

Mark Scheme (Results)

January 2025

Pearson Edexcel International Advanced Level In Statistics S3 (WST03) Paper 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

January 2025
Question Paper Log Number P76395A
Publications Code WST03_01_2501_MS
All the material in this publication is copyright
© Pearson Education Ltd 2025

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.

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 $\surd$ 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. 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

Special notes for marking Statistics exams (for AAs only)

- Any correct method should gain credit. If you cannot see how to apply the mark scheme but believe the method to be correct then please send to review.
- For method marks, we generally allow or condone a slip or transcription error if these are seen in an expression. We do not, however, condone or allow these errors in accuracy marks.

Question Number	Scheme	Marks
1 (a)	Rankings 2, 9, 7, 8, 6, 5, 1, 4, 3, 10	B1
	$\sum d^2 = 9+0+16+0+16+16+25+0+16+0[=98]$	M1
	$r_s = 1-\dfrac{6\times'98'}{10\left(10^2-1\right)} = 0.4060...$ awrt 0.406	M1 A1
		(4)
(b)	$H_0 : \rho = 0 \quad H_1 : \rho > 0$	B1
	Critical Value $r_s = 0.7455$ or CR: r_s …0.7455	B1
	Not in the critical region/not significant/Do not reject H_0	M1
	There is insufficient evidence of a **positive correlation** between the final **position** of a football team in the English Premier League and their average match day **attendance**.	A1ft
		(4)

	Notes	**Total 8**
(a)	**B1**	For all 8 correct missing rankings. If in the table and in the working space and different then award the highest scoring response.
	M1	For an attempt at $\sum d^2$ (at least 5 correct values seen, with an attempt to add) May be implied by 98
	M1	For using $1-\dfrac{6\sum d^2}{10(99)}$ with their $\sum d^2$ (you will need to check their $\sum d^2$ if no value shown)
	A1	awrt 0.406 Allow $\frac{67}{165}$ **NB** awrt 0.406 or $\frac{67}{165}$ scores 4/4
(b)	**B1**	For both hypotheses correct. Must be in terms of ρ or ρ_s (Condone *p)*. Must be attached to H_0 and H_1
	B1	For CV of 0.7455
	M1	A correct statement ft part (a) and their CV– no context needed but do not allow contradicting non contextual statements. This may be implied by a correct contextual conclusion on its own.
	A1ft	Correct conclusion in context. Must mention words in bold oe, ft their r in part (a) and their critical value.

Question Number	Scheme	Marks
2 (a)	$\frac{[0\times5]+1\times38+2\times32+3\times17+4\times7+5\times1}{100}\left[=1.86\right]$*	B1*
		(1)
(b)	[r = 1.203] because total **expected** frequency must equal 100	B1
		(1)
(c)	[The manager needed to do this] to ensure that [all] **expected** frequencies were greater than 5	B1
		(1)
(d)	H_0 : Poisson (distribution) is [a] suitable/ sensible (model) H_1 : Poisson (distribution) is not [a] suitable/ sensible (model)	B1
	$v = [5-1-1] = 3$	B1
	$c^2_3(0.01) = 11.345 \Rightarrow$ CR: $X^2 \ldots 11.345$	M1
	[Lies in the CR/Reject H_0] Sufficient evidence to say that **Poisson** is not a suitable model	A1ft
		(4)

	Notes	**Total 7**
(a)	**B1***	For a correct method to show the mean is 1.86 (Ignore use of 6 × 0) Allow $\frac{[0]+38+64+51+28+5}{100}$
(b)	**B1**	A correct explanation referring to the fact that total/sum **expected** frequency/ E_i must equal total observed frequency e.g. 100 – (15.567 + 28.955 +26.928 + 16.696 + 7.763 + 2.888) = r
(c)	**B1**	A correct explanation referring to the fact that [all] E_i /**expected** frequencies/values need to be greater than 5 e.g because **expected** 5 customers and [**expected]** 6 or more customers are both less than 5 Allow 2.88 < 5 and 1.203/r < 5 or 4.091 < 5
(d)	**B1**	Both hypotheses correct. Must mention Poisson/Po at least once.
	B1	v = 3 This mark can be implied by a correct critical value of 11.345 if no DoF given
	M1	For 11.345 or ft their degrees of freedom $\left[c^2_4(0.01) = 13.277\right]$
	A1ft	A correct conclusion based on their χ^2 critical value. Must mention Poisson

