Date: 21.06.2016  Instructor: Prof. Dr. Hüseyin Öğuz
Time: 11:00-12:30  Room: CL009
Student Registration No: ___________________
Student Name-Surname: ____________________________________

Important Note: Your own calculator and appendix tables/guidelines are allowable to use during exam with the prohibition of their exchanges.

1. (30 p) Suppose a random sample of 100 companies taken in 2005 showed that 21 offered high-deductible health insurance plans to their workers. A separate random sample of 120 firms taken in 2006 showed that 30 offered high-deductible health insurance plans to their workers. Based on the sample results, can you conclude that there is a higher proportion of the companies offering high-deductible health insurance plans to their workers in 2006 than in 2005? Conduct your hypothesis test at a level of significance α=0.05.

2. (35 p) At a university, a study was done to establish whether a relationship exists between students’ graduating grade point average (GPA) and the SAT verbal score when the student originally entered the university. The sample data are reported as follows:

<table>
<thead>
<tr>
<th>Y(GPA)</th>
<th>2.5</th>
<th>3.2</th>
<th>3.5</th>
<th>2.8</th>
<th>3.0</th>
<th>2.4</th>
<th>3.4</th>
<th>2.9</th>
<th>2.7</th>
<th>3.8</th>
</tr>
</thead>
<tbody>
<tr>
<td>X(SAT)</td>
<td>640</td>
<td>700</td>
<td>550</td>
<td>540</td>
<td>620</td>
<td>490</td>
<td>710</td>
<td>600</td>
<td>505</td>
<td>710</td>
</tr>
</tbody>
</table>

a) Prepare a scatter plot of the data and show the regression line fitted
b) Compute covariance and correlation coefficient
c) Compute and interpret least square coefficients $b_1$ and $b_0$
d) Determine the predicted GPA -value when SAT score = 650.

3. (35 p) The following table classifies a stock’s price change as up, down, or no change for both today’s and yesterday’s prices. Price changes were examined for 100 days. A financial theory states that stock prices follow what is called a ‘random walk’. This means, in part, that the price change today for a stock must be independent of yesterday’s price change. Test the hypothesis that daily stock price for this stock are independent. Let α=0.05.

<table>
<thead>
<tr>
<th>Price Change Today</th>
<th>Price Change Previous Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up</td>
<td>No Change</td>
</tr>
<tr>
<td>Up</td>
<td>14</td>
</tr>
<tr>
<td>No Change</td>
<td>6</td>
</tr>
<tr>
<td>Down</td>
<td>16</td>
</tr>
</tbody>
</table>

Recall:

$$S_X = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n - 1}}; S_Y = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}{n - 1}}; S_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{n - 1};$$

$$r_{XY} = \frac{S_{XY}}{S_X S_Y}; b_1 = \frac{\text{Cov}(X,Y)}{S_X^2}; b_0 = \bar{Y} - b_1 \bar{X}; \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}; \bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n};$$

$$x^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}; E_{ij} = \frac{R_i C_j}{n}$$
1. Let $A =$ Year 2005; $B =$ Year 2006. First identify the null and alternative hypothesis

$$H_0: P_A - P_B = 0$$
$$H_1: P_A - P_B < 0$$

To test $H_0: P_A - P_B = 0$ against the alternative $H_1: P_A - P_B < 0$ the decision rule is to reject $H_0$ and accept $H_1$ if

$$z < -z_{\alpha} \text{ (one - tail test, From Appendix Guidelines Diagram E - 4)}$$

$$z = \frac{(\hat{p}_A - \hat{p}_B) - (P_A - P_B)}{\sqrt{\hat{p}_o(1 - \hat{p}_o)\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}} < -z_{\alpha} \text{ (Guidelines E - 4) or}$$

$$\hat{p}_o = \frac{n_A\hat{p}_A + n_B\hat{p}_B}{n_A + n_B} \text{ (correspond to } \bar{p} \text{ in E - 4)}$$

$$\hat{p}_A = \frac{21}{100} = 0.21, \hat{p}_B = \frac{30}{120} = 0.25, n_A = 100, n_B = 120$$

$$\hat{p}_o = \frac{n_A\hat{p}_A + n_B\hat{p}_B}{n_A + n_B} = \frac{100(0.21) + 120(0.25)}{100 + 120} = \frac{21 + 30}{220} = 0.2318 \text{ (common proportion)}$$

Now substitute these values into the formula to find the test statistic, rounding to three decimal places.

$$z = \frac{(0.21 - 0.25) - 0}{\sqrt{0.2318(1 - 0.2318)\left(\frac{1}{100} + \frac{1}{120}\right)}} = -0.700$$

$$\alpha = 5\% , z_{0.05} = 1.645 \text{ corresponding to } F(z) = 1 - 0.05 = 0.95$$

(Read from Appendix Table 1, p.815)

Compare the test statistic to $z_{0.05}$

$$If \ z < -z_{\alpha}, \text{Reject } H_0$$

$$-0.700 < -1.645 \rightarrow NO$$

Since the test statistic, -0.700, is not less than the critical value of -1.645, do not reject the null hypothesis. Conclusion: There is no difference between the population proportions of the companies offering high-deductible health insurance plans to their workers in 2006 than in 2005.

