edexcel

Mark Scheme (Results)

Summer 2012

GCE Statistics S3 (6691) Paper 1

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 2012
Publications Code UA033146

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.

Hypothesis Tests (Final M1A1)

For an incorrect comparison (e.g. probability with *z* value) even with a correct statement and/or comment award M0A0

For a correct or no comparison with more than one statement one of which is false
Award M0A0 (This is compatible with the principle above of contradictory statements being penalised)

Apply these rules to all questions

June 2012
6691 Statistics S3
Mark Scheme

Question Number	Scheme	Marks
1 (a)		

X	Y	Rank X	Rank Y	d	d^2
62	54	3	2	1	1
56	47	4	5	-1	1
87	71	1	1	0	0
54	50	5	3	2	4
65	49	2	4	-2	4
15	25	6	8	-2	4
12	30	7	7	0	0
10	44	8	6	2	4

M1
M1

$\sum d^2 = 18$ A1

$r_s = 1 - \frac{6\times 18}{8(64-1)} = 0.7857...$ awrt 0.786 M1A1

(5)

1(b)

$H_0 : \rho = 0$ B1

$H_0 : \rho > 0$ B1

Critical region $r_s > 0.6429$ B1

(0.7857>0.6429 sufficient evidence to) reject H_0 M1

There is evidence of agreement between the scores awarded by each manager A1ft

(5)

1(c)

(*A* and *D* are now) tied ranks (for Manager *Y*) B1

Average rank (awarded to *A* and *D*) **and** use $r_s = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$ B1

(2)

Total 12

Notes

1(a)

1st M1 for an attempt to rank score *X* and score *Y*

2nd M1 for attempting d^2 for their ranks. Must be using ranks.

1st A1 for sum of 18

3rd M1 for use of the correct formula with their $\sum d^2$. If answer is not correct an expression is required.

2nd A1 for awrt 0.786

1(b)

1st B1 for null hypotheses in terms of ρ or ρ_s

2nd B1 for alt hyp as given

3rd B1 for cv of +0.6429 (or 0.7381 if two tailed from hyp)

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

A1ft for a correct contextualised comment. Must mention "scores / rankings" and "manager"

Follow through their r_s and their cv (provided it is |cv| <1

Use of "association" is A0

1(c)

1st B1 Tied ranks can be implied by 2.5, 6.5 or **both** 2 or 6 or description.

2nd B1 Average rank implied by 2.5 or 6.5 or description and 'use of pmcc'.

Question Number	Scheme	Marks
2(a)	Sampling frame within each species of fish in the lake impossible to obtain.	B1 **(1)**
2(b)	Quota sampling	B1 **(1)**
2(c)	Advantages: Sample can be obtained quickly Costs are kept to a minimum Administration of survey is easy Disadvantages: Not possible to estimate sampling errors Process not random Surveyor may not be able to identify species of fish easily	B1 B1 **(2)**
2(d)	<table><tr><th>Species</th><th>Quota</th></tr><tr><td>Trout</td><td>$\frac{1400}{2450}\times 30 = 17.14$</td></tr><tr><td>Bass</td><td>$\frac{600}{2450}\times 30 = 7.35$</td></tr><tr><td>Pike</td><td>$\frac{450}{2450}\times 30 = 5.51$</td></tr></table> Fish are caught from the lake until the quota of 17 trout, 7 bass and 6 pike are reached. If a fish is caught and the species quota is full, then this is ignored.	B1B1B1 B1 (4) **Total 8**
	Notes	
2(a)	'You can't / it's very difficult to number all the fish' or equivalent	
2(c)	Correct answer to (b) required. Some detail required.	
2(d)	1st B1 any one correct calculation seen or implied 2nd B1 all correct to at least 1 dp 3rd B1 for 17,7,6 4th B1 accept equivalent statement. Require comment on what to do with 'extra fish'.	

