edexcel
advancing learning, changing lives

# Mark Scheme (Results)
# Summer 2010

GCE

## GCE Statistics S3 (6691/01)

Edexcel Limited. Registered in England and Wales No. 4496750
Registered Office: One90 High Holborn, London WC1V 7BH

A PEARSON COMPANY

Edexcel is one of the leading examining and awarding bodies in the UK and throughout the world. We provide a wide range of qualifications including academic, vocational, occupational and specific programmes for employers.

Through a network of UK and overseas offices, Edexcel's centres receive the support they need to help them deliver their education and training programmes to learners.

For further information, please call our GCE line on 0844 576 0025, our GCSE team on 0844 576 0027, or visit our website at www.edexcel.com.

If you have any subject specific questions about the content of this Mark Scheme that require the help of a subject specialist, you may find our Ask The Expert email service helpful.

Ask The Expert can be accessed online at the following link:

http://www.edexcel.com/Aboutus/contact-us/

Summer 2010

Publications Code UA024774

All the material in this publication is copyright
© Edexcel Ltd 2010

**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

edexcel

# June 2010
# Statistics S3 6691
# Mark Scheme

| Question Number | Scheme | Marks |
|---|---|---|
| Q1 | $H_0$: $\mu = 80$, $H_1$: $\mu > 80$ | B1,B1 |
| | $z = \dfrac{83-80}{\frac{15}{\sqrt{100}}} = 2$ | M1A1 |
| | $2 > 1.6449$ (accept 1.645 or better) | B1 |
| | Reject $H_0$ or significant result or in the critical region | M1 |
| | Managing director's claim is supported. | A1 |
| | | **7** |

1st B1 for $H_0$. They must use $\mu$ not $x$, $p$, $\lambda$ or $\bar{x}$ etc
2nd B1 for $H_1$ (must be > 80). Same rules about $\mu$.
1st M1 for attempt at standardising using 83, 80 and $\frac{15}{\sqrt{100}}$. Can accept $\pm$.
May be implied by $z = \pm 2$
1st A1 for + 2 only

3rd B1 for $\pm 1.6449$ seen (or probability of 0.0228 or better)
2nd M1 for a correct statement about "significance" or rejecting $H_0$ (or $H_1$) based on their $z$ value and their 1.6449 (provided it is a recognizable critical value from normal tables) or their probability (< 0.5) and significance level of 0.05.
Condone their probability > 0.5 compared with 0.95 for the 2nd M1
2nd A1 for a correct contextualised comment. Must mention "director" and "claim" or "time" and "use of Internet". No follow through.

2nd M1A1

If no comparison or statement is made but a correct contextualised comment is given the M1 can be implied.
If a comparison is made it must be compatible with statement otherwise M0
e.g. comparing 0.0228 with 1.6449 is M0 or comparing probability 0.9772 with 0.05 is M0
comparing -2 with - 1.6449 is OK provided a correct statement accompanies it
condone -2 >-1.6449 provided their statement correctly rejects $H_0$.

Critical Region

They may find a critical region for $\bar{X}$: $\bar{X} > 80 + \frac{15}{\sqrt{100}} \times 1.6449$ = awrt 82.5

1st M1 for $80 + \frac{15}{\sqrt{100}} \times$ (z value)
3rd B1 for 1.645 or better
1st A1 for awrt 82.5
The rest of the marks are as per the scheme.

| Question Number | Scheme | Marks |
|---|---|---|
| Q2 | [ $P \sim N(90,9)$ and $J \sim N(91,12)$] | |
| (a) | $(J-P) \sim N(1,21)$ | M1, A1 |
| | $P(J<P) = P(J-P<0)$ | |
| | $= P\left(Z < \frac{0-1}{\sqrt{21}}\right)$ | dM1 |
| | $= P(Z<-0.2182...)$ | |
| | $= 1-0.5871 = 0.4129$ awrt (**0.413 ~ 0.414**) | A1 |
| | calculator (0.4136….) | (4) |
| (b) | $X = (J_1 + J_2 + .... + J_{60}) - (P_1 + P_2 + .... + P_{60})$ | M1 |
| | $E(X) = 60\times 91 - 60\times 90 = 60$ [stated as E($X$) = 60 or $X$~N(60, …)] | B1 |
| | $Var(X) = 60\times 9 + 60\times 12 = 1260$ | A1 |
| | $P(X>120) = P\left(Z > \frac{120-60}{\sqrt{1260}}\right)$ | M1 |
| | $= P(Z>1.69030...)$ | |
| | $= 1-0.9545 = 0.0455$ awrt (**0.0455**) | A1 |
| | | (5) |
| | | **9** |

