

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 gualifications 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 touch with the details on contact get in us using our 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: <u>www.pearson.com/uk</u>

Summer 2015 Publications Code UA042717 All the material in this publication is copyright © Pearson Education Ltd 2015 • 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 $\sqrt{}$ 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)
- d... or dep dependent
- indep independent
- dp decimal places
- sf significant figures
- * The answer is printed on the paper or ag- answer given
- _ 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

Ques	tion bor	Scheme									Mark	s			
1.	(a)	Label all the books from $1 - 160$ (o.e.)									B1				
	()	Use random numbers to select the 10 books									B1	(2)			
	(b)	Book A B C D E F G H I J											(2)		
		Borrow rank	1	2	3	4	5	6	7	8	9	10			
		Page rank	1	6	4	2	8	3	10	7	5	9	-	M1	
		$\begin{array}{c ccccccccccccccccccccccccccccccccccc$									M1				
		$r_s = 1 - \frac{6 \times 66}{10(100 - 1)}, [= 1 - 0.4] = 0.6$ <u>0.6</u>									M1,A1				
															(4)
	(c)	H ₀ : $\rho = 0$ H	$[1: \rho >$	> 0										B1	
		Critical value	is 0.56	536										B1	
		0.6 > cv so sig	0.6 > cv so significant result and sufficient evidence to reject H ₀									DIG			
		There is support for the librarian's belief									Blft				
		or there is evidence of a correlation between the number of pages in a book and the number of times it is borrowed											(3)		
		and the function of three it is contowed.											Total 9)	
						No	otes								
	(a)	 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 													
	(h)	1 st M1 for a	1^{st} M1 for an attempt to rank the number of pages (at least 4 correct) Allow reverse ranks								anks				
	()	2^{nd} M1 for attempt at d^2 row (may be implied by sight of $\sum d^2 = 66$ or 264 for reverse ra								nks)					
		3^{rd} 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.	A1 for 0.6 (or -0.6 for reverse ranks)												
	(c)	1 st B1 for bo	oth hyp	othese	s in ter	ms of	ρ , one	tail H	ı (com	patible	with r	anks)	Allo	ow use of	ρ_s
		. Hyj	pothes	es just	in wo	ords e.	g. "no	correl	lation'	' score	B0.				
		2 nd B1 for	cv of	0.5636	5 [If they	have have	a two	tail I	\mathbf{H}_1 then	allow	0.64	85]		
		All	ow <u>+</u> f	for rev	erse ra	anking	but n	nust be	e same	sign a	as r_s				
		If h	ypoth	eses ai	e the	wrong	way a	around	l this r	nust b	e B0 t	out 3 rd	B 1	is possibl	le.
		3 rd B1ft for	a corr	ect co	ntextu	alised	comn	ient. I	Must r	nentio	n "lib	rarian'	' (01	he)	
		<u>or</u> Fo	nume 11 nume	ber of j	pages	and r and	DOITO d their	wing	rovide	d it is		1)			
		Do	n't ind	nougi	the w	r_s and r_s		ev (pi	'negat	un 15	$ \mathbf{c}\mathbf{v} < \mathbf{c}$	r) na taile	nd to	et	
		Use of "association" is B0													
		Independent of 1 st 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													

