

# Mark Scheme (Results)

Summer 2016

Pearson Edexcel GCE Statistics S3 (6691/01)



ALWAYS LEARNING

### **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 <u>www.edexcel.com</u> or <u>www.btec.co.uk</u>. Alternatively, you can get in touch with us using the details on our contact us page at <u>www.edexcel.com/contactus</u>.

#### 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: <a href="https://www.pearson.com/uk">www.pearson.com/uk</a>

Summer 2016 Publications Code 6691\_01\_1606\_MS All the material in this publication is copyright © Pearson Education Ltd 2016

#### **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.

#### **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 2016 6691 Statistics S3 Mark Scheme

Question Number	Scheme	Marks
1(a) e.g.	Analyse / find estimates for a particular <b>subgroup</b> of the population.	
	Stratified guarantees representation of all groups, srs does not.	
	Observe <b>relationships between subgroups</b> – srs does not guarantee equal or proportionate representation.	
	Rare or extreme cases as part of a <b>small subgroups</b> can be represented <b>proportionately</b> in stratified i.e. stratified represents the structure of the population– srs does not allow this.	
	Stratified typically require <b>large sample size</b> compared to srs due to lower variability within subgroups compared to entire population.	
	Any 2 distinct reasons	B1B1
		(2)
(b) e.g.	It (a stratified sample) is <b>not biased</b> as the members are chosen randomly.	
	You can estimate the sampling errors (for a stratified sample)	
	It (a stratrified sample) gives <b>more accurate estimates</b> as it is a random process.	
	A quota sample may be (interviewer / process) biased.	
	It's <b>not possible to estimate/find the sampling errors</b> for a quota sample (whereas you can for a stratified sample)	
	Any 2 distinct reasons	B1B1
		(2)
		Total 4
Notes	Award B1B1 two correct, B1B0 one correct.	
	Allow 'it' for 'stratified'.	
	Do not award marks for vague responses such as 'cheap', 'easy' 'quick' 'random' etc.	
	Mentioning 'sampling frame' alone is not sufficient for a mark.	
	Mentioning 'non-response are not recorded' alone is not sufficient for a mark.	

		Scheme				Marks	
2	$H_0$ : <b>Drug conce</b>	ntration and catching	influenza are indepen	dent / not associated			
	$H_1$ : Drug concentration and catching influenza are not independent / associated						
		A	В	С			
	Influenza	$\frac{50 \times 27}{110} = 12.272$	$\frac{50 \times 52}{110} = 23.636$	$\frac{50\times31}{110}$ = 14.090.	50		
	No influenza	$\frac{60 \times 27}{110} = 14.727$	$\frac{60 \times 52}{110} = 28.363$	$\frac{60 \times 31}{110} = 16.909.$	60	M1A1	
		27	52	31	110		
		•					
	0	E	$\frac{(O-E)^2}{E}$	$\frac{2}{E}$ $\frac{O^2}{E}$			
	12	12.272	0.0060	. 11.7333			
	<u> </u>	23.636	<u> </u>				
	15	14.727	0.0050				
	23	28.363	1.0142			M1A1	
	22	16.909	1.5327	. 28.6236			
	$\simeq$ E	$ \simeq E $	v =115.62−110 = 5	av av	vrt 5.61-5.62	A1	
	v = (3-1)(2-1) Reject H <sub>0</sub> <b>Drug concentrati</b>	1) = 2, $\chi_2^2(10\%) = 4$	4.605		vrt 5.01-5.02	A1 B1B1ft M1 A1cao (10 Total 10	
Notes	v = (3-1)(2-1) Reject H <sub>0</sub>	1) = 2, $\chi_2^2(10\%) = 4$	4.605		vrt 5.01-5.02	B1B1ft M1 A1cao (10 Total	
Notes	v = (3-1)(2-1) Reject H <sub>0</sub> <b>Drug concentrati</b> B1 hyps correct w	1) = 2, $\chi_2^2(10\%) = 4$	4.605		vrt 5.01-5.02	B1B1ft M1 A1cao (10 Total	
Notes	v = (3-1)(2-1) Reject H <sub>0</sub> <b>Drug concentrati</b> B1 hyps correct w M1 for correct exp	1) = 2, $\chi_2^2(10\%) = 4$ for and catching influe	4.605 e <b>nza</b> are not independe	ent / are associated.		B1B1ft M1 A1cao (10 Total	
Notes	v = (3-1)(2-1) Reject H <sub>0</sub> <b>Drug concentrati</b> B1 hyps correct w M1 for correct exp A1 <b>all seen</b> and co	1) = 2, $\chi_2^2(10\%) = 4$ fon and catching influe ay around pression at least once	4.605 e <b>nza</b> are not independe	ent / are associated.		B1B1ft M1 A1cao (10 Total	
Notes	v = (3-1)(2-1) Reject H <sub>0</sub> Drug concentration B1 hyps correct w M1 for correct exp A1 all seen and co M1 either method	1) = 2, $\chi_2^2(10\%) = 4$ fon and catching <b>influe</b> ray around pression at least once prrect 2dp or better. Ca	4.605 enza are not independe n be implied by test st	ent / are associated.		B1B1ft M1 A1cao (10 Total	
Notes	v = (3-1)(2-1) Reject H <sub>0</sub> Drug concentration B1 hyps correct w M1 for correct exp A1 all seen and co M1 either method	1) = 2, $\chi_2^2(10\%) = 4$ fon and catching influe ray around pression at least once prrect 2dp or better. Cat at least one correct ext values. Can be impl	4.605 enza are not independe n be implied by test st	ent / are associated.		B1B1ft M1 A1cao (10 Total	
Notes	v = (3-1)(2-1) Reject H <sub>0</sub> Drug concentration B1 hyps correct w M1 for correct exp A1 all seen and co M1 either method A1 at least 3 correct	1) = 2, $\chi_2^2(10\%) = 4$ fon and catching influe ray around pression at least once prrect 2dp or better. Cat at least one correct ext values. Can be impl	4.605 enza are not independe n be implied by test st	ent / are associated.		B1B1ft M1 A1cao (10 Total	
Notes	v = (3-1)(2-1) Reject H <sub>0</sub> <b>Drug concentrati</b> B1 hyps correct w M1 for correct exp A1 <b>all seen</b> and co M1 either method A1 at least 3 correct A1 awrt 5.61-5.62	1) = 2, $\chi_2^2(10\%) = 4$ for and catching <b>influe</b> ray around pression at least once prrect 2dp or better. Cat at least one correct pet values. Can be impl	4.605 enza are not independe n be implied by test st	ent / are associated.		B1B1ft M1 A1cao (10 Total	

