28 + 0.24 \times 0.1}{(\text{their answer to } d)} \]

\[ = 0.1264 \quad \frac{(their \text{ answer to } d)}{\text{awrt 0.329}} \quad \text{A1ft} \]

A1

### Notes

(a) B1 for sight of 1 referring to $d$ (may be in table or in the question)

(b) M1 for any correct calculation seen (may be implied by one correct answer)

A1 for at least two values correct

A2 for all 3 values correct

(d) M1 for $(0.1 + 0.28 + 0.24)^2$ oe i.e. must be a complete correct expression

\[ \text{e.g. } (1 - ['a' + 'c'])^2 \text{ and ft their values for 'a' and 'c'} \]

A1 for awrt 0.384 or exact fraction \( \frac{961}{2500} \)

(e) M1 for attempt at correct conditional probability i.e. a correct ratio of probabilities stated in words that mentions both $X_1$ and $X_2$

May be implied by a numerical ratio with correct num’ and their “(d)” on denom’

This would score M1A1ft

1st A1ft for \( \frac{\text{correct numerator}}{0.384..} \) or correct numerator and denominator of their ‘d’

2nd A1 for awrt 0.329 or exact fraction \( \frac{316}{961} \)
2. (a) Year 7 median = 29  
    Year 11 median = 54  
    1st B1 for 29 seen  
    2nd B1 for 54 seen  
    (2)

(b) [Lower quartile =] 22  
    [Upper quartile =] 42  
    1st B1 for 22 and 2nd B1 for 42 (these values may be circled on the diagram)  
    (2)

(c) Year 7  
    \( Q_3 - Q_2 = 13 > Q_2 - Q_1 = 7 \) 
    Positive skew  
    1st A1 for Year 7 clearly labelled “positive skew” (both words) (“correlation” is A0)  
    (3)

Year 11  
    \( Q_3 - Q_2 = 5 < Q_2 - Q_1 = 16 \) 
    Negative skew  
    2nd A1 for Year 11 clearly labelled “negative skew” (both words) (“correlation” is A0)  

(d) Data is skewed  
    Data is not continuous  
    1st B1 for a statement mentioning (or implying) that the data is skewed (or not symmetric) Ignore ref to +ve or -ve  
    2nd B1 for a statement mentioning data is not continuous (allow identifiable spelling)  
    (2) (9 marks)

Notes

(a) In (a) at least one of the values should be assigned to a Year group  
If you see just “29” and “54” award SC B1B0  
1st B1 for 29 seen (may be circled on diag.)  
2nd B1 for 54 seen  
(b) 1st B1 for 22 and 2nd B1 for 42 (these values may be circled on the diagram)  
(c) M1 for a comparison for either year using quartiles only.  
For either “\( Q_3 - Q_2 > Q_2 - Q_1 \) and positive skew” or “\( Q_3 - Q_2 < Q_2 - Q_1 \) and negative skew”  
Statements should be compatible with their values  
1st A1 for Year 7 clearly labelled “positive skew” (both words) (“correlation” is A0)  
2nd A1 for Year 11 clearly labelled “negative skew” (both words) (“correlation” is A0)  
Ans. only If no comparison is stated then award M1A1A1 only if both statements are correct and compatible with their medians and quartiles so score is 0 or 3  
(d) 1st B1 for a statement mentioning (or implying) that the data is skewed (or not symmetric) Ignore ref to +ve or -ve  
SC Allow for statement “mean ≠ median” if mean = 48.8 and median = 54 or 53 seen  
2nd B1 for a statement mentioning data is not continuous (allow identifiable spelling)  
Allow “this data is discrete” for 2nd B1  

NB means  
Year 7 \( \bar{x} = 31.5 \)  
Year 11 \( \bar{x} = 48.8 \)
### 3. (a)  
\[ 29 \times 75 + 29 \times 83 + \ldots + 46 \times 126 = 33856 \]  
33856  
B1cao  
(1)

(b)  
\[ \sum m = 306 \text{ and } \sum b = 861 \]  
\[ S_{bm} = \left( \frac{33856 - (861 \times 306)}{8} \right) = 922.75 \]  
awrt 923  
M1 A1  
(3)

(c)  
\[ r = \frac{"922.75"}{\sqrt{3083.875 \times 305.5}} = 0.9506706\ldots \]  
awrt 0.951  
M1 A1  
(2)

(d) As milk price increase, so does bread price.  
B1  
(1)

(e) Since bread price increases but milk price stays the same  
Therefore the correlation will decrease (or be weaker)  
B1  
dB1  
(2)  
(9 marks)