2.a) A scatter plot of $Y$ (GPA value) and $X$ (SAT score) data with the regression line fitted (solid line) is given as follows:
b) Use tabulated method to compute covariance and correlation coefficient as follows:

<table>
<thead>
<tr>
<th>i</th>
<th>$X_i$</th>
<th>$Y_i$</th>
<th>$(X_i - \bar{X})$</th>
<th>$(Y_i - \bar{Y})$</th>
<th>$(X_i - \bar{X})^2$</th>
<th>$(Y_i - \bar{Y})^2$</th>
<th>$(X_i - \bar{X})(Y_i - \bar{Y})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>640</td>
<td>2.5</td>
<td>640-606.5=33.5</td>
<td>2.5-3.02=-0.52</td>
<td>1122.25</td>
<td>0.2704</td>
<td>-17.42</td>
</tr>
<tr>
<td>2</td>
<td>700</td>
<td>3.2</td>
<td>700-606.5=93.5</td>
<td>3.2-3.02=0.18</td>
<td>8742.25</td>
<td>0.0324</td>
<td>26.18</td>
</tr>
<tr>
<td>3</td>
<td>550</td>
<td>3.5</td>
<td>550-606.5=-56.5</td>
<td>3.5-3.02=-0.48</td>
<td>3192.25</td>
<td>0.2304</td>
<td>12.65</td>
</tr>
<tr>
<td>4</td>
<td>540</td>
<td>2.8</td>
<td>540-606.5=-66.5</td>
<td>2.8-3.02=-0.22</td>
<td>4422.25</td>
<td>0.0484</td>
<td>-14.63</td>
</tr>
<tr>
<td>5</td>
<td>620</td>
<td>3.0</td>
<td>620-606.5=13.5</td>
<td>3.0-3.02=0.00</td>
<td>182.25</td>
<td>0.0004</td>
<td>-0.27</td>
</tr>
<tr>
<td>6</td>
<td>490</td>
<td>2.4</td>
<td>490-606.5=-116.5</td>
<td>2.4-3.02=-0.62</td>
<td>13572.25</td>
<td>0.3844</td>
<td>72.23</td>
</tr>
<tr>
<td>7</td>
<td>710</td>
<td>3.4</td>
<td>710-606.5=103.5</td>
<td>3.4-3.02=-0.38</td>
<td>10712.25</td>
<td>0.1444</td>
<td>39.33</td>
</tr>
<tr>
<td>8</td>
<td>600</td>
<td>2.9</td>
<td>600-606.5=-6.5</td>
<td>2.9-3.02=-0.12</td>
<td>42.25</td>
<td>0.0144</td>
<td>0.78</td>
</tr>
<tr>
<td>9</td>
<td>505</td>
<td>2.7</td>
<td>505-606.5=-101.5</td>
<td>2.7-3.02=-0.32</td>
<td>10302.25</td>
<td>0.1024</td>
<td>32.48</td>
</tr>
<tr>
<td>10</td>
<td>710</td>
<td>3.8</td>
<td>710-606.5=103.5</td>
<td>3.8-3.02=-0.78</td>
<td>10712.25</td>
<td>0.6084</td>
<td>80.73</td>
</tr>
<tr>
<td>$\Sigma$</td>
<td></td>
<td></td>
<td>$6065$</td>
<td>$30.2$</td>
<td>$63002.5$</td>
<td>$1.836$</td>
<td>$212.2$</td>
</tr>
</tbody>
</table>


\[
\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{6065}{10} = 606.5; \bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} = \frac{30.2}{10} = 3.02
\]

\[
S_{XY} = Cov(X,Y) = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{n-1} = \frac{212.2}{10-1} = 23.578
\]

\[
S_X = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}} = \sqrt{\frac{63002.5}{10-1}} = 83.668; S_X^2 = 7000.28
\]

\[
S_Y = \sqrt{\frac{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}{n-1}} = \sqrt{\frac{1.836}{10-1}} = \sqrt{0.204} = 0.4517
\]

\[
r_{XY} = \frac{S_{XY}}{S_X S_Y} = \frac{23.578}{(83.668)(0.4517)} = 0.6239
\]

c) \[
b_1 = \frac{Cov(X,Y)}{S_X^2} = \frac{23.578}{7000.28} = 0.0034
\]

