

**Goodness of Fit Tests (Year 1) (From OCR 4768)**

**Q1, (Jun 2010, Q3b)**

$H_0$ : Stock market prices can be modelled by Benford's Law.

$H_1$ : Stock market prices can not be modelled by Benford's Law.

Prob	0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.046
Exp f	60.2	35.2	25.0	19.4	15.8	13.4	11.6	10.2	9.2
Obs f	55	34	27	16	15	17	12	15	9

$$\begin{aligned} \chi^2 &= 0.44917 + 0.04091 + 0.16 + 0.59588 + 0.04051 \\ &\quad + 0.96716 + 0.01379 + 2.25882 + 0.00435 \\ &= 4.5305(9) \end{aligned}$$

Refer to  $\chi^2_8$ .

Upper 5% point is 13.36.

Not significant.

Suggests Benford's Law provides a reasonable model in the context of share prices.

- M1 Probs  $\times$  200 for expected frequencies.  
All correct.
- M1 Calculation of  $\chi^2$ .
- A1 c.a.o.
- M1 Allow correct df (= cells - 1) from wrongly grouped table and ft. Otherwise, no ft if wrong.  
 $P(\chi^2 > 4.53059) = 0.80636$ .
- A1 No ft from here if wrong.
- A1 ft only c's test statistic.
- A1 ft only c's test statistic.

[7]

**Q2, (Jan 2012, Q3b)**

(i)	A paired test is used in this context in order to eliminate differences between health authorities.	E1 <b>[1]</b>	oe																		
(ii)	<table border="1" data-bbox="197 331 1086 411"> <tr> <td>Diff</td> <td>11</td> <td>26</td> <td>-15</td> <td>4</td> <td>-9</td> <td>-1</td> <td>23</td> <td>5</td> </tr> <tr> <td>Rank</td> <td>5</td> <td>8</td> <td>6</td> <td>2</td> <td>4</td> <td>1</td> <td>7</td> <td>3</td> </tr> </table> <p data-bbox="197 571 952 879"> <math>W_- = 1 + 4 + 6 = 11</math>                      Refer to tables of Wilcoxon paired (/single sample) statistic for <math>n = 8</math>.                      Lower 5% tail is 5 (or upper is 31 if 25 used).  <math>11 &gt; 5 \therefore</math> Result is not significant.                      No evidence to suggest a difference between the incidences of myocardial infarction in men and women on the whole.                 </p>	Diff	11	26	-15	4	-9	-1	23	5	Rank	5	8	6	2	4	1	7	3	<p data-bbox="994 453 1086 643">                         M1                          M1                          A1                          B1                          M1                     </p> <p data-bbox="994 687 1086 802">                         A1                          A1                          A1                     </p> <p data-bbox="994 884 1086 916"><b>[8]</b></p>	<p data-bbox="1106 453 2139 643">                         For differences. ZERO in this section if differences not used.                          For ranks.                          ft from here if ranks wrong.                          (or <math>W_+ = 2 + 3 + 5 + 7 + 8 = 25</math>)                          No ft from here if wrong.                     </p> <p data-bbox="1106 687 2139 879">                         ie a 2-tail test. No ft from here if wrong.                          ft only c's test statistic.                          ft only c's test statistic.                          "Non-assertive" conclusion in context to include "on the whole" oe.                     </p>
Diff	11	26	-15	4	-9	-1	23	5													
Rank	5	8	6	2	4	1	7	3													

**Q3, (Jun 2012, Q4i)**

$H_0$ : The model for the number of callouts fits the data  
 $H_1$ : The model for the number of callouts does not fit the data.

B1  
 B1

Do not allow "Data fit the model" o.e for either hypothesis.