Question Number	Scheme							Mar	ks	
3 (a)	Variables do no	ot have a	(joint) norm	al distributio	n					
	Relationship is The given data						А	any 1	B1	(1)
(b)	Day	Sun	Mon	Tues	Weds	Thurs	Fri	Sat		(1)
	Ice-cream	6	4	7	5	3	2	1		
	Sunglasses rank	6	5	7	2	3	4	1		
	d	0	-1	0	3	0	-2	0		
	$d^2$	0	1	0	9	0	4	0		
	$\sum d^2 = 14$								M1	
	$r_s = 1 - \frac{6 \times 14}{7(49 - 14)}$	$\frac{4}{-1} = 0.7$	5						M1A1	
(c)	$H_0: \rho = 0, H$	$\mathbf{H}_1: \boldsymbol{\rho} > 0$	)						B1	(3)
	5% cv 0.7143								<b>B</b> 1	
	Reject H <sub>0</sub>								M1	
	Evidence of <b>p</b>	ositive con	relation bet	ween sales of	<b>ice cream</b> a	nd sales of	sunglasse	s.	A1cao	
	(50)	`								(4)
(d)	(5% cv 0.6694) Accept H <sub>0</sub>	)							M1	
	Insufficient evi sunglasses.	idence of	<b>positive</b> cor	relation betw	veen sales of	ice cream a	nd sales of	f	A1cao	
(e)	Suggests relation	onship mi	ght be non-	linear.					B1	(2)
									То	(1) tal 11
Notes										
(a)	Accept 'alread	-								
(b)	Accept one van M1 attempt to		•		ied by sight	of $\sum d^2$ –	14			
	M1 for use of							d If		
	answer is not A1 0.75 cao				_		urry state	<b></b>		
(c)	1st B1for both	hypothes	es in terms o	of $\rho$ , one tail.	Allow use of	f $\rho_{\rm s}$ .				
	Only award if Hypotheses jus B1 0.7143 cao	no errors	seen in hyp	otheses in pa	rt(c) and pa	- 5				
	M1 must corre		•	• -		on their tes	t stat and	cv oe		
(4)	A1 Conclusion			-	lasses					
(d)	M1 for not reje A1 must menti	-		• •						

