Q1) Define Normal distribution and describe its properties.

Q2) Sketch normal distribution with
   a) 90% and 95% respectively,
   b) variance 1, 5 and 7 respectively.

Q3) Explain any three of the following terms:
   a) Null and alternative Hypothesis
   b) Simple and Composite Hypothesis
   c) Simple and Multiple Regression
   d) Standard and t-Parameter
   e) t-Test

Q4) A manufacturer of car batteries claims that the life of the batteries has a standard deviation of 1.7 years. If a random sample of 12 of these batteries have a standard deviation of 2.7 years, do you think that a 0.05 level of significance.

Q5) Test and explain the important applications of $\chi^2$ statistic.

Q6) A small manager of XY Drive is concerned that his brand's plant may be broken down throughout the country. A survey in which the country was divided into four geographic regions. A random sampling of 100 consumers in each region was surveyed, with the following result:

<table>
<thead>
<tr>
<th>Region</th>
<th>Status of brand</th>
<th>Purchase the brand</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>S</td>
<td>40</td>
</tr>
<tr>
<td>S</td>
<td>B</td>
<td>35</td>
</tr>
<tr>
<td>S</td>
<td>NWF</td>
<td>5</td>
</tr>
<tr>
<td>S</td>
<td>Purshae</td>
<td>35</td>
</tr>
<tr>
<td>NWF</td>
<td>NWF</td>
<td>5</td>
</tr>
<tr>
<td>B</td>
<td>NWF</td>
<td>5</td>
</tr>
</tbody>
</table>

Using a 0.05 significance level, test whether brand, status and region are independent of each other.

Q7) (a) Define the following terms:
   i) Simple Linear Regression
   ii) Partial and Multiple Regression.
   b) A multiple regression was carried out to relate Y (allowance granted to a student), the variables $X_1$ (number of classes held) and $X_2$ (drying time). The data set consisted of 12 observations. The estimated coefficients are $\beta_0 = 100; \beta_1 = 1.5; \beta_2 = 10.5$. Calculate $R^2$-value for the regression using the formula $R^2 = 1 - SSE/SST$, where $SST = 173.5, SSE = 117.4$. Compute the coefficient of multiple determination.
KARACHI UNIVERSITY BUSINESS SCHOOL
UNIVERSITY OF KARACHI
FINAL EXAMINATION, JUNE 8-7, 2009: HIGHER COLLEGES
ADVANCED BUSINESS STATISTICS: BA (M) 601

MBA-III

Date: June 30, 2009.
Time Allowed: 3 Hours

Max. Marks: 60

Instructions: Attempt only Five (5) questions. All questions carry equal marks.

Q1 a) Define Normal distribution and describe its properties.
   b) Given a normal distribution with \( \mu = 40 \) and \( \sigma = 6 \), find the probability and shade the area
      i) below \( x \)
      ii) above \( 27 \)
      iii) between 42 and 51

Q2 a) Define simple random, stratified and systematic sampling schemes. Give at least two practical
   examples where the three sampling schemes are used.
   b) Fifty-eight of 2,000 randomly sampled corporations had their 2005 federal income tax returns
      audited. In another sample of 2,500 corporations, 61 had their 2004 returns audited. Was the
      fraction of corporate returns audited in 2005 significantly different from 2004 fractions? Test
      the appropriate hypothesis at \( \alpha = 0.05 \)?

Q3 a) Explain any three of the following terms:
   i) Null and Alternative Hypotheses
   ii) Simple and Composite Hypotheses
   iii) Statistical and Parameter
   iv) Type I and Type II errors
   b) One of the larger components of the cost of living is the amount spent on housing. Housing cost
      include rent (for tenants), mortgage payments (for home owners), heating, electricity and water. An
      economist undertook a five years study to determine how housing costs have changed. Five years ago
      he took a random sample of 200 households and recorded the percentage of total income spent on
      housing. This year he took a sample of 100 households. The statistics from the two samples are:
      \( x_1 = 32.42 \), \( x_2 = 6.08 \), \( x_3 = 33.72 \), \( x_4 = 6.75 \). Test the hypothesis at 5% level, to determine
      whether the economist can infer that housing cost as a percent of total income has increased
      over the last five years.

Q4 a) Two floppy manufacturer are competing for our business. To guarantee a ready supply of
   floppies, we desire to have any of our purchases in each firm. However, we, a preliminary
   check, want assurance that their products are not markedly different in quality. One of the
   primary measures of quality is the ease accessibility to the data saved in the floppies. Ten cases
   of floppies from each vendor are purchased, and each is tested to assess the performance of
   the respective floppies. The test has given \( x_1 = 1210 \), \( x_2 = 2550 \), \( x_3 = 1775 \), and \( x_4 = 5600 \).
   b) Can you conclude that the true means are equal at \( \alpha = 0.05 \)?
   c) Can you conclude that the true variances are also equal at \( \alpha = 0.05 \)?
   d) Construct a 95% confidence interval for part (i) and (ii).

