Mark Scheme (Results)

Summer 2014

Pearson Edexcel GCE in Statistics 3R (6691/01R)

Edexcel and BTEC Qualifications

Edexcel and BTEC qualifications come from Pearson, the world's leading learning company. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers. For further information, please visit our website at www.edexcel.com.

Our website subject pages hold useful resources, support material and live feeds from our subject advisors giving you access to a portal of information. If you have any subject specific questions about this specification that require the help of a subject specialist, you may find our Ask The Expert email service helpful.

www.edexcel.com/contactus

Pearson: helping people progress, everywhere

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

Summer 2014
Publications Code UA040138

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.
- 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 $\surd$ 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)
   - d… or dep – dependent
   - indep – independent
   - dp decimal places
   - sf significant figures
   - $*$ The answer is printed on the paper or ag- answer given
   - $\lfloor$ 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.

Question Number	Scheme	Marks
1.		
a)		M1
	$\sum d^2 = 50$	M1A1
	$r_s = 1 - \frac{6\sum d^2}{8(64-1)} = 1 - \frac{6 \times 50}{8 \times 63}$	M1
	$r_s = \frac{204}{504} = 0.40476 \ldots$ awrt 0.405	A1
		(5)
b)	$H_0: \rho_s = 0$ $H_1: \rho_s > 0$ (accept ρ_s or ρ)	B1
	1 tail critical value $\rho = 0.6429$	B1
	Test value is not in critical region so insufficient evidence to reject H_0	M1A1ft
	No significant evidence at 5% level to support journalist's belief	
		(4)
c)	Underlying (bivariate) **Normal** distribution	B1
		(1)
d)	Evidence does not support Normal distribution since mean< median or (negative) skew, oe	B1
		(1)
		(11 marks)

Car model	A	B	C	D	E	F	G	H
Sales rank	8	6	1	5	4	7	2	3
Fuel efficiency rank	8	1	5	6	2	7	4	3
d^2	0	25	16	1	4	0	4	0

Notes	
a)	M1 for attempting to rank at least one set of data A1 for at least one set of data ranked correctly(NB this mark comes after 2nd M1) M1 for attempting Σd^2 M1 for correct use of formula for r_s
b)	B1 for H_0 **and** H_1 correct (condone $\leq$ for H_0) 2nd B1 allow 0.7381 if their H_1:$\rho_s \neq 0$ M1 for correct statement relating their test statistic and critical value A1ft their test statistic, H_1 and critical value but must be in context.
c)	B1 require **Normal** distribution, ignore additional assumptions
d)	B1 require not Normal **and** valid reason

Question Number	Scheme	Marks
2) **(a)** **(i)** **(ii)** **(b)**	Expected value = $\frac{50\times74}{200} = 18.5$ χ^2 contribution = $\frac{(27-18.5)^2}{18.5} = 3.905405405 = 3.91$ to 3sfs H_0: users age and main mobile phone use are independent/ no association between users age and main mobile phone use H_1: users age and main mobile phone use are not independent/ some association between users age and main mobile phone use $\nu = 4$ Critical value χ^2 = 9.488 Test statistic is in critical region therefore significant evidence to reject H_0 and accept H_1. Evidence at 5% level that age and main phone use are not independent.	B1 cso B1 cso (2) B1 B1 B1ft M1 A1ft (5) (7 marks)
Notes		
(b)	3rd B1 ft on their value of ν M1 for attempt to compare test statistic and their critical value A1 ft on test statistic and critical value but must be comment in context. (A0 if hypotheses are the wrong way around)	