Question Number	Scheme	Marks
3(a)	($X_1, X_2, X_3 ..., X_n$ is a random) **sample** of size n , for n is **large**, (from a population with mean μ and variance σ^2) then $\bar{X}$ is (approximately) Normal.	B1 B1 **(2)**
3 (b)	$\bar{x} = \frac{1740000}{100} = 17400$	B1
	$\bar{x} \pm z\frac{\sigma}{\sqrt{n}}, = 17400 \pm 1.96 \times \frac{5000}{\sqrt{100}}$	M1, B1
	[16420,18380]	A1A1 **(5)**
3(c)	$\bar{X}$: Normal (approx) by CLT, and normal needed to find CI.	B1,B1 **(2)**
3 (d)	20000 **above** upper confidence limit (**not** just outside) Complaint justified.	B1ft dB1ft **(2)**
		Total 11
3(b)	Notes Recognisable z value required for method. 2nd B1 1.96 or better seen award Final A1s accept 3sf if correct expression seen. 5/5 for [16420,18380]	

Question Number	Scheme	Marks
4	H_0 : Egg yield and breed of chicken are independent (not associated) H_1 : Egg yield and breed of chicken are dependent (associated)	B1

Egg Yield / Breed	Low	Medium	High	Total
Leghorn	$\frac{100\times36}{150}=24$	$\frac{100\times84}{150}=56$	$\frac{100\times30}{150}=20$	100
Cornish	$\frac{50\times36}{150}=12$	$\frac{50\times84}{150}=28$	$\frac{50\times30}{150}=10$	50
Total	36	84	30	150

M1A1

O	E	$\sum\frac{(O-E)^2}{E}$	$\sum\frac{O^2}{E}$
22	24	0.166667	20.2
52	56	0.285714	48.3
26	20	1.8	33.8
14	12	0.333333	16.3
32	28	0.571429	36.6
4	10	3.6	1.6

M1A1

$\sum\frac{(O-E)^2}{E}=6.757...$ or $\sum\frac{O^2}{E}-100=6.757...$ A1

$\nu=2, \chi^2_2(5\%)=5.991$ B1B1ft

(6.757>5.991 so sufficient evidence to) reject H_0 M1

Egg yield and breed of chicken are dependent (associated) A1

(10)

Total 10

Notes

B1 for both hypotheses. Must mention "yield" and "breed" in both but condone ditto marks.

Use of "relationship" or "correlation" or "connection"is B0

1st M1 for some use of $\frac{\text{Row Total}\times\text{Col.Total}}{\text{Grand Total}}$. May be implied by a correct E_i

1st A1 for all expected frequencies correct

2nd M1 for at least two correct terms or correct expressions with their E_i

2nd A1 for all correct terms. May be implied by a correct answer (2 sf or better)

3rd M1 for a correct statement linking their test statistic and their cv . Must be χ^2 not normal.

4th A1 for a correct comment in context - must mention "egg yield" and "breed of chicken" - condone "relationship" or "connection" here but not "correlation". No follow through e.g. "There is no evidence of a relationship between egg yield and breed of chicken" is A0 whatever their test stat or cv.

Question Number	Scheme	Marks
5(a)	$H_0 : \mu_A = \mu_B$ $H_1 : \mu_A \neq \mu_B$	B1
	$z = \dfrac{\pm(80-74)}{\sqrt{\dfrac{100}{29}+\dfrac{225}{26}}}$	M1A1
	$z = \pm 1.7247...$ awrt ± 1.72	A1
	1.7247>1.6449 o.e. so reject H_0	dM1
	There is evidence of a difference in the (mean) scores of their students.	A1 **(6)**
5(b)	(For z=1.6, test above not significant so no evidence of a difference.) For Mr A's claim, $H_0 : \mu_A = \mu_B$, $H_1 : \mu_A > \mu_B$, and critical value is z=1.2816	B1, B1
	(Both z values significant,) Mr Alan's claim supported.	B1 **(3)**
		Total 9
5(a)	Notes 1st M1 for attempt at s.e. (condone one number wrong) and for using their s.e. in correct formula for test statistic. 1st A1 for correct expression for se 2nd dM1 dep. on 1st M1 for a correct statement based on their normal cv and their test statistic 3rd A1 for correct comment in context. Must mention "scores" and "students / groups/classes" Award A0 for a one-tailed comment.	
5(b)	1st B1 Alternative hyp should be clearly defined	