Q5 a) What is meant by a statistical hypothesis? Explain the situation in which one sided and two
   sided tests are used.
   b) Twelve secretaries participated in an experiment to improve the typing speed. The following table
   shows the speed of 12 selected secretaries at the beginning of program (Before) and at the end of the
   program (After).

<table>
<thead>
<tr>
<th>Subject</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>261</td>
<td>254</td>
<td>254</td>
<td>266</td>
<td>228</td>
<td>235</td>
<td>236</td>
<td>235</td>
<td>240</td>
<td>247</td>
<td>268</td>
<td>201</td>
</tr>
<tr>
<td>After</td>
<td>260</td>
<td>236</td>
<td>236</td>
<td>256</td>
<td>224</td>
<td>216</td>
<td>216</td>
<td>206</td>
<td>207</td>
<td>247</td>
<td>210</td>
<td>209</td>
</tr>
</tbody>
</table>

Do the data provide sufficient evidence to conclude that the program is effective at 5% level of
significance?
Q6 a) Describe the three uses of Chi square Test.
b) A study of 800 graduates at a university was asked questions regarding their interest and attitudes. Some of these questions from scale called PEOPLE that measures interest in the welfare of others. Each student was thus classified as Low, Medium or High on this scale. Is there an association between PEOPLE score and Field of Study? Use 5% level of significance.

<table>
<thead>
<tr>
<th>Degree</th>
<th>Low</th>
<th>Medium</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arts</td>
<td>20</td>
<td>100</td>
<td>190</td>
</tr>
<tr>
<td>Commerce</td>
<td>20</td>
<td>120</td>
<td>70</td>
</tr>
<tr>
<td>Science</td>
<td>60</td>
<td>120</td>
<td>100</td>
</tr>
</tbody>
</table>

Q7 A book publisher has produced seven comparable text-books with the following costs:

<table>
<thead>
<tr>
<th>Quantity produced (000):</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturing costs (000)</td>
<td>5</td>
<td>5.5</td>
<td>6.5</td>
<td>7.5</td>
<td>8</td>
<td>9.5</td>
<td>10.8</td>
</tr>
</tbody>
</table>

a) Calculate the correlation coefficient for the association between quantity produced and manufacturing costs. Test it for significance at 5%.
b) Plot the data on a scatter diagram.
c) Calculate the regression line for predicting manufacturing costs from quantity produced, interpret its coefficients and add it to the scatter diagram.
d) Estimate how much of the variation in costs is accounted for by the model.
e) Predict the manufacturing cost of an eighth text-book which has an
Q.1 (a) Differentiate between Parameters and Statistic(s).
(b) The mean and variance of a normal random variable x are 50 and 25 respectively.
Find P (37.5 ≤ x ≤ 62.5)
(3) 
(9)

Q.2 (a) Define central limit theorem.
(b) What is the probability of drawing a random sample with a mean of 30 or more from a population with mean 28? The sample size is 100 and the population variance is 81.
(3) 
(9)

Q.3 (a) Find a 95% confidence interval for the mean of a normal population if a random sample of 16 values with mean 41.5 inches and standard deviation 9 inches is drawn from this population.
(b) Given two random samples of size n₁ = 9 and n₂ = 16 from two independent normal populations with
\[ \bar{x}_1 = 64; \bar{x}_2 = 52; s_1^2 = 36; \text{ and } s_2^2 = 25 \]
Find a 95% confidence interval for (μ₁ - μ₂). Assume σ₁ = σ₂.
(3) 
(9)

Q.4 (a) Define Type - I and Type - II errors.
(b) An ambulance service claims that it takes it on the average not more than 10 minutes to reach its destination in emergency calls. To test this claim, the time taken to reach the destination on 50 randomly chosen emergency calls were observed. The mean and standard deviation of the sample were computed as \[ \bar{x} = 11.2 \text{ minutes and } s = 1.8 \text{ minutes.} \] At 0.05 level of significance, does this constitute evidence that the figure claimed is too low?
(3) 
(9)

Q.5 (a) Define Pooled Variance.
(b) Two machines are used to cut steel bars of equal length. A random sample of 50 bars, cut on machine 1, gives mean of 55.6 inches with a standard deviation of 0.10 inches. Another random sample of 100 bars, cut on machine 2, gives a mean of 55.9 inches with standard deviation of 0.12 inches. At the 0.05 level of significance, are the machines cutting bars of equal length?
(3) 
(9)
Q.6 (a) Define Null and Alternate hypotheses.

