edexcel

Mark Scheme (Results)

Summer 2013

GCE Statistics 3 (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 2013
Publications Code UA037014

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.

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)
- 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.

8. In some instances, the mark distributions (e.g. M1, B1 and A1) printed on the candidate's response may differ from the final mark scheme.

Question Number	Scheme	Marks
1.	Label females 1 – 100 (or 0 – 99) and males 1 – 300 (or 0 – 299) Using *random numbers* for each group in range 1 – 100 (0 – 99) select 15 females and using 1 – 300 (or 0 – 299) select 45 males	B1 B1 B1 **[Total 3]**
	Notes	
	1st B1 for labelling\numbering\listing females and males 2nd B1 for use of random numbers or "randomly select" in each group (may be implied) 3rd B1 for selecting the correct number of females and males e.g. randomly select 45 males and 15 females scores 2nd and 3rd B marks since randomly selecting and the "each group" is implied, If using systematic sampling within each strata allow 1st B1 and 3rd B1 (if earned) but 2nd B0	

Question Number	Scheme	Marks
2.	$X \sim N(40, 3^2)$ $\quad \overline{X} \sim N(40, \frac{9}{n})$ $\qquad$ (Condone $Y \sim N(40, \frac{9}{n})$	B1
	$P(\overline{X} > 42) = P(Z > \frac{42-40}{\sqrt{\frac{9}{n}}})$	M1
	$\frac{42-40}{\sqrt{\frac{9}{n}}} \geq 1.6449$	B1 dM1
	$n \geq 6.087$ $n = 7$	A1 **[Total 5]**
	1st B1 for stating the correct distribution for $\overline{X}$. May be implied if correctly used in line 2 and no incorrect version seen elsewhere. 1st M1 for an attempt to standardise with 42, 40 and their $\sqrt{\frac{9}{n}}$, must have n. Allow $\pm$ 2nd B1 for using $z = \pm 1.6449$ (or better) 2nd dM1 Dep on 1st M1 for forming an equation in n or $\sqrt{n}$. Allow "=" or "<" i.e. setting their standardised expression = their z value ($\lvert z \rvert > 1.5$) A1 for $n = 7$ only The A1 must follow from correct working so e.g. $n < 6.087$ leading to $n = 7$ is A0	

Question Number	Scheme	Marks
3 (a)		

Town	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Pop	1	2	3	4	5	6	7	8	9	10
Empl	2	1	3	5	4	6	10	8	9	7
$\lvert d \rvert$	1	1	0	1	1	0	3	0	0	3
d^2	1	1	0	1	1	0	9	0	0	9

M1

$\sum d^2 = 22$ M1A1

$r_s = 1 - \dfrac{6 \times 22}{10 \times 99}$ dM1

$= \dfrac{143}{165} = 0.866\dot{}$ awrt **0.867** A1 (5)

(b) $H_0 : \rho = 0$; $H_1 : \rho > 0$ B1

CV = 0.6485 B1

in critical region / significant/ reject H_0 M1

evidence of positive correlation between population and no. of employees A1 (4)

(c) CV = 0.6319 B1

[not in critical region / not significant/ do not reject H_0]

No evidence of positive correlation B1 (2)

(d) No evidence to suggest that as pop' increased the no. of employees increased linearly. B1

Villages ranked highly for pop' were also ranked highly for the no. of employees. B1

(2)

ALT Alternate for part (d) if different conclusions in part (b) and part (c)

Data probably not (bivariate) normal therefore Spearman's coefficient is more suitable than the product moment correlation coefficient.

[Total 13]

Notes

(a)

1st M1 for an attempt to rank no of employees against the populations

2nd M1 for attempting $\sum d^2$ (must be using ranks) ft their ranks

1st A1 for 22

3rd dM1 dep on 1st M1 for use of the correct formula with their $\sum d^2$.

If ans. is not correct an expr' is required.

(b)

1st B1 for both hypotheses in terms of ρ, H_1 must be one tail and compatible with their ranking

M1 for a correct statement relating their r_s ($|r_s| < 1$) with their cv but cv must be such that |cv|<1

A1 for a correct contextualised comment that is rejecting H_0

Must mention "population" and "no. of employees" and "positive correlation".