Question Number	Scheme									
2. (a)	H ₀ : $\mu_g - \mu_s = 1.5$ [$g = \text{ in a group, } s = \text{ on their own }$]									
	$H_1: \mu_g - \mu_s > 1.5$									
	s.e. = $\sqrt{\frac{2.1^2}{80} + \frac{1.4^2}{65}} = \left[\sqrt{0.08527}\right] = [0.292]$									
	$z = \frac{8.7 - 6.6 - 1.5}{"\sqrt{\frac{2.1^2}{80} + \frac{1.4^2}{65}}"}$									
	= 2.0546 awrt $2.05(5)$	A1								
	cv 1% one tailed = 2.3263									
	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	A1ft								
	or researcher's belief not supported	(8)								
(b)) Since sample is large Central Limit Theorem (CLT) applies									
	No need to <u>assume</u> normal distribution									
		(2) Total 10								
	Notes									
	1 st & 2 nd B1 for hypotheses. Accept μ_1, μ_2 or μ_A, μ_B etc if there is some indication of									
(a)	which is which e.g. $G \sim N(\mu_g, 8.7)$									
	1 st M1 for an attempt at se with 3 out of 4 values correct. Condone switching 2.1 and 1.4									
	$\sqrt{\frac{2.1^2 \text{ or } 1.4^2}{90} + \frac{1.4^2 \text{ or } 2.1^2}{65}}$									
	$\gamma = \delta 0$ = 0.5 $2^{nd} dM1$ dependent on $1^{st} M1$ for a correct numerator(must have -1.5) and ft th	eir se								
	1^{st} A1 for awrt 2.05									
	3^{rd} B1 for ± 2.3263 or better seen or probability of awrt 0.02									
	2^{nd} Alft dep. on 1 st M1 for a correct statement based on their normal cv and their te 2^{nd} Alft for correct comment in context. Must mention "plan" and "group or i and "1.5" or "researcher" and "belief or claim"	st statistic individual"								
	NB Use of cv for difference in means <i>D</i> will have $D = 1.5 + 2.3263 \times s.e. = awrt 2.18$ and requires sight of $d = 2.1$ with a comment for the 3 rd M1									
(b)	$ \begin{array}{l} 1^{st} B1 & \text{for mentioning "large samples" and "CLT"} \\ 2^{nd} dB1 & \text{dependent on } 1^{st} B1 \text{ for stating no need to assume normality (since CLT assures it) } \end{array} $									

Questic Numbe	n Scheme	Marks						
3. () Label staff (from $1 - 16$) and children (from $1 - 40$) Use rendem numbers to select	B1 P1						
	4 staff and 10 children	B1 B1						
0) $\overline{x} = \hat{\mu} = 31.2142$ awrt <u>31.2</u>	(3) B1						
	$s^2 = \frac{20983 - 14 \times 31.2}{13}$	Alft						
	= 1026.33 awrt <u>1030</u>	A1						
		(4)						
() $\frac{\sqrt[n]{1026.33}}{\sqrt{14}}$, = 8.562 awrt <u>8.56</u>	M1, A1						
() The variation within each stratum is quite small (o.e.)	(2) B1						
	The difference in the means will be quite large , (so variations from the	B1						
	overall mean will be large giving a larger overall s.e.)	(2)						
	Notos	Total 11						
(;	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,							
đ	B1 for awrt 31.2 M1 for a correct expression ft their \overline{x} and allow transcription error in $\sum x^2$ e.g. 29683 1 st A1ft for a fully correct expression ft their \overline{x} only 2 nd A1 for awrt 1030							
() M1 for attempting $\frac{\text{"their s"}}{\sqrt{14}}$ (must have 14) A1 for awrt 8.56							
((1st B1 for a suitable comment about variation (se) suggesting that variation (se) within strata is less than that overall 2nd 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							
4. (a)	$H_0: \mu = 0.5$ $H_1: \mu \neq 0.5$							
	(Significance level = $)10\%$							
	(0.5 is in the interval so not significant, accept H ₀ , can accept) $\mu = 0.5$							
(b)	$1.6449 \times \frac{o}{100} = 0.0247$	MI B1						
	$\sqrt{100}$ 10×0.0247	DI						
	$\sigma = 0.15016 \text{ or } \frac{10 \times 0.0247}{1.6449}$ (awrt 0.15)	A1						
	$0.479 + 1.96 \times \sigma''$	M1						
	$0.479 \pm 1.90 \land \frac{1}{\sqrt{150}}$	B1						
	awrt (0.455, 0.503)	A1						
		(6) Total 9						
	Notes	10001 2						
(a)	1 st B1 for both hypotheses in terms of μ .							
	2^{nd} dB1 for 10% but accept 5% if they have a one-tail test as H ₁							
	3 rd B1 for a correct comment leading to accepting H ₀							
	Ignore any 'further calculations'.							
(b)	1 st M1 for $z \frac{0}{\sqrt{100}} = k$, using $n = 100$ and where $ z > 1.5$ and $0.02 < k < 0.03$							
	1 st B1 for 1.6449 or better in an attempt (could be 1.6449 $\sigma = k$ or even 1.6449 $\sigma^2 = k$)							
	1 st A1 for a correct expression for σ e.g. awrt 0.15							
	2 nd M1 for $\overline{x} \pm z \times \frac{\sigma}{\sqrt{150}}$ for any z (>1) and ft their σ and allow $\overline{x} \in (0.4633, 0.5127)$							
	Allow use of letter σ without a value.							
	2^{nd} B1 for 1.96 or better in an attempt (could be 1.96 σ or even 1.96 σ^2)							
	2^{na} A1 for awrt 0.455 and awrt 0.503							