Question Number	Scheme	Marks
3 (a)	$\left[p = \frac{118}{40} = \right] 2.95$	B1
	$[q =] \frac{350.05 - 40('2.95')^2}{39} = 0.05$	M1 A1
		(3)
(b)	$H_0 : \mu_A = \mu_B \quad H_1 : \mu_A < \mu_B$	B1
	$z = \pm \frac{2.65 - '2.95'}{\sqrt{\frac{0.07}{50} + \frac{'0.05'}{40}}}$	M1 M1
	$= 5.827...$ or $= -5.827...$ awrt ± 5.83	A1
	CV = 1.6449	B1
	Reject H_0 There is significant evidence to support the biologist's **belief**	M1 A1ft
		(7)
(c)	Large sample sizes so …	
	both sample means are normally distributed (CLT)	B1
	$s_A^2 = \sigma_A^2$ **and** $s_B^2 = \sigma_B^2$	B1
		(2)

	Notes	**Total 12**
(a)	**B1**	2.95 only
	M1	For use of $\frac{\sum x^2 - n\bar{x}^2}{n-1}$ oe ft their $\bar{x}$ May be implied 0.05 provided no incorrect working seen
	A1	cao
(b)	**B1**	Both hypotheses correct. Allow equivalent hypotheses. Must be in terms of μ
	M1	For the denominator. Ft their 0.05
	M1	Fully correct. Ft their 2.95 and their 0.05
	A1	awrt 5.83 allow $\lvert z \rvert = 5.827...$ accept $p = 2.8(1) \times 10^{-9}$
	B1	|CV| =1.6449 or better
	M1	A correct conclusion not in context ft their z value and CV or a correct p value (2 sf)
	A1ft	ft their z value and their CV (NB their CV must be consistent with their z value) or a correct p value (2 sf). Correct conclusion in context, need **belief/claim**. May be in words with **weights** and **region** e.g. the **weights** in **region *A*** are smaller
(c)	**B1**	Must comment on **both** sample means e.g. the sample means are normally distributed
	B1	Must comment on **both** variances/standard deviations e.g. sample variances can be used as values for the population variances

Question Number	Scheme	Marks
4 (a)	$2\times\text{awrt } 2.5758\times\text{SE}=0.964-0.9$ or $\text{awrt } 2.5758\times x=0.032$	M1 B1
	$\Rightarrow\dfrac{0.964-0.9}{2\times\text{awrt } 2.5758}\left[=0.0124\right]*$ or $x=\dfrac{0.032}{\text{awrt } 2.5758}\left[=0.0124\right]*$	A1*
		(3)
(b)	$\left[\overline{x}=\right]\dfrac{0.964+0.9}{2}\left[=0.932\right]$ or $\left[\overline{x}=\right]0.964-'2.5758'\times 0.0124[=\text{awrt}0.932]$ or $\left[\overline{x}=\right]0.9+'2.5758'\times 0.0124[=\text{awrt}0.932]$	M1
	$'0.932'\pm 1.96\times 0.0124$	M1 B1
	$(0.9076...,\ 0.9563...)$ awrt $(0.908,\ 0.956)$	A1
		(4)
(c)	$2\times z\times 0.0124=0.04$	M1
	$z=1.612...$ awrt 1.61	A1
	$\text{P}(Z>'1.61')=\text{P}(Z<-'1.61')=1-'0.9463'$	M1
	$=0.0537$ (Calculator gives 0.05371….) awrt 0.0537	
	Confidence level $=\left[100\times\right](1-2\times'0.0537')$ or $\left[100\times\right]('0.9463'\times 2-1)$	M1
	$=89.26$ awrt 89.3	A1
		(5)

		Notes	**Total 12**
(a)	**M1**	For $2\times z$ value $\times\text{SE}=0.964-0.9$ oe or z value × x = 0.032 oe where $2<z<3$ May be implied by $\dfrac{0.964-0.9}{2\times\text{awrt } 2.5758}$ or $\dfrac{0.032}{\text{awrt } 2.5758}$	
	B1	awrt 2.5758	
	A1*	Answer is given so no incorrect working must be seen. Must be at least one line of correct working between M1 and the final answer. Must use awrt 2.5758 May be implied by awrt 0.01242…	
(b)	**M1**	Accept awrt 0.932 to imply a correct method. If using a z value, then this must be awrt 2.5758 or consistent with the z value used in part (a)	
	M1	For $\overline{x}\pm z$ value $\times 0.0124$ ft their $\overline{x}$ and where $1.5<z<2$	
	B1	awrt 1.96	
	A1	for (awrt 0.908, awrt 0.956) Allow awrt $0.908<\mu<$ awrt 0.956	
(c)	**M1**	For $2\times z\times 0.0124=0.04$ oe May be implied by awrt 1.61	
	A1	For z = awrt 1.61	
	M1	For awrt 0.946 or awrt 0.947 or awrt 0.053 or awrt 0.054 **NB awrt 0.946 or or awrt 0.947 or awrt 0.053 or awrt 0.054 scores M1A1M1**	
	M1	For $\left[100\times\right](1-2\times'0.0537')$ or $\left[100\times\right]('0.9463'\times 2-1)$ ft their $\text{P}(Z>'1.61')$ (May be implied by 89.26 or awrt 89.2 or awrt 89.3 or 0.8926 or awrt 0.892 or awrt 0.893)	
	A1	For awrt 89.3 **NB** An answer of 89.2 or 89 can score M1A1M1M1A0	

