Mark Scheme (Results)

Summer 2019

Pearson Edexcel GCSE
In Statistics (1ST0)
Paper 2H

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

General marking guidance

These notes offer general guidance, but the specific notes for examiners appertaining to individual questions take precedence.

1 All candidates must receive the same treatment. Examiners must mark the last candidate in exactly the same way as they mark the first.

Where some judgement is required, mark schemes will provide the principles by which marks will be awarded; exemplification/indicative content will not be exhaustive. When examiners are in doubt regarding the application of the mark scheme to a candidate's response, the response should be sent to review.

2 All the marks on the mark scheme are designed to be awarded; mark schemes should be applied positively. Examiners should also be prepared to award zero marks if the candidate's response is not worthy of credit according to the mark scheme. If there is a wrong answer (or no answer) indicated on the answer line always check the working in the body of the script (and on any diagrams), and award any marks appropriate from the mark scheme.

Questions where working is not required: In general, the correct answer should be given full marks.
Questions that specifically require working: In general, candidates who do not show working on this type of question will get no marks – full details will be given in the mark scheme for each individual question.

3 **Crossed out work**
This should be marked **unless** the candidate has replaced it with an alternative response.

4 **Choice of method**
If there is a choice of methods shown, mark the method that leads to the answer given on the answer line.

If no answer appears on the answer line then mark both methods **as far as they are identical** and award these marks.

5 **Incorrect method**
If it is clear from the working that the "correct" answer has been obtained from incorrect working, award 0 marks.

6 **Follow through marks**
Follow through marks which involve a single stage calculation can be awarded without working as you can check the answer, but if ambiguous do not award.
Follow through marks which involve more than one stage of calculation can only be awarded on sight of the relevant working, even if it appears obvious that there is only one way you could get the answer given.

7   **Ignoring subsequent work**
It is appropriate to ignore subsequent work when the additional work does not change the answer in a way that is inappropriate for the question or its context. (eg an incorrectly cancelled fraction when the unsimplified fraction would gain full marks).
It is not appropriate to ignore subsequent work when the additional work essentially makes the answer incorrect (eg incorrect algebraic simplification).

8   **Probability**
Probability answers must be given as a fraction, percentage or decimal. If a candidate gives a decimal equivalent to a probability, this should be written to at least 2 decimal places (unless tenths).
Incorrect notation should lose the accuracy marks, but be awarded any implied method marks.
If a probability fraction is given then cancelled incorrectly, ignore the incorrectly cancelled answer.

9   **Range of answers**
Unless otherwise stated, when an answer is given as a range (eg 3.5 – 4.2) then this is inclusive of the end points (eg 3.5, 4.2) and all numbers within the range.

Guidance on the use of abbreviations within this mark scheme

M   method mark awarded for a correct method or partial method

A   accuracy mark (awarded after a correct method; if no method or process is seen then full marks for the question are implied but see individual mark schemes for more details)

B   unconditional accuracy mark (no method needed)

oe   or equivalent

cao   correct answer only

ft   follow through (when appropriate as per mark scheme)

sc   special case

dep   dependent (on a previous mark)

indep independent

awrt answer which rounds to

isw   ignore subsequent working

Question number		Answer	Additional guidance	Mark
1	(a)	B1B1 Any two from: • Repeated random numbers • Random numbers out of range/may not correspond to students' numbers • Selected students may not (want to) participate • Some students may have left the university	B1 for each bullet point up to a maximum of 2 Accept each bullet point only once Students may have joined the university is B0. There may not be 100 students at the university is B0. Random numbers may not be whole numbers is B0. Database may not be up to date **on its own** is B0. Do not accept (random) numbers may be more than 100 for the second bullet point. Ignore extraneous non-contradictory comments.	(2)
	(b)	B1B1 Any two advantages from: • Easy/convenient/quick/efficient/cheap • Represents population (proportions) • Allows for comparison (between undergraduates and postgraduates) • No sample frame required	B1 for each bullet point up to a maximum of 2 Accept each bullet point only once For 2nd bullet point allow e.g. 'fair number of each (group)' 'Unbiased' on its own is B0. Ignore extraneous non-contradictory comments.	(2)
	(c)	B1 Any one from: • Not every student has an **equal**(o.e.) chance of being selected • Only those in the main building can be selected/not every student has a chance of being selected • Robert is **choosing** the students	B1 for a reason which states or implies 'equal likelihood' of being selected **or** that Robert is doing the choosing Do not allow 'even' chance or 'its biased' for the first bullet point, but condone 'fair chance'.	(1)