### Notes

(a) B1 for 33856 as their final answer

(b) B1 for both \( \sum m \) and \( \sum b \) seen or implied by \( 861 \times 306 = 263466 \) or a correct answer  
These must be seen in (b) do not allow for \( \sum m + \sum b = 306 + 861 = 1167 \) just in (a)  
M1 for use of correct formula ft their answer to (a)  
A1 for awrt 923  
[Answer only scores B1M1A1]

(c) M1 for attempt at correct formula. Must have their \( S_{bm} \) and the given values of \( S_{bb} \) and \( S_{mm} \) (3sf or better) in the correct places. NB \( \sqrt{3083.875 \times 305.5} = 970.63\ldots \)  
(0.95 with no working score M1 A0). Allow M1 even if \(|r| > 1\)  
A1 for awrt 0.951  
[Answer only of awrt 0.951 scores M1A1]

(d) B1 for a contextual description of positive correlation.  
Must use words “milk” and “bread” so “as \( m \) increases \( b \) increases” is B0  
Ignore any mention of correlation or skewness if a correct interpretation is given.

(e) 1st B1 for a suitable reason  
e.g. \( m = 46, b = 175 \) does not follow trend/pattern or is an outlier  
or new point will be further from the (regression) line or 175 is more than expected  
NB “175 is larger than all values in table” is B0 since it makes no ref. to reg line or milk price.  
BUT “175 is an extreme value (or outlier)” implies the point is being considered and is B1  

2nd dB1 dep. on 1st B1 for saying \( r \) (or “it”) will decrease (allow weaker correlation)  
Mention of “skew decreases” is B0 unless there is a correct statement as well.

NB The new value of \( r = 0.86767\ldots \). You may see this but it does not score anything.
4. (a)(i) \( x + 0.1 \) \[ P(x + 0.1) \text{ is B0} \]  
(b) \( x + y + 0.1 \) \[ P(x + y + 0.1) \text{ is B0} \]  
(c) \[ x + y + 0.1 + 0.32 = 1 \] or \[ x + y + 0.1 = 0.68 \] or \( \text{“(b)”} + 0.32 = 1 \) o.e.  
Eliminating \( x \) gives \( 3y = 0.48 \)  
\( x = \boxed{0.42} \)  
\( y = \boxed{0.16} \)  

Notes  
(a)(ii) M1 for a correct ratio of probabilities formula with at least one correct probability value (may fit their (a)(i) in the denominator) or a prob ratio of the form \( \frac{0.1}{(a)(i)} \)  
If num’ > denom’ score M0. NB \( P(A) = 0.68 - y \) and \( P(B|A) = \frac{0.1}{0.68 - y} \) is B0M1A0  
A1 for \( \frac{0.1}{x + 0.1} \) as their final answer  
(b) B1 for any correct expression in \( x \) and \( y \) e.g. \( 0.1 + x + 0.1 + y - 0.1 \)  
Condone \( x + y + 0.1 = 1 - 0.32 \) or 0.68 since LHS is a correct expression  
(c) 1\(^{st} \) M1 for using sum of probs. = 1 to form a “correct” linear equ’n in \( x \) and \( y \) \( [x + y = 0.58] \)  
Ft their (b) and or their (a)(i) e.g. “(a)(i)” +0.32 + \( y = 1 \)  
2\(^{nd} \) M1 for using \( P(A) = 2P(B) \) to form a “correct” linear equ’n in \( x \) and \( y \) \( [x - 2y = 0.1] \)  
Ft their \( P(A) \) from part (a)  
If they use 2\(^{nd} \) \( P(A) = P(B) \) or swap \( x \) and \( y \) score 2\(^{nd} \) M0 but allow access to 3\(^{rd} \) M  
3\(^{rd} \) M1 for an attempt to solve their 2 linear equations. Implied by 1\(^{st} \) 2 Ms and correct ans.  
Requires correct algebraic steps leading to an equation in one variable.  
If there are not 2 equations this cannot be scored (but see SC)  
1\(^{st} \) A1 for \( x = 0.42 \) (following correct working and dep. on 1\(^{st} \) 2 Ms)  
2\(^{nd} \) A1 for \( y = 0.16 \) (following correct working and dep. on 1\(^{st} \) 2 Ms)  
Beware  
\( 0.42 = 0.32 + 0.1 \) so answer only does not score full marks  
SC  
\( P(A) = 0.68 - y = 2(y + 0.1) \) score M2 (2\(^{nd} \) and 3\(^{rd} \) Ms) and 2\(^{nd} \) A1 when \( y = 0.16 \) seen  
Sight of \( x + y + 0.1 = 0.68 \) (o.e.) (scores 1\(^{st} \) M1) and then 1\(^{st} \) A1 if \( x = 0.42 \) follows.  
or  
\( P(A) = x + 0.1 = 2(0.68 - x) \) score M2 (2\(^{nd} \) and 3\(^{rd} \) Ms) and 1\(^{st} \) A1 when \( x = 0.42 \) seen  
Sight of \( x + y + 0.1 = 0.68 \) (o.e.) (scores 1\(^{st} \) M1) and then 2\(^{nd} \) A1 if \( y = 0.16 \) follows.
5. (a) Resting heart rate, \( h \), is being measured (you can’t control it)
   So it is the response variable

