



Mark Scheme (Results)

Autumn 2020

Pearson Edexcel GCE In AS Level Statistics
(8ST0/01)

Paper 1

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

Autumn 2020

Publications Code 8ST0_01_2010_MS

All the material in this publication is copyright

© Pearson Education Ltd 2020

General Marking Guidance

Total marks

The total number of marks for the paper is 60.

Mark types

The Edexcel Statistics mark schemes use the following types of marks:

- **M** **Method** marks, awarded for 'knowing a method and attempting to apply it',
 unless otherwise indicated.
- **A** **Accuracy** marks can only be awarded if the relevant method (M) marks have been earned.
- **B** **Unconditional accuracy** marks are independent of M marks
- **E** **Explanation** marks

NOTE: Marks should not be subdivided.

Abbreviations

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

- ft follow through
- PI possibly implied
- cao correct answer only
- cso correct solution only
 (There must be no errors in this part of the question)
- awrt answers which round to
- awfw answers which fall within (a given range)
- SC special case
- nms no method shown
- oe or equivalent
- dep dependent (on a given mark or objective)
- dp decimal places
- sf significant figures
- ts test statistic
- cv critical value
- * The answer is printed on the paper

Further notes

- 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.
- All A marks are 'correct answer only' (cao.), unless shown, for example, as A1ft to indicate that previous wrong working is to be followed through.
- After a **misread**, the subsequent A marks affected are treated as A1ft, but manifestly absurd answers should never be awarded A marks.
- **Crossed out** work should be marked UNLESS the candidate has replaced it with an alternative response.
- If **two solutions** are given, each should be marked, and the resultant mark is the mean of the two marks, rounded down to the nearest integer if needed.

| Question | Scheme | Marks | AO | Notes |
|--------------|---|--------------------|------|--------------------------------------|
| 1(a) | John has a median of 51 & Nadiya of 48 so (they are very similar or John's is higher) | | | |
| | Nadiya has the student who scored the lowest mark | | | |
| | John has the student who scored the highest mark | | | |
| | John's IQR=22 < Nadiya's IQR=25 | | | |
| | John's range = 63, Nadiya's range =88 or 60 | | | |
| | As John's $Q_3=70$ and Nadiya's $Q_3<70$, John has more students over 70. | | | Allow any true specific comment |
| | | B1, B1, B1, B1, B1 | 1.1 | Any 5 comments from above or similar |
| 1(b) | John's sample is much smaller | | | |
| | Assessment might have been easier this year than previously | | | |
| | John and Nadiya may be assessing students with work of different standards where it may be easier for John's students to achieve higher marks | | | |
| | John's students may have tutors | | | |
| | Test may be different in different years | | | |
| | Higher marks doesn't necessarily mean better teacher because they may have different ability students | | | |
| | | E1, E1 | 3.1a | Any 2 reasons |
| Total | | 7 | | |

| Question | Scheme | Marks | AO | Notes |
|--------------|--|----------|------|-------|
| 2(a) | Better drivers might pass on first attempt/earlier/when younger | E1 | 3.1b | |
| | Driving test could change over time/someone who is 50 now could have sat a different test before | E1 | 3.1b | |
| 2(b) | Data is only for people who have passed a driving test – some people are excluded e.g. under 17 or people who have sat a test and not passed | E1 | 3.1b | |
| 2(c) | Calculate $E[X]$ with $p_4=0.15$ giving $E[X]=1.95$ | B1 | 1.2 | PI |
| | But 4 was 4+ so final answer will be 1.95 or more | E1 | 2.1b | |
| Total | | 5 | | |

| Question | Scheme | Marks | AO | Notes |
|--------------|--|----------|------|---|
| 3(a) | 1 head in first 5 and 1 head in second 5 is contained within 2 heads in all 10 | E1 | 2.1b | oe |
| | So Tim is correct | B1dep | 2.1b | Dep on valid reason |
| 3(b) | Probability is always a half regardless of previous flips | E1 | 1.2 | oe or result is independent of previous flips |
| | So Alexandra is wrong | B1dep | 3.1b | Dep on valid reason |
| Total | | 4 | | |

| Question | Scheme | Marks | AO | Notes |
|--------------|--|----------|------|--|
| 4(a) | Data not paired or Different number in each category | E1 | 3.1a | |
| 4(b) | Wilcoxon rank-sum test | B1 | 2.1a | Accept Mann-Whitney test |
| 4(c) | Colin used a larger sample so his result is more reliable | E1 | 3.1a | |
| 4(d) | The Wilcoxon rank-sum test is non-parametric (and requires no distributional assumption for its validity) | E1 | 2.1a | Accept Mann-Whitney is non-parametric |
| Total | | 4 | | |