Question number	Answer	Additional guidance	Mark
2 (a)	B1 9 (thousand)	B1 Accept 9000	(1)

(b)	B1 e.g.
	   • age group is narrower/doesn't start at 10
	   • there are fewer drivers (under 20) o.e.
		B1 for equivalent wording suggesting that the class is 'smaller'. (e.g. drivers start from age 17)
		Condone sensible contextual comments, e.g. 'they have not been driving long', 'they are learning', 'they are more careful', etc.
			(1)

(c)	M1 25(000) + 21(000) + 19(000) A1 65 (thousand) B1 They have decreased / there were more claims (in 2014 by male drivers aged 20–49)/ there were **fewer** claims **in 2015** o.e.	M1 for addition of 3 correct figures from population pyramid (o.e.) A1 for 65 or 65 000 (may be implied by 1700 or 1.7) B1 for a correct conclusion (This mark is independent of M1A1)	(3)

(d)	B1 B1 for **two** correct statements:
	   • Bars get shorter as age increases, o.e.
	
	   • Bars are shorter for females, o.e.
		NOTE: Condone use of e.g. 'accidents' for 'claims'.
		1st B1 for a correct statement comparing **age**. Accept e.g. '20 – 29 make most claims'. (Condone 'older drivers make fewer claims' o.e. but 'there are fewer older drivers' o.e. is B0)
		
		2nd B1 for a correct statement comparing **gender**. (Condone 'male drivers make more claims' o.e. and reference to single age-groups, but 'there are more male drivers' o.e. is B0)
		
		If both age and gender are included in **a** single comment, award this comment for the feature that is **different**.
		
		NB: If B0 scored, then a single incomplete comment e.g. 'young males make more claims' can score B1B0 (as we don't know if they are comparing age or gender).
			(2)

Question number	Answer	Additional guidance	Mark
2 (e)	B1 Not appropriate since, e.g. • the population pyramid only shows drivers who made claims • not all drivers are included	B1 for Not appropriate/No with valid reason	(1)
(f)	B1 Cannot be used to support the statement as, e.g. • we only have data for one year/2015 • there is no data for 2019 • (distribution of claims) may be different to 2015 • future claims may differ from past claims, etc	B1 for Not supported/No with valid reason Condone, e.g. 'No, as it is extrapolation' for B1. Condone, e.g. 'No, it is (only) a prediction' for B1. Do not allow 'No, the trend may not continue' on its own.	(1)
3 (a)		Time	Frequency
---	---
$53 < t \leq 54$	6		B1 for both 53 and 54 in correct order	(1)

(b)	M1 $4 + 16 + 20$ or $49 - (6 + 3)$
	
	A1 40
		M1 for the addition **3 figures** with at least two correct from 4, 16 and 20
		   or for the subtraction of the sum of **2 figures** from 49 with at least
		   one correct from 6 and 3
		A1 for 40
			(2)

(c)	M1 $4 \times 50.5 + 16 \times 51.5 + 20 \times 52.5 + 6 \times 53.5 + 3 \times 54.5 \ (= 2560.5)$ M1 $2560.5 \div 49$ A1 52.255…	M1 for Σft with at least 3 correct products **or** $2500 < \Sigma ft < 2600$ M1 for 'Σft' $\div 49$ (must be dividing a sum of products by 49) A1 for 52.2 to 52.3 Allow 52 **from correct working seen** to score 3 out of 3	(3)

Question number	Answer	Additional guidance	Mark
4 (a)	B1B1 for each of: • order the leaves/numbers • add a key	B1 for each improvement. Ignore extraneous comments. Do not allow comments referring only to titles or labels.	(2)
(b)	B1 for reliable and any one of • comes from known/trusted website • as they won't vary from one website to another/scores can be found all over the internet • data (on scores) very likely to be true	B1 for reliable **and** supporting reason Ignore extraneous comments once B1 is scored.	(1)
(c)	B1 e.g. 'not a suitable diagram as it does not show how data changes over time'	B1 for not suitable **and** appropriate supporting reason Allow 'Not suitable' and any reason which states or implies that 'time' is not shown in the diagram, e.g. 'no years given' Allow 'Not suitable' and a diagram which would show this e.g. 'time series' (condone scatter diagram)	(1)