Question Number	Scheme	Marks
5 (a)(i)	Quota sampling would remove the need for a sampling frame oe	B1
(ii)	Quota sampling [can be/introduce] bias	B1
		(2)
(b)(i)	$\frac{(66+40)\times 120}{200}=63.6$	M1 A1
(ii)	$(66+40)-63.6=42.4$ or $\frac{(66+40)\times 80}{200}=42.4$	A1
		(3)
(c)	H_0: Students favourite science **subject** and **place** lived are independent/not associated H_1: Students favourite science **subject** and **placed** lived are not independent/associated	B1
	Observed / Expected / $\frac{(O-E)^2}{E}$ 66 / 63.6 / $\frac{(66-63.6)^2}{63.6}\left[=\frac{24}{265}=0.09056...\right]$ 40 / '42.4' / $\frac{(40-'42.4')^2}{'42.4'}\left[=\frac{36}{265}=0.13584...\right]$	M1
	$\sum\frac{(O-E)^2}{E}=4.549+'0.09056...'+'0.13584...'$	M1
	$=4.775\ldots$ awrt 4.78	A1
	$\nu=(2-1)(3-1)=2$	B1
	$c^2_2(0.1)=4.605\Rightarrow$ CR: $X^2\ldots 4.605$	B1ft
	[in the CR/significant/Reject H_0] There is sufficient evidence to suggest that students' favourite science **subject** is not independent of the **place** they live.	dA1ft
		(7)
	Notes	**Total 12**
(a)(i)	**B1**	For a correct advantage. Possible advantages (but not an exhaustive list): includes all key subgroups, effective for small populations (Do not allow quick oe or cheap oe or easy oe)
(ii)	**B1**	For a correct disadvantage. Possible disadvantages (but not an exhaustive list): [risk of] non-random [selection], difficulty in setting quotas
(b)(i)	**M1**	For a correct method to find either expected frequency May be implied by 63.6 or 42.4
(ii)	**A1**	For either 63.6 or 42.4
	A1	For both 63.6 and 42.4
(c)	**B1**	For both hypotheses correct. Must mention subject and place at least once. Do not allow correlation to imply association. Allow dependent to imply not independent
	M1	A correct method for finding both contributions to the χ^2 value ft their 63.6 and their 42.4
	M1	Adding their two values to 4.549 (may be implied by a full χ^2 calculation, do not ISW)
	A1	awrt 4.78 **NB** This implies M1M1A1
	B1	$\nu=2$ This mark can be implied by a correct critical value of 4.605
	B1ft	4.605 or better ft their degrees of freedom [$c^2_3(0.1)=6.251$]
	dA1ft	Dependent on both M marks being awarded. A correct contextual conclusion, which has the words subject and place (Allow 'where they live' to imply 'the place they live'). Allow an answer in terms of association. Do not allow correlation to imply association. Allow dependent to imply not independent ft their $\sum\frac{(O-E)^2}{E}$ and their χ^2 critical value This mark is independent of hypotheses

Question Number	Scheme	Marks
6 (a)	$\left[\mathrm{E}\left(\bar{X}\right)=\right]\dfrac{2a+3+4a+9}{2}$	M1
	$=\dfrac{6a+12}{2}=3a+6\neq a$ * (So biased)	A1*
		(2)
(b)	$'(3a+6)'-a=2a+6$	B1ft
		(1)
(c)	$c=\dfrac{1}{'3'}$	B1ft
	$'\dfrac{1}{3}'\times'\left(3a+6\right)'+d=a$	M1
	$d=-2$	A1
		(3)
(d)	$'\dfrac{1}{3}'\times 7.32-'2'\left[=0.44\right]$ or $3a+6=7.32[\Rightarrow a=0.44]$	M1
	$4\times'0.44'+9$	M1
	$=10.76$	A1
		(3)

