### AMA1501 Introduction to Statistics for Business

**Miscellaneous Questions Set 5 Outline Suggested Solution**

1. (a)

<table>
<thead>
<tr>
<th>Class mark (x)</th>
<th>Frequency, f</th>
</tr>
</thead>
<tbody>
<tr>
<td>150</td>
<td>3</td>
</tr>
<tr>
<td>250</td>
<td>7</td>
</tr>
<tr>
<td>350</td>
<td>16</td>
</tr>
<tr>
<td>450</td>
<td>26</td>
</tr>
<tr>
<td>600</td>
<td>21</td>
</tr>
<tr>
<td>775</td>
<td>13</td>
</tr>
<tr>
<td>925</td>
<td>9</td>
</tr>
<tr>
<td>1075</td>
<td>5</td>
</tr>
</tbody>
</table>

\[ \sum f = 100 \quad \sum fx = 55875 \quad \sum fx^2 = 36576875 \]

Mean = \( \frac{55875}{100} \) = 558.75 (S’000)

Standard deviation = \( \sqrt{\frac{100(36576875) - 55875^2}{100(100 - 1)}} \) = 232.61184 (S’000)

Mode = 400 + \( \frac{26 - 16}{(26 - 16) + (26 - 21/2)} \) \( (500 - 400) \) = 439.21568 (S’000)

(b) \( SK_1 = \frac{558.75 - 439.21568}{232.61184} \) = 0.5139

(c)

<table>
<thead>
<tr>
<th>Monthly revenue less than (S’000)</th>
<th>Cumulative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>200</td>
<td>3</td>
</tr>
<tr>
<td>300</td>
<td>10</td>
</tr>
<tr>
<td>400</td>
<td>26</td>
</tr>
<tr>
<td>500</td>
<td>52</td>
</tr>
<tr>
<td>700</td>
<td>73</td>
</tr>
</tbody>
</table>
### Monthly revenue less than (S'000) 

<table>
<thead>
<tr>
<th>Monthly revenue less than (S'000)</th>
<th>Cumulative frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>850</td>
<td>86</td>
</tr>
<tr>
<td>1000</td>
<td>95</td>
</tr>
<tr>
<td>1150</td>
<td>100</td>
</tr>
</tbody>
</table>

\[ Q_1 = 300 + \frac{15}{16}(400 - 300) = 393.75 \ (S'000) \]

\[ Q_3 = 700 + \frac{2}{13}(850 - 700) = 723.07692 \ (S'000) \]

\[ IQR = 723.07692 - 393.75 = 329.32692 \ (S'000) \]

(d) \[ \hat{p} = \left( \frac{400 - 350}{400 - 300} 
\times 16 + 26 + 21 + \frac{780 - 700}{850 - 700} \times 13 \right) / 100 = 0.6193 \]

A 95% C.I. for \( p \) is \[ 0.6193 \pm 1.96 \sqrt{\frac{0.6193 \times (1 - 0.6193)}{100}} \], i.e. \( 0.5242 < p < 0.7145 \)

2. (a) (i) No. of panels can be formed = \( \frac{30}{8} \times \frac{25}{20} \times 3 \times 5 = 85503600 \)

(ii) \[ \frac{10 \times 2 \times 8 \times 3 \times 20 \times 4}{85503600} = 0.0408 \]

(b) A: highly motivated supervisor
B: hard working supervisor

\[ \Pr(A) = 0.83 \quad \Pr(B) = 0.9 \quad \Pr(A \cap B) = 0.05 \]

(i) \[ \Pr(A \cup B) = 0.83 + 0.9 - (1 - 0.05) = 0.78 \]

(ii) \[ \Pr(A | B) = 1 - \frac{0.78}{0.9} = 0.1333 \]

(c) HR – selected respondent works in HR Dept.
IT – selected respondent works in IT Dept.
GA – selected respondent works in GA Dept.
A – selected respondent satisfied with the working environment

\[ \Pr(HR) = 0.25 \quad \Pr(IT) = 0.35 \quad \Pr(GA) = 0.4 \]

\[ \Pr(A | HR) = 0.9 \quad \Pr(A | IT) = 0.85 \quad \Pr(A | GA) = 0.88 \]
3. (a) $X$ – weight of package (grams), \[ X \sim N(310,8^2) \]

(i) \[ \Pr(X < 300) = \Pr(Z < -1.25) = 0.1056 \]

(ii) \[ \Pr(X < a) = \Pr\left(Z < \frac{a - 310}{8}\right) = 0.05 \Rightarrow \frac{a - 310}{8} = -1.645 \Rightarrow a = 296.84 \text{ grams} \]

(iii) \[ \bar{X} \sim N(310,8^2/20) \]
\[ \Pr(308 < \bar{X} < 314) \approx \Pr(-1.12 < Z < 2.24) = 0.856 \]

(b) $X$ – number of enquiries that were classified as hardware problem
\[ X \sim B(60, 0.4) \]

Since $n>30$, $np=24>5$, $nq=36>5$ and $0.1<p<0.9$, \[ X \sim N(24,14.4) \] approximately
\[ \Pr(X > 30) = \Pr(X > 30.5) \approx \Pr(Z > 1.71) = 0.0436 \]

(c) $X$ – number of complaints per day
\[ X \sim Po(2) \]
\[ \Pr(X \leq 1) = \sum_{x=0}^{1} \frac{e^{-2}2^x}{x!} = 0.4060 \]