Question Number	Scheme	Marks						
5 (i)	Let $R = B_1 + B_2 + B_3 + B_4 + B_5 - 5H$ so $E(R) = -25$ (o.e.)	B1						
	$Var(R) = 5 \times 6^2 + 5^2 \times 4^2$ $R \sim N(-25, \sqrt{580}^2)$	M1A1						
	$P(R > 0) = P(Z > \frac{0 - 25}{\sqrt{580}}) = P(Z > 1.04), = 0.149619(calc) or 0.1492 (tables)$	dM1 A1						
		(5)						
(ii)(a)	$\mathbf{a} \mid \overline{X} \sim \mathbf{N}\left(\mu, \frac{\sigma^2}{5}\right)$							
	$\operatorname{Var}(D) = \sigma^2 + \left\ \frac{\sigma^2}{5} \right\ = \frac{6\sigma^2}{5}, \text{so} \qquad D \sim \operatorname{N}\left(0, \frac{6\sigma^2}{5}\right)$	M1, A1 (3)						
(b)	$P(Y_1 > \overline{X} + \sigma) = P(D > \sigma) = P\left(Z > \frac{\sigma}{\sqrt{\frac{6}{5}\sigma}}\right)$	M1						
	= P(Z > 0.912) = 0.181(3 dp) (*)	A1cso (2)						
(c)	Since U_1 and \overline{U} are not independent (so variance formula cannot be used)	B1						
	Can be implied e.g. U_1 used to calculate \overline{U} , U_1 and \overline{U} from same sample o.e.	(1)						
(d)	Let $F = U_1 - \overline{U} = U_1 - \frac{(U_1 + U_2 + U_3 + U_4 + U_5)}{5}, = \frac{4U_1 - (U_2 + U_3 + U_4 + U_5)}{5}$	M1, A1						
	Var(F) = $\frac{4^2 \sigma^2 + 4\sigma^2}{5^2} = 0.8 \sigma^2$, so $F \sim N(0, 0.8 \sigma^2)$	dM1, A1						
	$P(F > \sigma) = P\left(Z > \frac{\sigma}{\sigma\sqrt{0.8}}\right) = P(Z > 1.118)$	M1						
	= 0.1314 (tables) or 0.131776 (calc) awrt 0.131~0.132	Alcso						
		(6) Total 17						
	Notes	1000017						
(i)	1 st B1 for $E(R) = -25$ (or 25 if their R is defined the other way around) 1 st M1 for an attempt at $Var(R) = 5Var(R) + 25Var(U)$. Condens assuming of C^2	$a = 1 + 1^2$						
	1^{st} M1 for an attempt at Var(R) = 5Var(B) + 25Var(H). Condone swapping of 6^2 and 4^2 1 st A1 for normal and correct variance (ft their mean)							
	$2^{nd} dM1$ for attempting the correct probability and standardising with their m	nean and sd.						
	It is mark is dependent on 1° M1 so if R is not being used or M0 for variat If their method is not crystal clear then they must be attempting $P(Z > +$	ve value) o.e						
	2^{nd} A1 for answer in the range [0.149, 0.150]	,						
(ii)(a)	B1 for correct distribution of \overline{X} (may be implied for a correct answer for <i>D</i>) M1 for correct attempt at Var(<i>D</i>) (ft their Var(\overline{X})) [A1 needs must be fully correct]							
(ii)(b)	M1 for expressing probability in terms of D and standardising A1cso for seeing P($Z > 0.912$) or prob of 1 – 0.8186 (tables) or 0.180655(6)	calc)						
(c)	B1 correct statement that should mention U_1 and \overline{U}							
(d)	1 st M1 for forming an expression in terms of U_1U_5 only							
	1 st A1 for collecting U_1 terms and getting in a form where $Var(aX \pm bY)$ can	be used.						
	2^{na} dM1 for a correct expression for Var(their F). Dependent on 1^{st} M1. 2^{nd} A1 for a correct distribution for F							
	3^{rd} M1 attempting a correct prob and standardising using their Var(<i>F</i>), σ must cancel							
	3 rd A1cso for awrt 0.131 or 0.132							