Question Number	Scheme	Marks
6(a)	Mean= $\frac{1\times16+2\times20+...+6\times8}{100}=2.91$ **ag**	M1A1 **(2)**
6(b)	$p=\frac{2.91}{6}=0.485$	B1
	$a=100\times C_3^6\times0.485^3\times0.515^3=31.17$	M1A1
	$b=100\times0.485^6=1.3(0)$	A1 **(4)**
6(c)	H_0 : Binomial is a good fit	
	H_1 : Binomial is a not a good fit	B1

Number of defective items	0 or 1	2	3	4	5 or 6
O	22	20	23	17	18
E	12.41	24.82	31.17	22.01	9.59

M1

$\sum\frac{(O-E)^2}{E}=\frac{(22-12.41)^2}{12.41}+\frac{(20-24.82)^2}{24.82}+...+\frac{(18-9.59)^2}{9.59}=18.998...$ awrt 19.0 M1A1

ν =5-2=3 degrees of freedom B1

$\chi_3^2(5\%)=7.815$ B1ft

18.998...>7.815 so reject H_0 M1

Binomial is a not a good fit (and is not a good model for the number of defective items in samples of size 6) A1 **(8)**

Total 14

Notes

6(a) 1st M At least 2 correct terms on numerator and 100 for denominator.

6(b) 0.485 can be implied by at least 1 correct answer.

6(c) Accept awrt 2dp for final answers
Clear use of Binomial and x100 required for method.
Parameters in hyps award B0
1st M1 for combining either 0 and 1 or 5 and 6 or both. Require at least 1 value in a combined correct.
2nd M1 for attempting $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$, at least 2 correct expressions or values.

2nd A1 for a correct comment suggesting that Binomial model is not suitable. No ft
Condone parameters here.

Question Number	Scheme	Marks
7(a)	M : N$(177, 25)$, F : N$(163,16)$	
	$\mathrm{E}(M-F)=177-163=14$	B1
	$\mathrm{Var}(M-F)=25+16=41$	M1A1
	$M-F$: N$(14, 41)$	
	$\mathrm{P}(M-F>0)=\mathrm{P}\left(Z>\frac{-14}{\sqrt{41}}\right)$ or $\mathrm{P}\left(Z<\frac{14}{\sqrt{41}}\right)$	M1
	$=\mathrm{P}(Z<2.186...)$	
	$=0.9854$ or 0.9856 by calculator awrt 0.985 or 0.986	A1
		(5)
7(b)	$W=M_1+M_2+...M_6+F_1+F_2+...F_4$	
	$\mathrm{E}(W)=6\times177+4\times163$	
	$=1714$	B1
	$\mathrm{Var}(W)=6\times25+4\times16$	M1
	$=214$	A1
	$\mathrm{P}(W<1700)=\mathrm{P}\left(Z<\frac{1700-1714}{\sqrt{214}}\right)$ or $\mathrm{P}\left(Z>\frac{1714-1700}{\sqrt{214}}\right)$	M1
	$=\mathrm{P}(Z<-0.957..)$ awrt $Z<-0.96$ or $Z>0.96$	A1
	$=1-0.8315$	
	$=0.1685$ awrt 0.169	A1
	(0.1693 by calculator)	**(6)**
		Total 11
7(a)and (b)	Notes Condone reversed sds for method in (b) Accept metres: 2.14 award M1A0 in metres. 2nd M1s for identifying a correct probability and attempting to standardise with their mean and sd. Require explicit sd or accept 1156 for M1A0. This can be implied by the correct answer.	

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 UA033146 Summer 2012

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