Question number	Answer	Additional guidance	Mark
5	B1 (Mean =) 297 M1 $\sqrt{\dfrac{631217}{7} - {'297'}^2}$ A1 44.326… B1ft e.g. The mean for E(thiopia) > mean for P(akistan) **or** The s.d. for E(thiopia) < s.d. for P(akistan) B1ft On average, more aid/money(o.e.) is given to Ethiopia **and** the amount of aid/money(o.e.) given to Pakistan varies more.	B1 for correct mean (must be seen outside of s.d. formula) M1 for attempt at standard deviation (allow f.t. on their mean) A1 for awrt 44.3 (Ignore calculations of standardised scores) B1ft for a correct comparison of means **or** standard deviations (may be implied by one correct interpreted comparison of **means** or **standard deviation**) Do not allow this mark if it clearly comes from a comparison of greatest values. B1ft for correctly **interpreted** comparison of **means <u>and</u> standard deviations**	(5)

Question number	Answer	Additional guidance	Mark
6 (a)	B2 CPI has increased 3.2% (from July 2015 to July 2017)	B2 for both increase and 3.2% (B1 for increase or 3.2%) (Ignore reference to dates)	(2)
(b)	M1 $\dfrac{525 \times 100.7 + 475 \times x}{525 + 475} = 103.2$ A1 105.96…	M1 for either $525 \times 100.7 + 475 \times x$ or $(525 + 475) \times 103.2$ or $0.525 \times 100.7 + 0.475 \times x$ seen A1 for awrt 106 from correct working (A correct answer with no working scores 2 out of 2)	(2)
(c)	B2 Not appropriate with any one from • the weightings/proportions of individual items may vary • (the index number) for food is not known • (the index number/100.7) takes into account other things than just food • (the index number/100.7) represents goods not food	B2 for Not appropriate/No with correct supporting reason (B1 for Not appropriate/No and attempt at reason)	(2)

Question number	Answer	Additional guidance	Mark
7 (a)	B1 A suitable reason: • the data is continuous • the data is grouped	B1 for either suitable reason Condone 'categories have different class widths' for B1	(1)
(b)	B1 for two bars drawn at correct width (5.5-7.5 and 7.5-11.5) M1 for use of frequency density A1 for both bars drawn at correct height (2.5 and 1)	Allow ½ small square tolerance on the widths but not the heights. M1 for $\frac{5}{7.5-5.5}$ or $\frac{4}{11.5-7.5}$ (working may be seen in table) or one correct bar height A1 must come from correct working if seen, i.e. f.d. = 2.5 and f.d. =1 f.d. = 1.3(3…) or 4/3 seen or plotted is A0. Both bars drawn at correct height with no working scores M1A1	(3)
(c)	M1 $\dfrac{7}{14+11+9}$ A1 $\dfrac{7}{34}$	M1 for attempt at conditional probability in the form $\dfrac{k}{14+11+9}$ where $0 < k \le 14$ A1 for $\dfrac{7}{34}$ or awrt 0.21	(2)
(d)	B1B1 for each assessment, e.g.: • 'No response boxes for less than 2/more than 11 hours so estimate likely not to be accurate' • 'Grouped response boxes/exact values not known/uses midpoints so estimate won't be exact' • 'Loss of accuracy due to rounding to the nearest hour' • 'Not accurate as it is only for Saturday (or weekend)'	B1 for each assessment of the **accuracy** of the average up to a maximum of 2 Only need to mention 'not accurate' or 'not exact', etc. once. Do not condone 'reliable' for accuracy here.	(2)

Question number	Answer	Additional guidance	Mark
(e)	B1B1B1 for each assessment of the method, e.g.: • 'random sampling so may be representative/unbiased' • 'only asked students at her school so only valid for students at her school' • 'only asked about one day so only valid for Saturday' • '(large) non-response so estimate may be biased' • 'response boxes not-exhaustive so not valid' • 'sample size may affect the validity of her results'	B1 for each assessment of the validity of her conclusion (that the mean is 4 for all students) up to a maximum of 3 For bullet points 2 to 6 only need to mention 'not valid'/ 'affects the validity'/ 'not reliable'/ 'not representative', 'the mean isn't 4' etc. once. (condone not accurate)	(3)