$Y$ – number of days in a month of 31 days which have at most 1 complaint per day
\[ Y \sim B(31, 0.4060) \]

Since $n>30$, $np>5$, $nq>5$ and $0.1<p<0.9$, \[ Y \sim N(12.5862, 2.73425^2) \] approximately
\[ \Pr(Y \geq 16) = \Pr(Y > 15.5) \approx \Pr(Z > 1.07) = 0.1423 \]

4. (a) $\bar{X}_1 - \bar{X}_2$ - difference of average weekly sales ($$)
\[\bar{X}_1 - \bar{X}_2 \sim N \left(100000 - 100000, \frac{20000^2}{4} + \frac{20000^2}{4}\right)\]

\[\Pr(\bar{X}_1 - \bar{X}_2 < -10000) + \Pr(\bar{X}_1 - \bar{X}_2 > 10000) \approx \Pr(Z < -0.71) + \Pr(Z > 0.71) = 0.4778\]

(b) A 95% confidence interval for mean time spent per week is

\[1.5 \pm 2.064 \times \frac{0.8}{\sqrt{25}}, \text{ i.e., } 1.1698 < \mu < 1.8302\]

(c) \(X\) – weight of cake (grams)

\[\bar{X} \sim N \left(\mu, \frac{\sigma^2}{n}\right)\]

\(H_0 : \mu = 100\)

\(H_1 : \mu < 100\)

\(\alpha = 0.01\)

Critical region: \(z < -2.33\)

Under \(H_0\), test statistic \(z = \frac{94.6 - 100}{2.5/\sqrt{36}} = -12.96\)

Decision: Reject \(H_0\)

(d)

\(\bar{X}_1\) - sample mean score rated by senior staff

\(\bar{X}_2\) - sample mean score rated by junior staff

\[\bar{X}_1 - \bar{X}_2 \sim N \left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)\] approximately

\(H_0 : \mu_1 = \mu_2\)

\(H_1 : \mu_1 < \mu_2\)

\(\alpha = 0.01\)

Critical region: \(z < -2.33\)

Under \(H_0\), test statistic \(z = \left(\frac{78 - 81}{9^2} - 0\right)\) \(\approx \frac{1.706}{\sqrt{40 + 60}} = -1.706\)
Decision: Do not reject $H_0$

5. (a) D – paired difference of score

d: 6 11 6 6 9 9 0 -1 7

$n=10, \sum d = 58, \sum d^2 = 466$

$\bar{d} = \frac{58}{10} = 5.8, s_d = \sqrt{\frac{10(466)-58^2}{10(10-1)}} = 3.7947$

$H_0: \mu_d = 0$

$H_1: \mu_d > 0$

$\alpha = 0.05$

Critical region: $t > 1.833, \nu = 9$

Under $H_0$, test statistic $t = \frac{5.8 - 0}{3.7947 \sqrt{10}} = 4.8333$

Decision: Reject $H_0$

(b)

$H_0$: there is no preference between the models

$H_1$: there is preference between the models

$\alpha = 0.05$

$E_i = 125, i = 1, \ldots, 4$

Critical region: $\chi^2 > 7.815, \nu = 3$

Under $H_0$, test statistic $\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} = 9.2$

Decision: Reject $H_0$

(c) $H_0$: favouring the new design or not is independent of the age

$H_1$: favouring the new design or not and the age of customer are dependent

$\alpha = 0.01$

Critical region: $\chi^2 > 9.210, \nu = 2$

Expected frequencies:
<table>
<thead>
<tr>
<th></th>
<th>Teenager</th>
<th>Adult</th>
<th>Elderly</th>
</tr>
</thead>
<tbody>
<tr>
<td>Favour the new design</td>
<td>120.05</td>
<td>120.62</td>
<td>69.33</td>
</tr>
<tr>
<td>Do not favour the new design</td>
<td>92.95</td>
<td>93.38</td>
<td>53.67</td>
</tr>
</tbody>
</table>

Under $H_0$, test statistic $\chi^2 = \sum \frac{(O_y - E_y)^2}{E_y} = 32.47$

Decision: Reject $H_0$

6. (a) (i) 
\[ b = \frac{10(56089) - (683)(813)}{10(47405) - (683)^2} = 0.7421 \]
\[ a = \frac{813}{10} - 0.7421 \times \frac{683}{10} = 30.6147 \]
\[ \hat{y} = 30.6147 + 0.7421x \]

(ii) When $x = 80$, $\hat{y} = 30.6147 + 0.7421 \times 80 = 89.9825 \approx 90$

(b) (i) $\hat{y} = -29.6972 + 0.3209(\text{Population size}) + 4.8782(\text{Advertising cost})$

(ii) 
\[ a = 38798.92 - 2094.42 = 36704.5 \]
\[ b = 2 \]
\[ c = 12 - 2 - 1 = 9 \]
\[ d = 12 - 1 = 11 \]
\[ e = 36704.5/2 = 18352.25 \]
\[ f = 2094.42/9 = 232.713 \]
\[ g = 18352.25/232.713 = 78.86 \]

(iii) $R^2 = \frac{36704.5}{38798.92} = 0.946$

(iv) $H_0: \beta_i = 0$
\( H_1 : \beta_i \neq 0 \)

\( \alpha = 0.05 \)

Critical region: \( t < -2.262 \) and \( t > 2.262, \nu = 9 \)

Under \( H_0 \), test statistic \( t = \frac{0.3209 - 0}{0.0484} = 6.63 \)

Decision: Reject \( H_0 \)