Follow through their r_s and their cv (provided it is |cv| <1 Use of "association" is A0

(c)

1st B1 for 0.6319 2nd B1 does not require context just no positive correlation mentioned

(d)

1st B1 for a comment relating to pmcc

(i) no linear relationship or (ii) pmcc requires (joint) normal distribution

2nd B1 for a second comment relating to Spearman's

(i) there is a (non-linear) relationship between ranks..

or (ii) data not (joint) normal so Spearman's is better

Question Number	Scheme	Marks
4 (a)	$\frac{282\times100}{600}$ (Do not accept 282 – 114.2 – 90.2 – 30.6 (o.e.))	B1 **(1)**
(b)	9	B1 **(1)**
(c)	2.5 or better (Do not accept 0.025)	B1 **(1)**
(d)	H_0: hair colour occurs in the ratio 2:6:1:3 H_1: hair colour does not occur in the ratio 2:6:1:3	B1

	black	brown	red	blonde
observed	105	282	48	165
expected	100	300	50	150
$\frac{(O_i-E_i)^2}{E_i}$	0.25	1.08	0.08	1.5
$\frac{O_i^2}{E_i}$	110.25	265.08	46.08	181.5

B1expected
M1
A1

$\sum \frac{(O_i-E_i)^2}{E_i} = 2.91$ or $\sum \frac{O_I^2}{E_I} - 600 = 602.91 - 600 = 2.91$ (awrt **2.91**) A1

$\nu = 3$ B1

cv is 7.815 B1

[2.91< 7.815] so insufficient evidence to reject H_0 or not significant dM1

There is evidence to suggest that hair colour does occur in the given ratio. A1

(9)
[Total 12]

Notes

(d)

1st B1 for both hypotheses. Must mention hair colour and ratio
e.g. "hair colour in the given ratio" Allow use of ditto

2nd B1 for all 4 correct expected frequencies

1st M1 for at least 2 correct calculations from 3rd or 4th row

1st A1 for all correct calculations to at least 3sf if row 4
If awrt 2.91 is seen with no incorrect working award B1M1A1A1

2nd dM1 Dep on 1st M1 for a correct statement linking their test statistic and their cv (cv > 3.5)

3rd A1 for a correct comment in context - must mention "hair colour" and "ratios" or "model"
e.g. "There is evidence of to support the given model" No follow through
If hypotheses are the wrong way round score A0.

Question Number	Scheme	Marks
5 (a)	$\bar{x} = \frac{1}{2}(118.8 + 121.2) = 120$	B1
	1.6449 (or better)	B1
	" their 1.6449" $\frac{\sigma}{\sqrt{n}} = 121.2 - 120$	M1
	2.3263 (or better)	B1
	" their 2.3263" $\frac{\sigma}{\sqrt{n}} = 2.3263 \times \left(\frac{121.2 - 120}{1.6449}\right)$	dM1
	So 98% C.I. = $120 \pm 1.424\ldots$ = (118.3028…, 121.699…)	
	awrt **(118, 122)**	A1
		(6)
(b)	awrt $(118\pi, 122\pi)$ or (371/372, 382/383)	B1ft
		(1)
(c)	P (All) $= (0.98)^3$	M1
	$= 0.941$	A1
		(2)
		[Total 9]

Notes

(a)

NB in part (a) only lose one of the B1 marks for not using the percentage points table

1st B1 for $\bar{x} = 120$

2nd B1 for 1.6449 or better in an attempt (could be $1.6449\sigma = k$ or even $1.6449\ \sigma^2 = k$)
Condone strange notation for standard error (E) here if it is used correctly

1st M1 for an attempt to find "width" or "half-width" of a 90% CI ft their z value ($|z| > 1.5$)
e.g. for z $E = 121.2 - 120$ (o.e.) N.B. $E = 0.7295...$ Condone missing 2 here.