Question Number	Scheme	Marks	
4 (a)	$X_i$ be rv 'weight of $i^{\text{th}}$ randomly chosen egg'		
	$E(X_1 - X_2) = 0$	B1	
	$Var(X_1 - X_2) = 2 \times 5^2 = 50$	B1	
	$P( X_1 - X_2  > 2) = 2P(X_1 - X_2 > 2)$	M1	
	$=2P(Z > \frac{2}{\sqrt{50}})$	dM1	
	=2P(Z > 0.2828)		
	= 2(1-0.6103) = 0.7794 awrt 0.777-0.779	A1	
			(5)
(b)	$W = C + X_1 + X_2 + \dots + X_{12}$		
	$E(W) = 40 + 12 \times 60 = 760$	B1	
	$Var(W) = 1.5^2 + 12 \times 5^2$	M1	
	= 302.25	A1	
	Distribution is N(760, 302.25)		
			(3)
(c)	$P(W > 800) = P\left(Z > \frac{800 - 760}{\sqrt{302.25}}\right)$	M1	
	=1-P(Z < 2.3007)		
	= 0.0107 awrt 0.0107	A1	
			(2)
		Total 10	
Notes (a)	B1 for 0		
	B1 for 50		
	M1 for $ X_1 - X_2  > 2$ seen. Accept $X_1 - X_2 > 2$ provided a subsequent doubling of the probability is seen. i.e. 0.3897 x 2.		
	dM1 standardise with their 0 and their $\sqrt{50}$ dependent on previous M. A1 awrt 0.777-0.779		
(b)	B1 for 760		
	M1 requires squares		
	A1 cao		
(c)	Must be finding correct probability (ie $P(W > 800)$ or $P(Z > 2.3007)$ etc) and		
	standardise with 800 and their 760 and their $\sqrt{302.25}$ A1 awrt 0.0107 from correct working.		

Question Number	Scheme	Marks
5(a)	$H_0: \mu_e = \mu_n, H_1: \mu_e > \mu_n$	B1
	$z = \frac{26.3 - 24.8}{\sqrt{\frac{12.2}{35} + \frac{10.1}{42}}} = \frac{1.5}{\sqrt{0.58904}} = \frac{1.5}{0.76749}$	M1M1
	$\sqrt{35}$ 42 z = 1.9544 awrt 1.95	A1
	Critical value is 1.6449	B1
	Reject $H_0$ . Doctor's claim is supported.	Al
	5	(6)
(b)	Either assume $\overline{X}$ has a normal distribution (for both samples) or assume sample sizes are large enough to use CLT Assume individual <b>results</b> are independent	
	Assume $\sigma^2 = s^2$ for <b>both</b> populations or a single general population	B1 B1
		(2)
(c)	$\overline{x} = \left(\frac{35 \times 26.3 + 31.7}{36} = \frac{952.2}{36}\right) 26.45$	B1
	For $n = 35$ , $\sum x^2 = 34 \times 12.2 + 35 \times 26.3^2 (= 24623.95)$	M1
	For $n = 36$ , $s^2 = \frac{25628.84 - 36 \times 26.45^2}{35} = 12.661$ awrt 12.7	dM1A1
		(4)
		Total 12
Notes (a)	Both hyps, one tailed only oe. Accept $\mu_1, \mu_2$ or $\mu_A, \mu_B$ etc if there is some indication of which is which.	
	M1 for correct method for standard error	
	M1 for whole expression	
	A1 awrt 1.95	
	B1 1.6449 or $p = 0.974$ (>0.95)	
	A1 must mention doctor and claim or description of claim that includes ' <b>mean</b> lung capacity' and 'exercise'.	
ALT (a)	M1 for $\sqrt{\frac{12.2}{35} + \frac{10.1}{42}}$	
	M1 for 1.6449 = $\frac{c}{\sqrt{\frac{12.2}{35} + \frac{10.1}{42}}}$	
	A1 for awrt $c = 1.26$ seen	
	B1 1.5	
(c)	M1 Attempt $\sum x^2 = 34 \times 12.2 + 35 \times 26.3^2$	
	or $\sum (x - \bar{x})^2 = 34 \times 12.2 + 35(26.45 - 26.3)^2 (= 415.5875)$	
	$dM1 s^{2} = \frac{\sum x^{2} + 31.7^{2} - 36 \times 26.45^{2}}{35} \text{ or } s^{2} = \frac{415.5875 + (31.7 - 26.45)^{2}}{35}$	
	A1 awrt 12.7	

