GCE Examinations
Advanced Subsidiary / Advanced Level
Statistics
Module S3

Paper E

MARKING GUIDE

This guide is intended to be as helpful as possible to teachers by providing concise solutions and indicating how marks should be awarded. There are obviously alternative methods that would also gain full marks.

Method marks (M) are awarded for knowing and using a method.

Accuracy marks (A) can only be awarded when a correct method has been used.

(B) marks are independent of method marks.

Written by Shaun Armstrong & Chris Huffer
© Solomon Press

These sheets may be copied for use solely by the purchaser’s institute.
S3 Paper E – Marking Guide

1. (a) total = 500 \therefore require \frac{1}{5} \ M1
giving 33, 28, 21, 18 respectively \ A1

(b) e.g. know that each group is represented proportionately
provides data for each strata as well for whole \ B2 (4)

2. \ \ H_0 : \text{discrete uniform is a suitable model} \ B1
\ \ H_1 : \text{discrete uniform is not a suitable model} \ M1 \ A1
exp. freqs = 80 \div 5 = 16

\[
\begin{array}{c|c|c|c}
O & E & (O - E) & \frac{(O - E)^2}{E} \\
16 & 16 & 0 & 0 \\
20 & 16 & 4 & 1 \\
14 & 16 & -2 & 0.25 \\
17 & 16 & 1 & 0.0625 \\
13 & 16 & -3 & 0.5625
\end{array}
\]

\therefore \ \Sigma \frac{(O - E)^2}{E} = 1.875 \ \ M1 \ A2

\nu = 5 - 1 = 4, \chi^2_{crit}(10\%) = 7.779 \ \ M1 \ A1

1.875 < 7.779 \therefore do not reject H_0
discrete uniform is a suitable model supporting psychologist’s theory \ A1 (9)

3. (a) \ \ H_0 : \mu = 5’6” \ \ H_1 : \mu > 5’6” \ B1
5% level \therefore \ C.R. is z > 1.6449 \ M2 \ A1
require = \frac{\Sigma \frac{X}{\sqrt{n}}}{\sqrt{n}} > 1.6449

giving C.R. \ X > 66.31 inches \ A1

(b) \ \ \bar{X} = \frac{822+12}{150} = 66.56 \ M1 \ A1

66.56 > 66.31 \therefore reject H_0 \ M1
there is evidence that mean height of women is > 5’6” \ A1 (10)

4. (a)

<table>
<thead>
<tr>
<th>capacity</th>
<th>1.1</th>
<th>1.3</th>
<th>1.6</th>
<th>2.1</th>
<th>2.4</th>
<th>2.6</th>
<th>2.8</th>
<th>3.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>sales</td>
<td>527</td>
<td>632</td>
<td>840</td>
<td>619</td>
<td>425</td>
<td>487</td>
<td>401</td>
<td></td>
</tr>
<tr>
<td>cap. rank</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sales rank</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>\Sigma d^2</td>
<td>140</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

r_i = 1 - \frac{6x^2}{8x^2} = ~0.6667 \ M2 \ A2

(b) \ \ H_0 : \rho = 0 \ \ H_1 : \rho \neq 0 \ B1
n = 8, 5% level \therefore \ C.R. is r_i < ~0.7381 or r_i > 0.7381 \ M1 \ A1
not in C.R. \therefore no evidence of correlation \ A1

(c) need variables to be jointly normally distributed for pmcc test
engine capacities are discrete so use Spearman’s \ B2 (12)
5. (a) let \( A \) = amount side length of red cube is longer than blue cube
\[ \therefore A \sim N(14.5 - 12.2, 16.0 + 9.0) = N(2.3, 25) \]
\[
P(A > 3) = P\left(Z > \frac{3 - 2.3}{\frac{5}{\sqrt{25}}}\right) = P(Z > 0.14) = 1 - 0.5557 = 0.4443
\]

(b) let \( C \) = amount red tower is taller than blue tower
\[ \therefore C \sim N(4 \times 14.5 - 5 \times 12.2, 4 \times 16 + 5 \times 9) = N(-3, 109) \]
\[
P(C > 0) = P\left(Z > \frac{0 + 3}{\sqrt{109}}\right) = P(Z > 0.29) = 1 - 0.6141 = 0.3859
\]

(c) e.g. likely to use smaller blocks higher up the tower

6. (a) expected freq. family/ITV = \( \frac{101 \times 132}{240} = 55.55 \)
family/Ch4 = \( \frac{85 \times 132}{240} = 46.75 \)
sports/ITV = \( \frac{101 \times 66}{240} = 27.78 \)
sports/Ch4 = \( \frac{85 \times 66}{240} = 23.38 \)
giving expected freqs 55.55 46.75 29.70 27.78 23.38 17.67 14.87 9.46

H\(_0\) : no difference in proportion of adverts on different channels
H\(_1\) : difference in proportion of adverts on different channels

\[ \therefore \sum \frac{(O - E)^2}{E} = 15.535 \]
\[ \nu = 4, \chi^2_{0.05}(5\%) = 9.488 \]
\[ 15.535 > 9.488 \therefore \text{significant} \]
there is evidence of different proportion of adverts on different channels

(b) e.g. advertisers perception of the type of people who watch each channel

7. (a) when a sample from any dist. is large, the dist. of the sample mean is
approximately normal with same mean and variance \( \frac{\sigma^2}{n} \)

(b) binomial with \( n = 10, p = \frac{1}{6} \)

(c) mean = \( np = 10 \times \frac{1}{6} = \frac{5}{3} \)
variance = \( npq = 10 \times \frac{1}{6} \times \frac{5}{6} = \frac{25}{18} \)

(d) let \( X \) = no. of sixes when throw 10 dice \( \therefore X \sim B(10, \frac{1}{6}) \)
\[ \therefore \bar{X} \sim N\left(\frac{5}{3}, \frac{25}{18}\right) \]
\[
P(\bar{X} > 1.8) = P\left(Z > \frac{1.8 - \frac{5}{3}}{\sqrt{\frac{25}{18}}}\right) = P(Z > 1.13) = 1 - 0.8708 = 0.1292
\]
# Performance Record – S3 Paper E

<table>
<thead>
<tr>
<th>Question no.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic(s)</td>
<td>sampling goodness of fit, discrete uniform</td>
<td>hyp. test on mean</td>
<td>Spearman's, hyp. test</td>
<td>linear comb. of Normal r.v.</td>
<td>conting. table</td>
<td>CLT, dist. of sample mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marks</td>
<td>4</td>
<td>9</td>
<td>10</td>
<td>12</td>
<td>12</td>
<td>14</td>
<td>14</td>
<td>75</td>
</tr>
</tbody>
</table>

Student

© Solomon Press