Question number	Answer	Additional guidance	Mark
8 (a)	B1 '(Appropriate) since the data is quarterly/follows a pattern every 4 quarters'	B1 for any correct supporting reason stating or implying 'quarters' or 'seasons'	(1)
(b)	B1 Point plotted at (middle of Q1 and Q2 2017, 1525)		(1)
(c)	B1 Trend line drawn from (Q3 2015, [1650 to 1750]) to (Q4 2016, [1500 to 1600])		(1)
(d)(i) (d)(ii)	B1 e.g. 'on average the 4th quarter has 140(thousand) fewer **visitors**(o.e.) compared **to the trend**' M1 '1450'(000) – 140(000) A1 1225(000) to 1375(000)	B1 for a correct interpretation in context M1 for subtracting 140 from the value read off their trend line at (Q4 2017) A1 for an answer in the range 1225 to 1375 inclusive Note: 1690 – 140 is M0A0	(1) (2)
(e)	B1 for each assumption: • '**overall** trend continues' • '**seasonal** trend continues'	B1 for each assumption required to use the method If B0 scored, allow 'trend continues' or 'continues to decrease' to score B1	(2)

Question number	Answer	Additional guidance	Mark
9 (a)	M1 $\dfrac{68-55}{8}$ or $55+8\times1.5$ A1 1.625 or 67 depB1f.t. e.g. 'Mithra will get an interview'	M1 for standardising or for attempting to find minimum score required A1 for awrt 1.6 or 67 depB1ft reasoning statistically to form correct conclusion or correct ft conclusion (dep on M1 being scored)	(3)
(b)	B2 Alexi performed worse on the test since Alexi's (standardised) score was lower (o.e.)	B2 for Alexi performed worse since $-1.25 < -1$ or with 45 and 47 seen (B1 for Alexi performed worse with **incomplete** reasoning e.g. since Alexi's score is further away from 0/mean')	(2)

Question number	Answer	Additional guidance	Mark
10	B1 e.g. 'To clean data, Noah will need to determine if there are any outliers' B1 e.g. 'Appropriate to remove **outliers (5.3 and 5.7)** as they may affect the reliability/accuracy of the results' B1 e.g. 'He will need to use $\mu \pm 3\sigma$' M1 $\frac{88}{50}(=1.76) \pm 3 \times 1.05$ A1 Outlier limit 4.91 (and -1.39 (no lower outliers)) B1 5.3 and 5.7 are outliers	B1 for any correct reason why data might need to be cleaned… e.g. to remove outliers, to remove anomalous values, to put data in the same units, deal with missing data, improve accuracy, etc. B1 for a supported reason as to why it is appropriate to remove the outliers or why 5.3 and 5.7 should be removed. B1 for stating the rule for removing outliers, $\mu + 3\sigma$ or $\mu - 3\sigma$ M1 for use of correct summary statistics in $\mu + 3\sigma$ or $\mu - 3\sigma$ A1 for outlier limit identified as awrt 4.9 (Sight of awrt 4.9 implies B1M1A1) B1 for identifying the two outliers (this mark may be scored independently of all previous marks)	(6)

Question number	Answer	Additional guidance	Mark
11 (a)	B1 Any reason from: • faster/convenient/easier • (more) accurate/reduces human error	B1 for any suitable advantage of using technology Do not allow cheap(er).	(1)
(b)(i) **(b)(ii)**	B1 Positive (correlation) B1 'As age increases, salary increases' B1 Spearman's/0.95 (since they are both positive) as it is closer to 1	B1 for positive B1 for correct interpretation of positive correlation (allow converse statements) B1 for Spearman/0.95 with correct supporting reason Allow e.g. Spearman/0.95 is larger (than 0.77) as supporting reason	(2) (1)
(c)	B1 Figure 1 B1 e.g. 'Figure 1 as pmcc < Spearman (so that means the correlation will be less linear)'	B1 for Figure 1 depB1 (dependent upon 1st B1) either pmcc/0.77 < Spearman/0.95 or for understanding that **Spearman** shows **rank** correlation and **pmcc** shows **linear** correlation)	(2)
(d)	B1 e.g. 'Not appropriate since **bivariate** data is needed for the pmcc'	B1 for not appropriate with supporting reason stating or implying that the data is not bivariate/paired Allow e.g. 'not appropriate since it will only show the relationship between age and salary'	(1)