Obs'd frequency	145	79	22	6	3	0
Exp'd frequency	139.947	83.968	25.190	5.038	0.756	0.101

Merge last 3 cells. Obs 9 Exp 5.895  
 $\chi^2 = 0.1824 + 0.2939 + 0.4040 + 1.6355$   
 $= 2.515(8)$

M1  
 M1  
 A1  
 M1

Calculation of  $\chi^2$ .  
 Cao Require 3/4 sf; condone up to 6.  
 Allow correct df (= cells - 2) from wrongly grouped table and ft.  
 Otherwise, no ft if wrong.  $P(\chi^2 > 2.5158) = 0.2842$ .  
 No fit from here if wrong.  
 ft only c's test statistic.  
 ft only c's test statistic. "Non-assertive" conclusion in words (+context).  
 Do not allow "Data fit model" o.e.

Refer to  $\chi^2$ .

Upper 5% point is 5.991.  
 Not significant.

Suggests it is reasonable to suppose that the model fits the data.

[9]

**Q4, (Jan 2013, Q4b)**

$H_0$ : The random number function is performing as it should.  
 $H_1$ : The random number function is not performing as it should.

All expected frequencies are 10  

$$\chi^2 = 1.6 + 0.4 + 0.1 + 1.6 + 0.4 + 0.1 + 2.5 + 2.5 + 1.6 + 1.6$$

$$= 12.4$$

Refer to  $\chi_9^2$ .

Upper 10% point is 14.68.  
 Not significant.  
 Insufficient evidence to suggest that the random number function is not performing as it should.

B1	Both hypotheses. Must be the right way round. Allow use of the uniform distribution/model. Do not accept "data fit model" oe.
B1	
M1	Calculation of $\chi^2$ .
A1	c.a.o.
M1	Allow correct df (= cells - 1) from wrongly grouped table and ft. Otherwise, no ft if wrong. $P(\chi^2 > 12.4) = 0.1916$ .
A1	No ft from here if wrong.
A1	ft only c's test statistic.
A1	ft only c's test statistic. Conclusion in context. Allow in terms of the uniform distribution/model. Do not accept "data fit model" oe.
<b>[8]</b>	

**Q5, (Jun 2016, Q2a)**

$H_0$ : The (genetic) model fits the data.

$H_1$ : The (genetic) model does not fit the data

Observed	125	37	32	6
Expected	112.5	37.5	37.5	12.5

Cont's	1.3889	0.0067	0.8067	3.38
--------	--------	--------	--------	------

$$X^2 = 5.582$$

Degrees of freedom = 3

Critical value = 9.348

$5.582 < 9.348 \rightarrow$  cannot reject  $H_0$

The data give no reason to doubt the genetic model

B1	Both hypotheses; Not 'data fits model'
B1	Expected values correct
M1	use of $(O-E)^2/E$ (at least one correct)
A1	all correct to 3dp where appropriate
A1	cao (3sf or 4sf)
B1	no FT if wrong (can be implied by 9.348)
B1	no FT if wrong
M1	FT their $X^2$
E1	Do not accept data fits model; but 'Evidence suggests that model first the data' is fine
[9]	

**Q6, (Jun 2017, Q2ii)**

$H_0$ : Judges awarded the same number of each mark.

$H_1$ : Judges did not award the same number of each mark.

observed	5	6	10	9	14	16	14	6
expected	10	10	10	10	10	10	10	10

$$\chi^2 = 2.5 + 1.6 + 0 + 0.1 + 1.6 + 3.6 + 1.6 + 1.6$$

$$= 12.6$$

Refer to  $\chi^2_7$

The 10% critical value is 12.02.

12.6 > 12.02 so significant

There is sufficient evidence that judges have not been awarding the same number of each mark.

B1

Both hypotheses. Must be in correct context. Allow 'uniform distribution' or 'in equal proportions'. 'Model fits data' or 'belief is justified' is ok. Do not accept 'data fits model' oe'.

B1

For expected frequencies.

M1

Calculation of  $\chi^2$ . (if 12.6 not seen, must see evidence of calculation)

A1

cao.

M1

No ft if wrong.

A1

No ft if wrong.

M1

ft their test statistic

A1

Must be in context and mention 'evidence'. ('organiser's belief' is sufficient context)

**[8]**