| Question | Scheme | Marks | AO | Notes | | | | | | |
|--|--|--------|-----------------|-------------------------------|--------|--------|--------|----|-----|---|
| 5(a) | Advantage: Cheaper/less time consuming to carry out as all clubs are based in London. | E1 | 1.1 | oe | | | | | | |
| | Disadvantage: As it only includes clubs in London it may not be representative of whole country | E1 | 1.1 | oe Each must be in context | | | | | | |
| 5(b) | Attempt at chi squared test | M1 | 2.1a | | | | | | | |
| | H ₀ : No association between league/nationality signed H ₁ : Association between league/nationality signed | B1 | 1.3 | Hypotheses in context | | | | | | |
| | Attempt to calculate expected values | M1 | 1.2 | PI | | | | | | |
| | Expected values: <table border="1" data-bbox="375 1003 842 1216"> <tr> <td>12.606</td> <td>13.394</td> </tr> <tr> <td>8.242</td> <td>8.758</td> </tr> <tr> <td>11.152</td> <td>11.848</td> </tr> </table> | 12.606 | 13.394 | 8.242 | 8.758 | 11.152 | 11.848 | A1 | 1.2 | PI Any 2 in different rows correct At least one decimal place |
| | 12.606 | 13.394 | | | | | | | | |
| | 8.242 | 8.758 | | | | | | | | |
| | 11.152 | 11.848 | | | | | | | | |
| | Test stat formula used correctly | M1 | 1.3 | PI | | | | | | |
| | Correct values before addition <table border="1" data-bbox="375 1357 842 1570"> <tr> <td>3.461</td> <td>3.258</td> </tr> <tr> <td>0.0697</td> <td>0.0656</td> </tr> <tr> <td>3.067</td> <td>2.886</td> </tr> </table> | 3.461 | 3.258 | 0.0697 | 0.0656 | 3.067 | 2.886 | | | |
| | 3.461 | 3.258 | | | | | | | | |
| 0.0697 | 0.0656 | | | | | | | | | |
| 3.067 | 2.886 | | | | | | | | | |
| ts = 12.808 | A1 | 1.3 | awfw 12.5- 13.0 | | | | | | | |
| 2 degrees of freedom | B1 | 1.3 | PI | | | | | | | |
| cv = 5.991 Or p-value = 0.00165 compared to 0.05 | B1 | 1.3 | | | | | | | | |

| Question | Scheme | Marks | AO | Notes |
|------------------------------|---|-----------|--------|---|
| 5(b) (cont) | Reject H_0 | B1dep | 2.1b | PI Dep on correct ts and cv |
| | There is significant evidence at the 5% level to suggest that there is an association between the league a team plays in and the nationality of the player signed | E1dep | 2.1a | Dep on correct ts and cv |
| 5(c) | Greatest source of association is in the Premier League where fewer British players were signed than would be expected | E1 | 2.1a | Condone Premier League clubs sign more players who are not British oe |
| 5(d) | The situation in the leagues has changed considerably during this period - there is now far more finance and international business involved | E1 | 3.1a | Accept e.g. Result can change over time oe |
| 5(e) | Possible suggestions (not exhaustive) | | | |
| | Stratify based on location | | | |
| | Use a sample covering more years | | | |
| | Use a sample with more clubs chosen at random from the 3 leagues throughout England and Wales | | | |
| | Random sample of all players | | | |
| | Take a larger sample | | | |
| | | | E1, E1 | 3.1a |
| Total | | 16 | | |

| Question | Scheme | Marks | AO | Notes |
|-----------------|---|----------|------|--|
| 6(a) | Possible comments | | | |
| | Data clearly not bell shaped | | | Or e.g. data dips below 12 & 14 |
| | Data appears to be bi-modal | | | |
| | Mean and median/mode are not equal so not symmetric | | | Mean=12.9 Or data is skewed |
| | | E1 | 2.1b | Sensible comment on shape of distribution |
| 6(b) | $11 - 2 \times 6 = -1$ | B1 | 1.2 | Or $11 - 3 \times 6 = -7$ |
| | Less than zero so normal distribution not appropriate | E1dep | 2.1b | Dep on some evidence |
| | | | | SC Rita may have misremembered scores B0E1 |
| 6(c)(i) | $P(X > 16) = 0.19568$ | B1 | 1.2 | awrt 0.196 |
| 6(c)(ii) | $Y \sim B(10, p)$ | M1 | 2.1a | Their p from (c)(i) |
| | $P(Y \geq 3)$ seen | M1 | 1.2 | PI Or $P(Y \leq 2)$ seen |
| | 0.3092 | A1 | 1.2 | awrt 0.309 or 0.310 |
| Total | | 7 | | |