	Notes		Total 9
(a)	**M1**	For using the formula $\left(\dfrac{a+b}{2}\right)$ May be implied by $\dfrac{6a+12}{2}$ or $3a+6$	
	A1*	For $\dfrac{6a+12}{2}$ or $3a+6$ and $\neq a$ (Allow $3a+6-a$ or $2a+6$ and $\neq/>0$)	
(b)	**B1ft**	For $2a+6$ or ft their part (a)	
(c)	**B1**	For $\dfrac{1}{3}$ or $\dfrac{1}{\text{coefficient of } a \text{ (from part a)}}$	
	M1	For c × their (3a + 6) + d = a oe written or used May be implied by d = –2	
	A1	Cao	
(d)	**M1**	For their c × 7.32 – their d oe or 7.32 = '3a + 6'	
	M1	For 4 × their 0.44 + 9	
	A1	cao Do not ISW but condone rounding	

Question Number	Scheme	Marks
7 (a)	$W = S_1 + S_2 + S_3 + L_1 + L_2 + L_3 + L_4$	
	$W \sim \mathrm{N}\left(3\times 7.7 + 4\times 20,\ 3\times 0.01^2 + 4\times 0.02^2\right)$ So $W \sim \mathrm{N}(103.1,\ 0.0019)$	M1 A1
	$\left[\mathrm{P}(W > 103.15) =\right] \mathrm{P}\left(Z > \dfrac{103.15 - '103.1'}{'\sqrt{0.0019}'}\right) \left[= \mathrm{P}(Z > 1.1470...)\right]$	M1
	$[1 - 0.8749] = 0.1251$ (Calculator gives 0.12567…) awrt 0.13	A1
		(4)
(b)	Let $Y = L_1 - L_2$	
	$Y \sim \mathrm{N}\left(0,\ 2\times 0.02^2\right)$ So $Y \sim \mathrm{N}(0,\ 0.0008)$	M1 A1
	$\mathrm{P}\left(Z > \dfrac{0.01 - '0'}{'\sqrt{0.0008}'}\right)$ or $\mathrm{P}\left(Z < \dfrac{-0.01 - '0'}{'\sqrt{0.0008}'}\right)$	M1
	$2\times(1 - 0.6368) = 0.7264$ (Calculator gives 2 × 0.36183…) awrt 0.724 – 0.726	M1 A1
		(5)
(c)	$T \sim \mathrm{N}\left(\mu, \sigma^2\right)$	
	$\mu = 7.7n - 7.7n \left[= 0\right]$	M1
	$\sigma^2 = 0.0001n^2 + 0.0001n$	M1
	$\dfrac{2 - '0'}{\sqrt{'0.0001n^2 + 0.0001n'}} = 1.99$	M1 B1
	$0.0001n^2 + 0.0001n - 1.01[00755...] = 0$	dM1
	$n = 100$	A1
		(6)

	Notes	**Total 15**
(a)	**M1**	For setting up a normal distribution with a mean 103.1
	A1	For a correct expression for variance (0.0019) or standard deviation (0.04358…) Implied by a correct variance or a correct standard deviation
	M1	For standardising using 103.15, their mean and their standard deviation If their mean and/or their standard deviation/variance are incorrect then working must be shown
	A1	awrt 0.13
(b)	**M1**	For $L_1 - L_2$ May be implied by a correct mean or variance
	A1	For $\mathrm{N}(0,\ 0.0008)$
	M1	For standardising using 0.01, their mean and their standard deviation (May be implied by awrt 0.6368 or awrt 0.3632 or awrt 0.3618 or awrt 0.6382)
	M1	For 2 times *p* where 2*p* is a probability (Calculator gives 2 × 0.36183…)
	A1	For answers in the range awrt 0.724 – awrt 0.726
(c)	**M1**	For a correct expression for μ Implied by a mean of 0
	M1	For a correct expression for σ^2
	M1	For standardising using 2, their mean and their standard deviation and set = to a *z* value where $1.95 < \lvert z \rvert < 2$
	B1	awrt 1.99 seen or used
	dM1	Dependent on 2nd M1. For rearranging to get a correct 3 term quadratic e.g. $n^2 + n - 10101$ or $n^2 + n - 10102$
	A1	cao (Must reject –101 if found)

Pearson Education Limited. Registered company number 872828
with its registered office at 80 Strand, London, WC2R 0RL, United Kingdom