Question Number			Scheme					Mark	S	
6(a)	$H_0$ : Binomial with $p = 0.3$ is a good fit. $H_1$ : Binomial with $p = 0.3$ is not a good fit.							B1		
		0	)	1		2 or mo	ore			
	Observed	6		25		19				
	Expected	50x0.		50x0.4116	50x0	).2646+50x0.07				
	(0.5)	=12.005 o 12.		=20.58		=13.23+3.78 =17.415 or <b>17.</b> 4		M1A1		
	$\frac{(O-E)^2}{E}$		3.003751	0.949291			0.144256			
	$\frac{O^2}{E}$		2.998751	30.36929			20.72926			
	$\sum \frac{(O-E)^2}{E}$	= 4.097 or	$\sum \frac{O^2}{E} - N =$	= 54.0975	0 = 4.	.097	awrt 4.09-4.1(0)	dM1A1		
	v = 3 - 1 = 2		E					B1ft		
	$\chi_2^2(5\%) = 5.$	991 (>4.1(0	))					B1ft		
	Insufficient ev	-	0 =	ot $H_0$ )						
	Binomial with	h $p = 0.3$ is a	good fit.					A1		
(b)	$\overline{x} = \frac{40 + 62}{10}$	$\frac{+54+24}{00} = 1.$	8					B1 cao	(8)	
	r = 26.78	00						B1 cao		
	<i>s</i> =16.07							B1 cao		
(c)	II . D	1.0%	II . D	• , 1	<b>C</b> ,			B1	(3	
(0)	$H_0$ : Poisson	$\frac{15 \text{ a good fit.}}{0}$	$H_1$ : Poisson	is not a good $2$	III.	3	4 or more	DI		
	Observed	5	40	31		18	6			
	Expected	16.53	29.75	26.78		16.07	10.87			
	$(O-E)^2$	$\frac{11.53^2}{16.53} = 8.042$	$10.25^2$	$4.22^{2}$		$\frac{1.93^2}{16.07} = 0.232$	$\frac{4.87^2}{10.87} = 2.182$			
	$\frac{(O - L)}{E}$	16.53	$\frac{10.25^2}{29.75} = 3.532$	$\dots \qquad \frac{4.22^2}{26.78} = 0.6$	665	16.07	10.87			
	$O^2$	$\frac{5^2}{16.53} = 1.512$	$\frac{40^2}{29.75} = 53.782$	2 31 <sup>2</sup>		18 <sup>2</sup>	$\frac{6^2}{10.87} = 3.312$			
	$\frac{O^2}{E}$	16.53	29.75	$\frac{31^2}{26.78} = 35.$	885	$\frac{18^2}{16.07} = 20.162$	10.87			
	$\sum \frac{(O-E)^2}{E}$	=14.65-14.	66 or $\sum \frac{O^2}{E}$	-N = 114.65	-100	=14.65-14.6	6	M1A1		
	v = 5 - 1 - 1 =	= 3						B1 cao		
	$\chi_3^2(1\%) = 11$	.345 (<14.6	5)					B1ft		
		dence to rejec	t H <sub>0</sub>					A1 cao		
	Poisson is not a good fit.								(6)	
								Total 17		
Notes (a)	B1 both including $p = 0.3$ M1 with some combined columns and at least one <i>E</i> correct to 2sf									
				d values is 50.		0 281				
	dM1 either m	-	1							
	A1 awrt 4.09-									
	B1 ft their col	lumns -1								
	B1 ft their A1 cao									
(c)	B1 no parame	eters included								
<-/	M1 either me									

Question Number	Scheme	Marks	5
7(a)	1.5		
/ (u)	$19.5 \pm 1.6449 \times \frac{1.5}{\sqrt{50}}$	M1B1	
	=(19.151, 19.848) awrt 19.2, awrt 19.8	A1A1	
			(4)
(b)	CI does not contain 20 oe	M1	
	Fast Food restaurant statement is too high; they should reduce the stated value.	A1	
			(2)
(c)	$P( \bar{X} - \mu  < 0.5) = 0.9$		
	$\frac{0.5}{0.5} = 1.6449$	M1A1	
	$\frac{0.5}{\frac{2}{\sqrt{n}}} = 1.6449$	1011731	
	$n = \left(2 \times \frac{1.6449}{0.5}\right)^2 = 43.29$	dM1A1	
	Sample size required is 44	A1	
			(5)
		Total 11	
Notes			
(a)	M1 correct with their z i.e. $19.5 \pm (z \text{ value}) \times \frac{1.5}{\sqrt{50}}$		
	B1 for 1.6449		
	A1 awrt 19.2, A1 awrt 19.8(5)		
(b)	M1 Require 20 compared to their interval		
	A1 Accept statement that relates to 20 being above the interval.		
(c)	M1 $\frac{0.5}{2} = z$ value or equivalent expression		
	$\frac{1}{\sqrt{n}}$ A1 All correct		
	dM1 Attempt to solve $\frac{0.5}{2}$ = their z value		
	$\overline{\sqrt{n}}$		
	A1 awrt 43.3	1	

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