Question Number	Scheme	Marks
3) **(a)**	$P(S > 2C) = P(S - 2C > 0)$	B1
	$E[S - 2C] = 4.9 - 2 \times 2.5 = -0.1$	M1A1
	$Var(S - 2C) = 0.64 + 4 \times 0.16 = 1.28$	M1, M1
	$P(S - 2C > 0), = P(Z > \frac{0--0.1}{\sqrt{1.28}})$	
	$= P(Z > 0.08838\ldots)$	A1
	$=0.4641$ (tables), or $0.4648\ldots$ (calculator) accept awrt 0.464 or 0.465	(6)
(b)	Let $T = S_1 + S_2 + \ldots + S_{100}$	M1A1
	$E[T] = 100 \times 4.9 = 490$	A1
	$Var(T) = 100 \times 0.64 = 64$	M1
	$P(T < 500) = P(Z < \frac{500-490}{\sqrt{64}})$	
	$= P(Z < 1.25)$	A1
	$= 0.8944$	(5)
		(11 marks)

Notes

(a)

1st M1 for …±4Var(*C*)
2nd M1 for P (*S* – 2*C* >0)
3rd M1 ft their expectation and variance but not if Var(*S* – 2*C*) is negative. (Should lead to P(*Z* > +ve)

(b)

1st M1 for attempt to find mean or variance of total

1st A1 either correct

2nd A1 both correct
2nd M1 for standardising using 500, their mean and their sd leading to P(*Z* < +ve) o.e.

Sample mean, $\bar{x} = \frac{660+\alpha}{5} = 132 + \frac{\alpha}{5}$

Question Number	Scheme	Marks
4)	Test statistic, $z = \frac{132+\frac{\alpha}{5}-160}{\frac{6}{\sqrt{5}}}$	M1A1ft
	Critical z values is 1.6449	B1
	Therefore the test statistic is significant if $\frac{132+\frac{\alpha}{5}-160}{\frac{6}{\sqrt{5}}} > 1.6449$	M1
	Therefore $132+\frac{\alpha}{5}-160 > 1.6449\times\frac{6}{\sqrt{5}}$ $\alpha > 5\left(1.6449\times\frac{6}{\sqrt{5}}+28\right)$ $\alpha > 162.0686493\ldots$ Accept awrt 162.1	A1 (6) (6 marks)
Notes		
	1st A1 ft on their $\bar{x}$ 1st B1 given for 1.6449 seen (condone sign) 3rd M1 inequality using their test statistic, accept incorrect signs for M1	

Question Number	Scheme	Marks
5)	$S_E{}^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right) = \frac{1}{119}\left(956909 - \frac{10650^2}{120}\right)$	M1
(a)	$= \frac{11721.5}{119} = 98.5$	A1 (2)
(b)	H_0: $\mu_F = \mu_E$, H_1: $\mu_F \neq \mu_E$,	B1
	$\bar{x}_E = \frac{10650}{120} = 88.75$ and $\bar{x}_F = \frac{6510}{70} = 93$	M1
	Test statistic, $z = \frac{93-88.75-0}{\sqrt{\frac{151}{70}+\frac{98.5}{120}}} = 2.4627$ ...	M1A1
	Critical values, $z = (\pm)2.5758$	B1ft
	Test stat is not in critical region	M1
	Insufficient evidence to reject H_0 at 1% level	A1ft
	No significant evidence of a difference in mean lengths of English and French films	(7)
(c)	By CLT we can assume that the mean of a large sample has a Normal distribution	B1 (1)
(d)	On a list, label English films 1 – 724 and French films 1-473 (oe)	B1
	Use random number table/generator to select	
	$\frac{724}{724+473} \times 190 = 115$ English films and	M1A1
	$\frac{473}{1197} \times 190 = 75$ French films	(3)
		(13 marks)

Notes		
	Alternative	
(a)	$S_E{}^2 = \frac{n}{n-1}\left(\frac{\sum x^2}{n} - \bar{x}^2\right) = \frac{120}{119}\left(\frac{956909}{120} - 88.75^2\right) = 98.5$	
(b)	1st B1 needs both H_0 and H_1, can be in words 2nd B1ft on their H_1 1st M1 for attempt @ both means ($\bar{x}_E$ may be in (a)) 2nd M1 for attempt at correct test statistic, ft their values 3rd M1 for attempt to compare their test stat and critical values A1 ft on their test and critical values but must include comment in context	
(c)	Require mention of mean of *E* or *F* and normal distribution	
(d)	M1 requires use of <u>random numbers</u> and attempt to find correct sample sizes A1 both 115 and 75 found.	