Interpretation of $b_1$: The slope of the equation indicates the change in the average value of $Y$ when $X$ variable increases by 1 unit. For a 1 unit increase in the value of $X$ variable, we estimate that the value of $Y$ would increase by 0.0034 units.

\[
b_0 = \bar{Y} - b_1 \bar{X} = 3.02 - (0.0034)(606.5) = 0.9772
\]

Interpretation of $b_0$: If the value of $X$ were 0 unit, we would expect to have the value of $Y$ 0.9772 units. Interpret with caution—note that we are extrapolating the results beyond the observed data.

d) Prediction (Estimation) from the least-square linear correlation ($Value$ of $r_{XY}$ = 0.878 shows moderately good correlation of $Y$ with $X$ positively.

\[
\hat{Y}_{n+1} = b_0 + b_1 X_{n+1} = 0.9772 + (0.0034)(650) = 3.19 units
\]

3. Null and alternative hypothesis are as follows:

$H_0$: Stock price changes today are independent of previous day price changes.

$H_1$: Stock price changes today are not independent of previous day price changes.

Complete the contingency table (3x3 matrix) given to calculate expected counts ($E_{ij}$) from the corresponding formula below. These values were written under the observed counts ($O_{ij}$) as **boldface and italic** with the row and column total as follows:
\[ E_{ij} = \frac{R_i C_j}{n} \rightarrow E_{11} = \frac{R_1 C_1}{n} = \frac{(42)(36)}{100} = 15.12 \]
\[ E_{12} = \frac{R_1 C_2}{n} = \frac{(42)(38)}{100} = 15.96 \]
\[ E_{13} = \frac{R_1 C_3}{n} = \frac{(42)(26)}{100} = 10.92 = 42 - (15.12 + 15.96) \]
\[ E_{21} = \frac{R_2 C_1}{n} = \frac{(20)(36)}{100} = 7.2 \]
\[ E_{22} = \frac{R_2 C_2}{n} = \frac{(20)(38)}{100} = 7.6 \]
\[ E_{23} = \frac{R_2 C_3}{n} = \frac{(20)(26)}{100} = 5.2 = 20 - (7.2 + 7.6) \]
\[ E_{31} = \frac{R_3 C_1}{n} = \frac{(38)(36)}{1000} = 13.68 \]
\[ E_{32} = \frac{R_3 C_2}{n} = \frac{(38)(38)}{1000} = 14.44 \]
\[ E_{33} = \frac{R_3 C_3}{n} = \frac{(38)(26)}{1000} = 9.88 = 38 - (13.68 + 14.44) \]

<table>
<thead>
<tr>
<th>Price Change Today</th>
<th>Price Change Previous Day</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up</td>
<td>No Change</td>
<td>Down</td>
</tr>
<tr>
<td>14</td>
<td>15.12</td>
<td>16</td>
</tr>
<tr>
<td>15.12</td>
<td>15.96</td>
<td>10.92</td>
</tr>
<tr>
<td>No Change</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7.2</td>
<td>6</td>
</tr>
<tr>
<td>7.2</td>
<td>7.6</td>
<td>5.2</td>
</tr>
<tr>
<td>Down</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>13.68</td>
<td>14</td>
</tr>
<tr>
<td>14</td>
<td>14.44</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>9.88</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>38</td>
<td>26</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[
\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \left( \frac{O_{ij} - E_{ij}}{E_{ij}} \right)^2
\]
\[
= \frac{(14 - 15.12)^2}{15.12} + \frac{(16 - 15.96)^2}{15.96} + \frac{(12 - 10.92)^2}{10.92} + \frac{(6 - 7.2)^2}{7.2} + \frac{(8 - 7.6)^2}{7.6} + \frac{(6 - 5.2)^2}{5.2}
\]
\[
= 0.08296 + 0.0001 + 0.1068 + 0.2 + 0.02105 + 0.1231 + 0.3935 + 0.01341 + 0.3577 = 1.2986
\]

Decision rule:

\[
\text{reject } H_0 \text{ if } \chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \left( \frac{O_{ij} - E_{ij}}{E_{ij}} \right)^2 > \chi^2_{(r-1)(c-1), \alpha}
\]
\[ \chi^2 = 1.2986 > \chi^2_{(3-1)(3-1), 0.05} \]
\[ \chi^2_{4, 0.05} = 9.49 \text{ (Read from Appendix Table 7)} \]
Since the calculated chi-square value is less than the critical chi-square value (1.2986 < 9.49), we do not reject the null hypothesis and conclude that daily stock price changes are independent.