3rd B1 for 2.3263 or better in an attempt at CI.
If score 2nd B0 for using 1.64 or 1.645 allow 3rd B1 for 2.32 or 2.33 here

2nd dM1 for a correct attempt at "width" or "half-width" of a 98% CI ft their z value ($|z| > 2$)
Dependent on 1st M1 and ft their value or expression for s.e.

A1 for lower limit in range [118, 118.35) and upper limit in range (121.65, 122]

Answer only of awrt (118, 122) with no incorrect working seen scores 6/6/ if 1.6449 and 2.3263 are seen and 5/6 (B1B1M1B0M1A1) otherwise.

(c)

M1 for a correct expression i.e. $(0.98)^3$

A1 for awrt 0.941

Question Number	Scheme	Marks
6 (a)	Var $(X) = \dfrac{(a+5-a+1)^2}{12}$ [=3]	M1
	$\overline{X} \sim \text{N}\left(a+2, \dfrac{3}{50}\right)$	A1, A1ft
		(3)
(b)	$17.2 - 1.96\times\sqrt{\dfrac{3}{50}} < \mu < 17.2 + 1.96\times\sqrt{\dfrac{3}{50}}$	B1 M1
	$17.2 - 1.96\times\sqrt{\dfrac{3}{50}} < a+2 < 17.2 + 1.96\times\sqrt{\dfrac{3}{50}}$	B1
	$14.7 < a < 15.7$	A1
		(4)
		[Total 7]

	Notes
(a)	M1 for a correct expression for Var(X) in terms of a or Var(X) = 3
	1st A1 for normal and correct mean must be $a + 2$
	NB N(17.2, ...) is A0 and N(17.2, $\frac{3}{50}$) is M1A0A1
	2nd A1ft for correct Var($\overline{X}$), i.e. (their "3")/50
(b)	1st B1 for correct use of $z = 1.96$ in an attempt e.g. $\overline{x} \pm z\sigma$ or $\overline{x} \pm z\sigma^2$
	M1 for $17.2 \pm z \times \sqrt{\dfrac{"3"}{50}}$ where $\lvert z\rvert > 1.5$ accept just + or just −
	Answer of (16.7, 17.7) scores B1M1B0A0
	2nd B1 for either of the inequalities with $a + 2$ and any z ($\lvert z\rvert > 1.5$) or $a = 15.2 \pm z \times \sqrt{\dfrac{"3"}{50}}$
	A1 for awrt 14.7 and 15.7

Question Number	Scheme	Marks
7 (a)	$H_0: \mu_a = \mu_b$, $H_1: \mu_a < \mu_b$	B1
	s.e. $= \sqrt{\frac{25^2}{100} + \frac{10^2}{150}}$, $z = \frac{67-60}{\sqrt{\frac{25^2}{100} + \frac{10^2}{150}}}$ CR $= 1.6449 \times \sqrt{\frac{25^2}{100} + \frac{10^2}{150}}$	M1,dM1
	$z = \pm 2.6616....$ $= \pm 4.326...$ (awrt **2.66/4.33**)	A1
	One tailed critical value $z = 1.6449$ (or prob of awrt 0.004 (<0.05))	B1
	[Condone 0.996 if compared correctly with 0.95 for the B1]	
	2.6616 > 1.6449 so] significant evidence to reject H_0	dM1
	There is evidence that the amount of lead present in the soil has decreased.	A1ft
		(7)
(b)	CLT enables you to assume that means are normally distributed	B1
		(1)
(c)	Have assumed $s^2 = \sigma^2$ or variance of sample = variance of population	B1
		(1)
		[Total 9]

	Notes
(a)	1st B1 for both hypotheses in terms of μ not words. Accept μ_1, μ_2 etc if there is some indication of which is which e.g $X \sim N(67, 25^2)$ implies X is "before". 1st M1 for attempt at s.e. - condone one number wrong or mis-matched variances i.e. $\sqrt{\frac{p}{q} + \frac{r}{s}}$ (3 of p,q,r& s correct) or $\sqrt{\frac{10^2}{100} + \frac{25^2}{150}}$ 2nd dM1 Dep on 1st M1 for using their s.e. in correct formula for test statistic. Num of $\pm (67 - 60)$ or for correct expression for CR 3rd dM1 dep. on 2nd M1 for a correct statement based on their normal cv ($\|cv\| > 1.5$) and their test statistic 2nd A1ft for correct comment in context. Must mention "lead" or "soil" and "factory". Allow ft If hypotheses are the wrong way round score A0 If hypotheses are not for a difference between 2 means award A0
(b)	B1 must mention mean and normal. In words or symbols e.g. $\bar{X} \sim N(...$