Question Number	Scheme	Marks
6)	Independence of each occurrence (of a fake coin)	B1
(a)	Constant probability for each occurrence (of a fake)	B1
		(2)
(b)	$r = 150 \times P(X = 2) = 150 \times \binom{20}{2} \times 0.05^2 \times 0.95^{18}$	M1
	r = 28.3015… awrt 28.3	A1
	s = 150 – (53.8 + 56.6 + 28.3 + 8.9) = 2.4	A1ft
		(3)
(c)	H_0: Bin(20, 0.05) is a suitable model H_1: Bin(20, 0.05) is not a suitable model	B1
	Combining last two groups	
	≥ 3 / **Observed frequency** 19 / **Expected frequency** 11.3	M1
	$\nu = 4 - 1 = 3$	B1
	Critical value, χ^2 (0.05) = 7.815 (accept 9.488 if their $\nu = 4$)	B1ft
	Test statistic, $\sum \frac{(O-E)^2}{E} = \frac{(43-53.8)^2}{53.8} + \frac{(62-56.6)^2}{56.6} + \cdots$	M1
	= 2.168…+ 0.515… + 0.186… + 5.246…	
	= 8.117 (accept 10.16 if groups not combined)	A1ft
	In critical region, sufficient evidence to reject H_0, accept H_1	
	Significant evidence at 5% level to reject the manager's model	A1ft
		(7)

Question Number	Scheme	Marks
(d)	$\nu = 4 - 2 = 2$	B1
	4 classes due to pooling 2 restrictions (equal total and mean/proportion)	B1 (2)
(e)	H_0: Binomial distribution is a good model H_1: Binomial distribution is not a good model	B1
	Critical value, χ^2 (0.05) = 5.991	B1
	Test statistic is not in critical region, insufficient evidence to reject H_0	B1
	Accept the assistant manager's model for the number of fake coins per bag.	(3) (17 marks)
	Notes	
(b)	M1A1 for one of *r* or *s* correct A1ft for other value if using 150 - … and answer must be >0	
(c)	1st B1 can be in words but must include p = 0.05	
	3rd B1 ft on their ν	
	Test statistic alternative method	
	Test stat = $\sum \frac{O^2}{E} - 150 = \frac{43^2}{53.8} + \frac{62^2}{56.6} + \cdots - 150 = 8.117$ ... 1st A1 ft if their groups not combined 2nd A1 ft their test and critical values but must be comment in context e.g. mention of "manager's model" <u>or</u> "fake coins"	
(d)	1st B1 evidence that pooling is required 2nd B1 must have correct reasons for restrictions.	

Question Number	Scheme	Marks
7) **(a)** **(i)**	$\bar{x} = \frac{10.01+9.97+9.93+\cdots}{8} = 9.9775$	M1
	95% CI $\bar{x} \pm 1.96 \times \frac{0.08}{\sqrt{8}}$	B1M1
	95% CI for μ (9.92, 10.03)	A1 (4)
(ii)	10.00 is within confidence interval so accept that pump may be performing correctly (although sample mean is low).	B1 (1)
(b)	Upper limit of CI is $9.96 + 1.6449 \times \frac{0.08}{\sqrt{n}} < 10.00$ $\frac{1.6449 \times 0.08}{\sqrt{n}} < 0.04$ $\sqrt{n} > \frac{1.6449 \times 0.08}{0.04}$	B1, M1A1ft M1
	$n > 10.82\ldots$ therefore minimum $n = 11$	A1 cao (5) (10 marks)

Notes		
(a) **(i)**	1st M1 attempt to find sample mean B1 for correct *z* value A1 limits correct to 2 decimal places (or more) B1 for correct z value	
(b)	1st M1A1, ft their z value	

Pearson Education Limited. Registered company number 872828
with its registered office at Edinburgh Gate, Harlow, Essex CM20 2JE