| Question | Scheme | Marks | AO | Notes | | | | | | | | | | | | | | | | | | | | | | |
|------------|--|--------|---|--|-----|-----|---|-----|---|---|-----|---|---|---|---|----|------|---|---|---|---|----|------|----|-----|------------------|
| 7(a) | Systematic Sample | B1 | 1.1 | | | | | | | | | | | | | | | | | | | | | | | |
| | Not all combinations are possible | E1 | 1.1 | Or other legitimate reason | | | | | | | | | | | | | | | | | | | | | | |
| 7(b)(i) | $P(17 \text{ in sample}) = \frac{1}{10}$ | B1 | 1.2 | | | | | | | | | | | | | | | | | | | | | | | |
| 7(b)(ii) | $P(\text{Both in sample}) = \frac{1}{10}$ | B1ft | 1.2 | Their answer to part (b)(i) | | | | | | | | | | | | | | | | | | | | | | |
| 7(b)(iii) | 0 | B1 | 1.2 | | | | | | | | | | | | | | | | | | | | | | | |
| 7(c) | Sort on Weight in Ascending order | B1, B1 | 1.1 | Need all 3 for 2 marks, 1 mark if omitting 1 | | | | | | | | | | | | | | | | | | | | | | |
| 7(d) | H ₀ : no association H ₁ : positive association 1 tail, 5% | B1 | 1.3 | | | | | | | | | | | | | | | | | | | | | | | |
| | Attempt at evaluation of a correlation coefficient | M1 | 2.1a | SC Condone PMCC | | | | | | | | | | | | | | | | | | | | | | |
| | Attempt at ranking data | M1 | 1.3 | | | | | | | | | | | | | | | | | | | | | | | |
| | Correct use of ties <table border="1" style="margin-left: auto; margin-right: auto;"> <tbody> <tr><td>3.5</td><td>4</td></tr> <tr><td>3.5</td><td>6.5</td></tr> <tr><td>1.5</td><td>1</td></tr> <tr><td>1.5</td><td>3</td></tr> <tr><td>5</td><td>6.5</td></tr> <tr><td>9</td><td>5</td></tr> <tr><td>6</td><td>9</td></tr> <tr><td>10</td><td>10.5</td></tr> <tr><td>7</td><td>2</td></tr> <tr><td>8</td><td>8</td></tr> <tr><td>11</td><td>10.5</td></tr> </tbody> </table> | 3.5 | 4 | 3.5 | 6.5 | 1.5 | 1 | 1.5 | 3 | 5 | 6.5 | 9 | 5 | 6 | 9 | 10 | 10.5 | 7 | 2 | 8 | 8 | 11 | 10.5 | M1 | 1.3 | Condone reversed |
| | 3.5 | 4 | | | | | | | | | | | | | | | | | | | | | | | | |
| 3.5 | 6.5 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1.5 | 1 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1.5 | 3 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 5 | 6.5 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 9 | 5 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 6 | 9 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 10 | 10.5 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 7 | 2 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 8 | 8 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 11 | 10.5 | | | | | | | | | | | | | | | | | | | | | | | | | |
| t.s.=0.704 | A1 | 1.3 | awrt 0.704 r _s =0.707 with formula No mark for PMCC= 0.751 | | | | | | | | | | | | | | | | | | | | | | | |

| Question | Scheme | Marks | AO | Notes |
|----------------|--|-----------|------|---|
| 7(d) (cont) | Critical value=0.5273 | B1 | 1.3 | Condone cv= -0.5273 SC Condone cv=0.521 when PMCC used |
| | (0.704>0.5273 so significant evidence to) reject H ₀ | A1 | 2.1b | Dep on consistent cv and ts SC Condone PMCC |
| | There is (significant) evidence that there is a positive correlation between weight of model of smartphone and battery energy charge | E1dep | 2.1a | Dependent on a comparison Or evidence to support Derek's observation SC Do not accept PMCC (max B0 M1 M0 M0 A0 B1 A1 E0 in this case) |
| 7(e) | Conclusion may be invalid as some phones were more likely to be selected than others (so not a true reflection of population) | E1 | 3.1a | Do not award just 'Biased' - Must be in context |
| 7(f) | Amina may be wrong because evidence of positive correlation does not imply causality Or Though correlation does not imply causality, it would seem reasonable that larger screens weigh more than smaller screens so her comments may be reasonable. | E1 | 2.1b | |
| Total | | 17 | | |