Question		Scheme									
6. (a)	$H_0: U[0,$	B1									
	D	O_i	E_i	$\frac{(O_i - E_i)^2}{E}$	$\frac{O_i^2}{F}$	Values of D	B1				
	0-4	22	40	$\frac{E_i}{8.1}$	$\frac{L_i}{12.1}$	Expected Freq	M1A1				
	4 – 7	39	30	2.7	50.7	4^{th} or 5^{th} col	M1				
	7 - 9	25	20	1.25	31.25	$\chi^2 = 13.65$	A1				
	y = 10	$\frac{14}{2(1\%) - 11}$	345	1.0	19.6		B1 B1				
	$\begin{bmatrix} v - 3, & \chi_3 \\ \text{[Reject H]} \end{bmatrix}$	$v = 5$, $\chi_3(1\%) = 11.545$ [Paieet H] the uniform distribution over [0, 10] is not a suitable model									
		[10] (10 , 10									
(b)	Area $\propto \pi R^2$	² so $r = 81$, – 49 = <u>32</u>				M1, A1				
				s = 100 - "3	32" – 49 <u>or</u> 1	100 - 81 = 19	B1ft	(2)			
(c)	Not signific	ant. Henry	's model is s	uitable			M1. A1	(3)			
			5 1110 001 10 5				,	(2)			
(d)	H_0 : The col	our/region c	hosen for the	points is inde	ependent of g	gender(or no assoc')	B1				
	\mathbf{H}_1 : The cold	our/region cl	nosen for the	points is depe	endent on gei	nder(or assoc')	D 1				
	30×65							(1)			
(e)	$\frac{39\times03}{100}$						B1				
	100							(1)			
(f)	(f) Expected frequency for Yellow and Boys is $4.9 < 5$ so col. must be										
	pooled/combined. [This gives a 2×3 table so $v = (2-1) \times (3-1) = 21$										
		$[2 m s g (s s u 2 n s u s s s (s - (2 - 1) \land (s - 1) - 2)]$									
(g)	cv = 4.605	B1									
		Ы	(2)								
	2 nd B1 for	Notes									
(a)	1^{st}M1 for	at least 2 e	xpected freq	uencies or c	lear use of a	a correct formula e.g	g. 0.4N				
	1 st A1 for	all the corr	ect E_i		, th ≂ th						
	2^{nd} M1 for 2^{nd} A1 for	at least 2 c	orrect calcu	lations from	4^{m} or 5^{m} c to 3 sf)	column					
	Awr	t 13.7 only	scores 2 nd E	81M1A1M1	A1						
	3 rd A1 for	a correct co	onclusion rej	ecting the u	niform mod	el. Award provided	their test				
(b)	M1 for sor	$suc > 11.5^2$ ne attempt	to use πR^2 t	to find r							
		r									
(c)	M1 for a c	orrect state	ment that it	is not signif	icant						
	B1 Indepen	B1 Independence or association mentioned at least once if ditto marks used.									
(d)	Allow conn	Allow connection but not correlation.									
(f)	BI for rec $2^{nd}B1$ for c	B1 for recognising there is an $Ei < 5$ and need for pooling/combining oe 2 nd B1 for correctly stating that Phoebe's belief is not supported by the data of (depends on									
(g)	their cv bein	ng > 1.411)	that I h		i is not supp		(uepenus	, 011			

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