(a)
1st M1 for attempting $J$ - $P$ and E($J$ - $P$) or $P$ - $J$ and E($P$ - $J$)
1st A1 for variance of 21 (Accept 9 + 12). Ignore any slip in $\mu$ here.
2nd dM1 for attempting the correct probability and standardising with their mean and sd.
This mark is dependent on previous M so if $J$ - $P$ ( or $P$ - $J$) is not being used score M0
If their method is not crystal clear then they must be attempting P($Z$< -ve value) or P($Z$ > +ve value) i.e. their probability <u>after</u> standardisation should lead to a prob. < 0.5 so e.g. P($J$ - $P$ < 0) leading to 0.5871 is M0A0 unless the M1 is clearly earned.
2nd A1 for awrt 0.413 or 0.414

**The first 3 marks may be implied by a correct answer**

(b)
1st M1 for a clear attempt to identify a correct form for $X$. This may be implied by correct variance of 1260
B1 for E($X$) = 60. Can be awarded even if they are using $X$ = 60$J$ - 60$P$. Allow $P$ - $J$ and -60
1st A1 for a correct variance. If 1260 is given the M1 is scored by implication.
2nd M1 for attempting a correct probability and standardising with 120 and their 60 and 1260
If the answer is incorrect a full <u>expression</u> must be seen following through their values for M1 e.g. $P\left(Z > \frac{120 - \text{ their } 60}{\sqrt{\text{their variance}}}\right)$. If using -60, should get $P\left(Z < \frac{-120 - -60}{\sqrt{\text{their variance}}}\right)$

Use of means
Attempt to use $\bar{J} - \bar{P}$ for 1st M1, E($\bar{J} - \bar{P}$) = 1 for B1 and Var($\bar{J} - \bar{P}$) = 0.35 for A1
Then 2nd M1 for standardisation with 2, and their 1 and 0.35

| Question Number | Scheme | Marks |
|---|---|---|
| Q3 (a) | $E \sim \mathrm{N}(0, 0.5^2)$ or $X \sim \mathrm{N}\left(w, 0.5^2\right)$ | |
| | $\mathrm{P}(\lvert E\rvert < 0.6) = \mathrm{P}\left(\lvert Z\rvert < \frac{0.6}{0.5}\right)$ or $\mathrm{P}\left(\lvert X - w\rvert < 0.6\right) = \mathrm{P}\left(\lvert Z\rvert < \frac{0.6}{0.5}\right)$ | M1 |
| | $= \mathrm{P}(\lvert Z\rvert < 1.2)$ | |
| | $= 2 \times 0.8849 - 1 = 0.7698$ awrt **0.770** | A1 |
| | | (2) |
| (b) | $\bar{E} \sim \mathrm{N}\left(0, \frac{1}{64}\right)$ or $\bar{X} \sim \mathrm{N}\left(w, \frac{0.5^2}{16}\right)$ | M1 |
| | $\mathrm{P}\left(\lvert \bar{E}\rvert < 0.3\right) = \mathrm{P}\left(\lvert Z\rvert < \frac{0.3}{\frac{1}{8}}\right)$ or $\mathrm{P}\left(\lvert \bar{X} - w\rvert < 0.3\right) = \mathrm{P}\left(\lvert Z\rvert < \frac{0.3}{\frac{1}{8}}\right)$ | M1, A1 |
| | $= \mathrm{P}(\lvert Z\rvert < 2.4)$ | |
| | $= 2 \times 0.9918 - 1 = 0.9836$ awrt **0.984** | A1 |
| | | (4) |
| (c) | $35.6 \pm 2.3263 \times \frac{1}{8}$ | M1 B1 |
| | (**35.3, 35.9**) | A1,A1 |
| | | (4) |
| | | **10** |