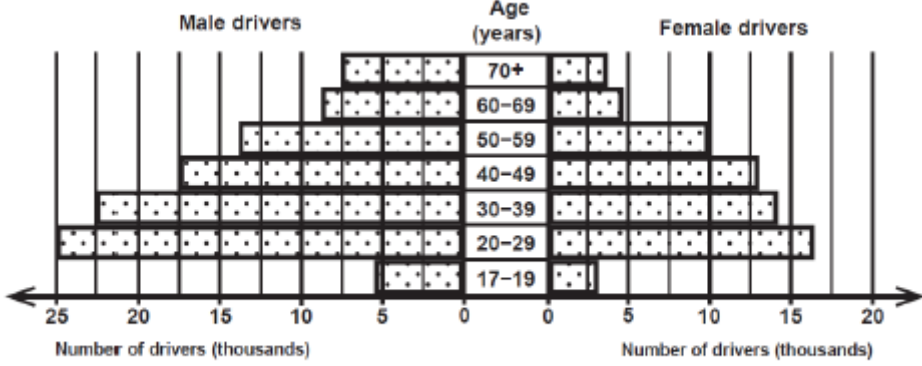| Question number	Answer	Additional guidance	Mark
12 **(a)**	B1 0.6^3 [= 0.216]	B1 for a correct calculation of the given probability Allow $^3C_0(0.4)^0(0.6)^3$	(1)
(b)	M1 $3 \times 0.4 \times 0.6^2$ A1 0.432 oe	M1 for a correct expression (allow $^3C_1$ for 3) A1 for awrt 0.43	(2)

(c)	M1 $3 \times 0.4^2 \times 0.6$ (=0.288) or 0.4^3 (=0.064)
	 A1 both awrt 0.29 and awrt 0.06
	 M1A1ft
	
	 <table><tr><td>Expected frequencies</td><td>Observed frequencies</td></tr><tr><td>$100 \times 0.216 = 21.6$</td><td>[24]</td></tr><tr><td>$100 \times 0.432 = 43.2$</td><td>[41]</td></tr><tr><td>$100 \times 0.288 = 28.8$</td><td>[30]</td></tr><tr><td>$100 \times 0.064 = 6.4$</td><td>[5]</td></tr></table>
	
	 depB1ft 'B(3, 0.4) is a suitable model/Jasper is correct'
		M1 for calculating either probability of obtaining 2 heads or 3 heads
		 A1 for both probabilities calculated (awrt 0.29 and awrt 0.06)
		 M1 for converting observed values and expected values into the same units (may be implied by at least one correct or correct ft)
		 i.e. multiplying binomial probabilities by 100
		     or for converting observed frequencies into probabilities/percentages
		 A1ft for all 4 expected frequencies correct or correct ft
		     awrt 22, '43', '29' and '6'
		     or 0.24, 0.41, 0.30, and 0.05
		
		 depB1ft (dependent upon both M marks) correct conclusion or correct ft conclusion
		
		 SC: $\dfrac{24 \times 0 + 41 \times 1 + 30 \times 2 + 5 \times 3}{300} = \dfrac{29}{75}$ (= awrt 0.39)
		 or
		
		 $\dfrac{24 \times 0 + 41 \times 1 + 30 \times 2 + 5 \times 3}{100} = 1.16$ **and** $3 \times 0.4 = 1.2$
		 scores M0A0M1A1B0
			(5)

Modifications to the mark scheme for Modified Large Print (MLP) papers: 1ST0 2H

Only mark scheme amendments are shown where the enlargement or modification of the paper requires a change in the mark scheme.

The following tolerances should be accepted on marking MLP papers, unless otherwise stated below:
Angles: ±5°
Measurements of length: ±5 mm

Question number	Modification	Mark scheme notes

2*	Diagram enlarged.
	Bar values changed as shown:
	
	(Adapted from Department for Transport)
	
	Source changed to 'Adapted from'.
		(a) B1 **10** (thousand)
		
		(c)
		M1 **25 + 22.5 + 17.5**
		      (or equivalent in thousands)
		      A1B1 as scheme

Question number		Modification	Mark scheme notes
3*		Diagram enlarged. Vertical axis scale changed to go up in 2s. Final point moved to 2. Right axis labelled. Wording '49' changed to '48'.	(c) M1 final product is **…+ 2 × 54.5** (= 2506) M1 **2506 ÷ 48** A1 **52.2**
3	(a)	Braille only: Spaces labelled (i) and (ii).	
7*		Diagram enlarged. Numbers for '4 to 5' and '6 to 7' changed to 8 and 6. Right axis labelled.	(b) f.d. for 6 to 7 = 3
8*		Diagram enlarged. 2015 removed. Wording 'Seven of the eight...' changed to 'Four of the five...'	(d)(ii) A1 1200 to 1400
11		Diagrams enlarged.	