Question Number	Scheme	Marks
8 (a)	Let $W = D_1 - D_2$	M1
	$W \sim N(0, 2.88)$	A1, A1
	$P(\lvert W \rvert > 3) = 2 \times P(W > 3)$	M1
	$= 2 \times P\left(Z > \frac{3-0}{\sqrt{2.88}}\right)$	dM1
	$= 2 \times P(Z > 1.76776\ldots)$	
	$= 2 \times (1 - 0.9616)$	
	$= 0.0768$ awrt **0.077**	A1
		(6)
(b)	Let $T = 5C - 4D$ or $4D - 5C$ or $C - \frac{4}{5}D$ or $\frac{4}{5}D - C$	M1
	$T \sim N(\pm 4, 39.04)$ or $N(\pm 0.8, 1.5616)$	A1 A1
	$P(T < 0) = P\left(Z < \frac{0-4}{\sqrt{39.04}}\right)$ or $P\left(Z < \frac{0-0.8}{\sqrt{1.5616}}\right)$	M1
	$= P(Z < -0.64018\ldots)$	
	$= (1 - 0.7389)$	
	$= 0.2611$ awrt **0.261**	A1
		(5)
(c)	Let $P = D + D + D + D + D + D + B$	M1
	Let $Q = C + C + C + C + C + C + B$	
	$P \sim N(352, 13.64)$ and $Q \sim N(292, 8.84)$	A1, A1
	[Let $R = P - Q$] $R \sim N(\pm 60, 22.48)$	M1
	$P(R > 50) = P\left(Z > \frac{50-60}{\sqrt{22.48}}\right)$	dM1
	$= P(Z > -2.10\ldots)$	
	$= 0.9821$ awrt **0.982 ~ 0.983**	A1
		(6) **[Total 17]**

Notes

Award full marks in each part for a correct answer with no incorrect working seen.

(a)
1st M1 for explicitly defining a suitable W and attempt to find the distribution of W. May be implied by sight of N(0, 2.88)
1st A1 for normal and mean of 0, 2nd A1 for variance of 2.88. Award M1A1A1 for N(0, 2.88) seen.
2nd M1 for realising need $2 \times P(W > 3)$
3rd dM1 Dep on 1st M1 for standardising with 3, 0 and their s.d. Must lead to $P(Z >$ +ve) (o.e.)

(b)
1st M1 for explicitly defining a suitable T but may be implied by sight of one of these normals
1st A1 for normal and correct mean, 2nd A1 for correct variance. Accept awrt 3sf i.e. 39.0, 1.56
2nd M1 for standardising with 0 and their mean and their s.d. Must lead to $P(Z <$ -ve) (o.e.)

(c)
1st M1 for explicitly defining a correct P or Q. May be implied by a correct distribution for P or Q
1st A1 for a correct distribution for P 2nd A1 for a correct distribution for Q
2nd M1 for attempting R and obtaining its distribution- ft their P and Q means and variances
3rd dM1 for attempting $P(R > 50)$ and standardising with 50 and their $E(R)$ and their $\sqrt{\text{Var}(R)}$
Dependent on 2nd M1. Must lead to a $P(Z >$ -ve) (o.e.)

Further copies of this publication are available from
Edexcel Publications, Adamsway, Mansfield, Notts, NG18 4FN

Telephone 01623 467467
Fax 01623 450481
Email publication.orders@edexcel.com

Order Code UA037014 Summer 2013

For more information on Edexcel qualifications, please visit our website
www.edexcel.com

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

Ofqual

Llywodraeth Cynulliad Cymru
Welsh Assembly Government

Rewarding Learning