(a) 1st M1 for identifying a correct probability (they must have the 0.6) and attempting to standardise. Need | |. This mark can be given for 0.8849 - 0.1151 seen as final answer.
1st A1 for awrt 0.770. NB an answer of 0.3849 or 0.8849 scores M0A0 (since it implies no | |)
**M1 may be implied by a correct answer**

(b) 1st M1 for a correct attempt to define $\bar{E}$ or $\bar{X}$ but must attempt $\frac{\sigma^2}{n}$. Condone labelling as *E* or *X*
This mark may be implied by standardisation in the next line.
2nd M1 for identifying a correct probability statement using $\bar{E}$ or $\bar{X}$ . Must have 0.3 and | |
1st A1 for correct standardisation as printed or better
2nd A1 for awrt 0.984
**The M marks may be implied by a correct answer.**

Sum of 16, not means
1st M1 for correct attempt at suitable sum distribution with correct variance ( = $16 \times \frac{1}{4}$)
2nd M1 for identifying a correct probability. Must have 4.8 and | |
1st A1 for correct standardisation i.e. need to see $\frac{4.8}{\sqrt{4}}$ or better

(c) M1 for $35.6 \pm z \times \frac{0.5}{\sqrt{16}}$
B1 for 2.3263 or better. Use of 2.33 will lose this mark but can still score ¾
1st A1 for awrt 35.3
2nd A1 for awrt 35.9

| Question Number | Scheme | Marks |
|---|---|---|
| Q4 (a) | <table><tr><td>Distance rank</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr><tr><td>Depth rank</td><td>1</td><td>2</td><td>4</td><td>3</td><td>6</td><td>7</td><td>5</td></tr><tr><td>$\lvert d\rvert$</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>2</td></tr><tr><td>$d^2$</td><td>0</td><td>0</td><td>1</td><td>1</td><td>1</td><td>1</td><td>4</td></tr></table> | M1<br>M1 |
| | $\sum d^2 = 8$ | M1A1 |
| | $r_s = 1 - \frac{6\times 8}{7\times 48}$ | M1 |
| | $= \frac{6}{7} = 0.\dot{8}57142\dot{}$ awrt **0.857** | A1<br>(6) |
| (b) | $H_0 : \rho = 0, H_1 : \rho > 0$ | B1 |
| | Critical value at 1% level is 0.8929 | B1 |
| | $r_s < 0.8929$ so not significant evidence to reject $H_0$, | M1 |
| | The researcher's claim is not correct (at 1% level).<br>or insufficient evidence for researcher's claim<br>or there is insufficient evidence that water gets deeper further from inner bank.<br>or no (positive) correlation between depth of water and distance from inner bank | A1ft<br>(4) |
| | | **10** |

(a)
1st M1 for an attempt to rank the depths against the distances
2nd M1 for attempting *d* for their ranks. Must be using ranks.
3rd M1 for attempting $\sum d^2$ (must be using ranks)
1st A1 for sum of 8 (or 104 for reverse ranking)
4th 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 (±) 0.857. Sign should correspond to ranking (so use of 104 should get -0.857)

(b)
1st B1 for both hypotheses in terms of $\rho$, $H_1$ must be one tail and compatible with their ranking
2nd B1 for cv of 0.8929 (accept ±)
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 "researcher" and "claim" or "distance (from bank)" and "depth (of water)"
Follow through their $r_s$ and their cv (provided it is |cv| <1)
Use of "association" is A0

| Question Number | Scheme | Marks |
|---|---|---|
| Q5 | | |

| Finances / Income | Worse | Same | Better | |
|---|---|---|---|---|
| Under £15 000 | 10.54 | 10.54 | 12.92 | 34 |
| £15 000 and above | 20.46 | 20.46 | 25.08 | 66 |
| | 31 | 31 | 38 | 100 |

M1
A1

$H_0$ : State of finances and income are independent (not associated)

$H_1$ : State of finances and income are not independent (associated)