(b) For every additional minute of exercise, heart rate decreases by 0.43 (bpm)

(c) \[ \bar{t} = 50 \quad \bar{h} = 72 \]

(d) \( h = 93.5 - 0.43 \times 50 \) so \( h = 72 \) or
   Allow: \( 72 = 93.5 - 0.43 \times 50 \)

(e) \[ h = 93.5 - 0.43 \times 60 \] \( h = 67.7 \) (allow 68 if a correct expression is seen)

(f) Since 1 hour (60 minutes) is within the range (of the \( t \)-values),
   The estimate is reliable

(g) \[
\frac{a - 73}{8} = -1.96 \quad \text{or} \quad \frac{b - 73}{8} = 1.96 \\
73 \pm 1.96 \times 8 \\
(57.32, 88.68)
\]

\textbf{awrt 57.3 and 88.7}

<table>
<thead>
<tr>
<th>Notes</th>
</tr>
</thead>
</table>
| (a) \( 1^{st} \) B1 for a reason that doesn’t use words “response” or “explanatory”  
   e.g. \( h \) is dependent on/ affected by/changed by/influenced by/determined by \( t \)  
   or \( t \) is being controlled  
   \( 2^{nd} \) dB1 dep. on \( 1^{st} \) B1 for choosing \( h \) as the response variable |
| (b) B1 for a correct interpretation in context. Need mention of “exercise” plus a unit of time  
   and “heart rate” or “beats” with a correct corresponding value. No need for bpm.  
   (Just saying “increase of \( t \) by 1 means decrease of \( h \) by 0.43 is B0...need words!) |
| (c) \( 1^{st} \) B1 for 50 and \( 2^{nd} \) B1 for 72 |
| (d) B1 cso allow a correct expr’ with all 4 numbers in the correct places without a comment |
| (f) \( 1^{st} \) B1 for a reason. Allow \( t \) or time or 60 is within data or “interpolation”.  
   “Its” is B0B0. If they say both \( t = 60 \) and \( h = 67.7 \) are within range then B0B0  
   unless they later specify that \( t \) is intended or mention “interpolation”  
   \( 2^{nd} \) dB1 dep. on \( 1^{st} \) B1, for saying it is reliable (o.e. e.g. “accurate”) |
| (g) \( 1^{st} \) M1 for \( \frac{a - 73}{8} = z \) or \( \frac{b - 73}{8} = z \) with \(|z| > 1\), must be a \( z \)-value  
   B1 for 1.96 seen and used as a \( z \) value. NB 1 – 1.96 is not a \( z \) value and scores B0  
   \( 2^{nd} \) dM1 dep. on \( 1^{st} \) M1 for rearranging to find \( a \) or \( b \)  
   \( 73 \pm z \times 8 \)  
   A1 for both \( a = \text{awrt} \ 57.3 \) and \( b = \text{awrt} \ 88.7 \)  
   Both values seen and correct then answer only scores 4/4 |

\textbf{Ans only}
### 6. (a) 
\[
\frac{1^2 + 2^2 + 3^2 + 4^2}{k} = 1, \quad \frac{30}{k} = 1,
\]
so \( k = 30 \)  

Or verify
\[
\frac{1^2 + 2^2 + 3^2 + 4^2}{30} = 1
\]

M1  

**(2)**

(b) 
\[
1 - P(X = 4), \quad 1 - \frac{16}{30} = \frac{7}{15} \quad \text{(or exact equiv. e.g. \( \frac{14}{30} \) or 0.46)}
\]

M1, A1

(c) 
\[
E(X) = \left[ 1 \times \frac{1}{30} + 2 \times \frac{4}{30} + 3 \times \frac{9}{30} + 4 \times \frac{16}{30} \right] = \frac{10}{3}
\]

M1, A1

(d) 
\[
E(X^2) = \left[ 1^2 \times \frac{1}{30} + 2^2 \times \frac{4}{30} + 3^2 \times \frac{9}{30} + 4^2 \times \frac{16}{30} \right] = \frac{354}{30} (= 11.8)
\]

E(X) = \( \frac{31}{45} \) (or exact equivalent e.g. 0.68)  

M1, A1

(e) 
\[
E(Y) = 3E(X) - 1 \quad (= 9)
\]

M1  

Var(Y) = 3^2 Var(X) \quad (= 6.2)

E(Y^2) = Var(Y) + E(Y)^2 = 6.2 + 9^2, \quad \text{= 87.2 (o.e. e.g. \( \frac{436}{5} \))}