B1

| $O_i$ | $E_i$ | $\frac{(O_i - E_i)^2}{E_i}$ | $\frac{O_i^2}{E_i}$ |
|---|---|---|---|
| 14 | 10.54 | 1.1358…. | 18.59.. |
| 11 | 10.54 | 0.0200…. | 11.48.. |
| 9 | 12.92 | 1.1893… | 6.269.. |
| 17 | 20.46 | 0.5851… | 14.12.. |
| 20 | 20.46 | 0.0103… | 19.55.. |
| 29 | 25.08 | 0.6126… | 33.53.. |

M1
A1

$\sum \frac{(O_i - E_i)^2}{E_i} = 3.553\ldots$ or $\sum \frac{O_i^2}{E_i} - 100 = 103.553... - 100 = 3.553...$ (awrt **3.55)**

A1

$\nu = (3-1)(2-1) = 2$

B1

cv is 5.991

B1

3.553 < 5.991 so insufficient evidence to reject $H_0$ or not significant

M1

There is no evidence of association between state of finances and income.

A1

**10**

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

1st A1 for all expected frequencies correct

B1 for both hypotheses. Must mention "state" or "finances" and "income" at least once
Use of "relationship" or "correlation" or "connection"is B0

2nd M1 for at least two correct terms (as in 3rd or 4th column) or correct expressions with their $E_i$

2nd A1 for all correct terms. May be implied by a correct answer.(2 dp or better-allow eg 1.13…)

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

4th A1 for a correct comment in context - must mention "state" or "finances" and "income"
condone "relationship" or "connection" here but **not** "correlation". No follow through.
e.g. "There is no evidence of a relationship between finances and income"

| Question Number | Scheme | Marks |
|---|---|---|

Q6

| Distance from centre of site (m) | 0-1 | 1-2 | 2-4 | 4-6 | 6-9 | 9-12 |
|---|---|---|---|---|---|---|
| $b-a$ | 1 | 1 | 2 | 2 | 3 | 3 |
| No of artefacts | 22 | 15 | 44 | 37 | 52 | 58 |
| $\mathrm{P}(a \le X < b)$ | $\frac{1}{12}$ | $\frac{1}{12}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{4}$ | $\frac{1}{4}$ |
| $228 \times \mathrm{P}(a \le X < b)$ | 19 | 19 | 38 | 38 | 57 | 57 |

M1

A1

A1

| Class | $O_i$ | $E_i$ | $\frac{(O_i - E_i)^2}{E_i}$ | $\frac{O_i^2}{E_i}$ |
|---|---|---|---|---|
| 0-1 | 22 | 19 | $\frac{9}{19} = 0.4736...$ | 25.57… |
| 1-2 | 15 | 19 | $\frac{16}{19} = 0.8421...$ | 11.84… |
| 2-4 | 44 | 38 | $\frac{36}{38} = 0.9473...$ | 50.94… |
| 4-6 | 37 | 38 | $\frac{1}{38} = 0.0263...$ | 36.02… |
| 6-9 | 52 | 57 | $\frac{25}{57} = 0.4385...$ | 47.43… |
| 9-12 | 58 | 57 | $\frac{1}{57} = 0.0175...$ | 59.01… |

M1

A1

$H_0$: <u>continuous uniform</u> distribution <u>is</u> a good fit

$H_1$: <u>continuous uniform</u> distribution <u>is not</u> a good fit

B1

$\sum \frac{(O_i - E_i)^2}{E_i} = \frac{313}{114} = 2.75$ <u>or</u> $\sum \frac{O_i^2}{E_i} - 228 = 230.745... - 228 = \ldots$ (awrt **2.75**)

dM1A1

$\nu = 6 - 1 = 5$

B1

$\chi_5^2(0.05) = 11.070$ (ft their $\nu$ i.e. $\chi_\nu^2(0.05)$)

B1ft

2.75<11.070, insufficient evidence to reject $H_0$

M1

Continuous uniform distribution is a suitable model

A1

**12**

1st M1 for calculation of at least 3 widths and attempting proportions/probs. <u>or</u> for 1:2:3 ratio seen

1st A1 for correct probabilities

2nd A1 for all correct expected frequencies

2nd M1 for attempting $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$, at least 3 correct expressions or values.

Follow through their $E_i$ provided they are not all = 38

3rd A1 for a correct set of calcs - 3rd or 4th column. (2 dp or better and allow e.g. 0.94…)

3rd dM1 **dependent on 2nd M1** for attempting a correct sum or calculation (must see at least 3 terms and +)

**The first three Ms and As can be implied by a test statistic of awrt 2.75**

4th M1 for a correct statement based on their test statistic ( > 1) and their cv (> 3.8)

Contradictory statements score M0 e.g. "significant" do not reject $H_0$.

5th A1 for a correct comment suggesting that continuous uniform model is suitable. No ft

| Question Number | | Scheme | Marks |
|---|---|---|---|
| Q7 | (a) | Label full time staff 1-6000, part time staff 1-4000 | M1 |
| | | Use random numbers to select | M1 |
| | | Simple random sample of 120 full time staff and 80 part time staff | A1 (3) |
| | (b) | Enables estimation of statistics / errors for each strata <u>or</u> "reduce variability" <u>or</u> "more representative" <u>or</u> "reflects population structure" **NOT** "more accurate" | B1 (1) |
| | (c) | $H_0: \mu_f = \mu_p$, $H_1: \mu_f \neq \mu_p$ (accept $\mu_1, \mu_2$) | B1 |
| | | s.e. $= \sqrt{\frac{21}{80}+\frac{19}{80}}$, $z = \frac{52-50}{\sqrt{\frac{21}{80}+\frac{19}{80}}} = \left(2\sqrt{2}\right)$ | M1,M1 |
| | | $= 2.828\ldots$ (awrt **2.83**) | A1 |
| | | Two tailed critical value z = 2.5758 (or prob of awrt 0.002 (<0.005) or 0.004 (<0.01)) | B1 |
| | | [2.828 > 2.5758 so] significant evidence to reject $H_0$ | dM1 |
| | | There is evidence of a difference in policy awareness between full time and part time staff | A1ft (7) |
| | (d) | Can use mean full time and mean part time | B1 |
| | | ~ Normal | B1 (2) |
| | (e) | Have assumed $s^2 = \sigma^2$ or variance of sample = variance of population | B1 (1) |
| | (f) | 2.53 < 2.5758, not significant <u>or</u> do not reject $H_0$ | M1 |
| | | So there is insufficient evidence of a difference in mean awareness | A1ft (2) |
| | (g) | Training course has closed the gap between full time staff and part time staff's mean awareness of company policy. | B1 (1) |
| | | | **17** |

| | | |
|---|---|---|
| | (a) | 1st M1 for attempt at labelling full-time and part-time staff. One set of correct numbers.<br>2nd M1 for mentioning use of random numbers<br>1st A1 for s.r.s. of 120 full-time and 80 part-time |
| | (c) | 1st M1 for attempt at s.e. - condone one number wrong . NB correct s.e. = $\sqrt{\frac{1}{2}}$<br>2nd M1 for using their s.e. in correct formula for test statistic. Must be $\frac{\pm(52-50)}{\sqrt{\frac{p}{q}+\frac{r}{s}}}$<br>3rd dM1 **dep. on 2nd M1** for a correct statement based on their normal cv and their test statistic<br>2nd A1 for correct comment in context. Must mention "scores" or " policy awareness" and types of "staff". Award **A0** for a one-tailed comment. Allow ft |
| | (d) | 1st B1 for mention of mean(s) <u>or</u> use of $\bar{X}$, provided $\bar{X}$ clearly refers to full-time or part-time<br>2nd B1 for stating that distribution can be assumed normal<br>e.g. "mean score of the test is normally distributed" gets B1B1 |
| | (f) | M1 for correct statement (may be implied by correct contextualised comment)<br>A1 for correct contextualised comment. Accept "no difference in mean scores". Allow ft |
| | (g) | B1 for correct comment in context that implies training was effective.<br>This must be supported by their (c) and (f). Condone one-tailed comment here. |

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

Telephone 01623 467467
Fax 01623 450481

Email publications@linneydirect.com

Order Code UA024774 Summer 2010

For more information on Edexcel qualifications, please visit www.edexcel.com/quals

Edexcel Limited. Registered in England and Wales no.4496750
Registered Office: One90 High Holborn, London, WC1V 7BH