M1, A1

### Notes

(a) M1 for clear use of sum of probs. = 1 (Minimum is \( k = 1 + 2^2 + 3^2 + 4^2 \))  
A1 for correct conclusion with no incorrect working seen

(b) M1 for \( 1 - P(X=4) \) or \( P(X=1) + P(X=2) + P(X=3) \)

(c) M1 for attempt at correct expression for \( E(X) \) (at least 3 correct products)

(d) \( 1 \)st M1 for attempt at correct expression for \( E(X^2) \) (at least 3 correct products)  
\( 1 \)st A1 for 11.8 o.e. may be implied by fully correct sol’n. Condone \( \text{Var}(X) = E(X^2) \) for M1A1  
\( 2 \)nd M1 for using \( \text{Var}(X) \) formula with correct substitution, may fit their \( E(X) \) and \( E(X^2) \)  
If \( \text{Var}(X) < 0 \) score \( 2 \)nd M0

(e) \( 1 \)st M1 for finding \( y = 2, 5, 8 \) and 11 (at least 3 correct)  
\( 2 \)nd M1 for a correct prob. distribution for \( Y \) so \( P(Y = 2) = \frac{1}{30}, P(Y = 5) = \frac{4}{30} \) etc  
\( 3 \)rd M1 for \( E(Y^2) = (2)^2 \times \frac{1}{30} + (5)^2 \times \frac{4}{30} + (8)^2 \times \frac{9}{30} + (11)^2 \times \frac{16}{30} \) (at least 3 correct)

ALT 1

Prob dist’n

2nd M1 for a correct prob. distribution for \( Y \) so \( P(Y = 2) = \frac{1}{30}, P(Y = 5) = \frac{4}{30} \) etc  
3rd M1 for \( E(Y^2) = (2)^2 \times \frac{1}{30} + (5)^2 \times \frac{4}{30} + (8)^2 \times \frac{9}{30} + (11)^2 \times \frac{16}{30} \) (at least 3 correct)

ALT 2

E[(3X – 1)^2]  

1st M1 attempt correct expression e.g. \( E(aX^2 + bX + c) \) for any \( a, b \) and \( c \)  
2nd M1 for \( 9 \text{E}(X^2) \)  
3rd M1 for \( E(Y^2) = 9 \text{E}(X^2) - 6 \text{E}(X) + 1 \)
7. (a) \[ P(W > 92) = P(Z > \frac{92 - 99}{3.6}) = P(Z > -1.94) \text{ or } P(Z < 1.94) = 0.9738 \text{ awrt 0.974} \quad \text{(3)} \]

(b) \[ P(W < k) = 3P(W > k) \text{ so } P(W < k) = 0.75 \text{ or } P(W > k) = 0.25 \]
\[ \frac{k - 99}{3.6} = 0.67 \quad (k = 101.4) \]

(c) \[ k \text{ is the upper quartile} \]

(d) \[ P(W < P_{20}) = 0.2 \]
\[ \frac{116 - 120}{\sigma} = -0.8416 \]
\[ \sigma = 4.7528517... \text{ awrt 4.75} \quad \text{(3)} \]

Notes

(a) M1 for standardising with 92, 99 and 3.6
1st A1 for either correct probability statement and z awrt ±1.94 (may be seen as a correct shading on a diagram).
2nd A1 for awrt 0.974

NB They may get \( z = 1.945 \) and round to 1.95 leading to 0.9744 (score M1A0A1)

(b) 1st B1 for \( P(W < k) = 0.75 \) or \( P(W > k) = 0.25 \) (o.e.) [May be implied by \( k = \text{awrt 101.4} \)]

NB B0M1B1A1 is possible if an incorrect statement e.g. \( P(W < k) = 0.25 \) is seen for an attempt to standardise with \( k \) (or any letter), 99 and 3.6 and set equal to \( \pm \) a \( z \)-value in range 0.6 ~ 0.7

2nd B1 for \( \pm 0.67 \) or better i.e. \( z \) in 0.670 ~ 0.678 (calc gives 0.674489...)
NB e.g. 0.68 is B0 but could score A1.

A1cao for 101.4 (must be given to 1dp) and must follow from compatible signs

Ans. only

(c) B1 for Upper quartile (allow \( Q_3 \) or third quartile or 75th percentile)

(d) M1 for an attempt to standardise and set equal to \( \pm \) a \( z \)-value in 0.8 ~ 0.9
B1 for \( \pm 0.8416 \) or better (calc gives 0.84162123...). Value must be used as a \( z \) value
NB 0.84 scores B0 but see SC

A1 for awrt 4.75 following from an equation with compatible signs

SC If they use \( z = 0.84 \) and get an answer of awrt 4.76 (with correct working) score M1B0A1