(b) A market researcher believes that in a certain population the proportion of persons preferring brands A, B, C and D of tooth paste are 0.30, 0.60, 0.08 and 0.02 respectively. A simple random sample of 600 persons' draws from the population shows the following preferences:

<table>
<thead>
<tr>
<th>Brand</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of persons</td>
<td>192</td>
<td>342</td>
<td>48</td>
<td>22</td>
</tr>
</tbody>
</table>

Do these data provide sufficient evidence to reject the researcher's belief at α = 0.01?

Q.7 In terms of Yule's notation if $r_{12} = 0.8$, $r_{13} = 0.4$, $r_{14} = 0.56$ then find the value of (i) $r_{123}$ (ii) $R_{14}$

Q.8 (a) What is a time series? Describe its components.

(b) Determine a least square equation of trend line fitted to the following data. Estimate the value for the year 2010.

<table>
<thead>
<tr>
<th>Years</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1999</td>
<td>72</td>
</tr>
<tr>
<td>2000</td>
<td>41</td>
</tr>
<tr>
<td>2001</td>
<td>43</td>
</tr>
<tr>
<td>2002</td>
<td>55</td>
</tr>
<tr>
<td>2003</td>
<td>56</td>
</tr>
<tr>
<td>2004</td>
<td>82</td>
</tr>
<tr>
<td>2005</td>
<td>76</td>
</tr>
<tr>
<td>2006</td>
<td>79</td>
</tr>
<tr>
<td>2007</td>
<td>88</td>
</tr>
</tbody>
</table>
KARACHI UNIVERSITY BUSINESS SCHOOL
UNIVERSITY OF KARACHI
FINA] EXAMINATION, JUNE - 2008: AFFILIATED COLLEGES
Advanced Business Statistics: BA (P) - 412
BBA - IV

Time: 3 Hours
Max. Marks: 60

Instructions: (1) Attempt any five (5) questions.
(2) Use of scientific calculator is allowed.

Q.No. 1 (a) Define Random Variable. Differentiating between discrete and continuous random variables.
(b) It is known that 10% of the production of a manufacturer is defective. A random sample of 50 items is selected from the production. The number of defective items in the sample is binomial random variable x. Find (i) P (x ≤ 5) (ii) E(x) (iii) V(x)

Q.No. 2 (a) Define (i) Sampling distribution of sample mean and
(ii) Central limit theorem.
(b) A random sample of size 64 is drawn from a normal population with mean 10 and standard deviation 3. What is the probability that the mean of the sample is at least 117

Q.No. 3 (a) In Karachi city a simple random sample of 500 shoppers revealed that 250 prefer reduction rates when purchasing clothing. Construct 95% confidence interval for the true proportion of all shoppers who prefer reduction rates when purchasing clothing.
(b) Given two random samples of size n1 = 9 and n2 = 16 from two independent normal populations with
\[ \bar{x}_1 = 64, \bar{x}_2 = 52, s_1^2 = 36, s_2^2 = 25 \]
Find 95% confidence interval for \( \mu_1 - \mu_2 \). Assume \( \sigma_1 = \sigma_2 \)

Q.No. 4 (a) Define the following terms:
(i) Null and Alternative Hypothesis
(ii) Type-I and Type-II error
(b) An ambulance service claims that it takes it on the average not more than 10 minutes to reach its destination in emergency calls. To test this claim, the time taken to reach the destination of 50 randomly chosen emergency calls were observed. The mean and standard deviation of the sample were computed as \( \bar{x} = 11.2 \) minutes and \( S = 1.8 \) minutes. At 0.05 level of significance, does this constitute evidence that the figure claimed is too low?

Q.No. 5 (a) A market researcher believes that in a certain population the proportion of persons preferring brand A, B, C and D of toothpaste are 0.30, 0.60, 0.08 and 0.02 respectively. A simple random sample of 500 persons drawn from the population shows the following preferences:

<table>
<thead>
<tr>
<th>Brand</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of persons</td>
<td>192</td>
<td>342</td>
<td>44</td>
<td>22</td>
</tr>
</tbody>
</table>

Do these data provide sufficient evidence to reject the researcher's belief at \( \alpha = 0.01 \).

PLEASE TURN OVER
4.a) It is generally believed that stature of children depend on the stature of their parent. To test this proposition a survey of 200 children revealed the following information.

<table>
<thead>
<tr>
<th>Children's Stature</th>
<th>Tall</th>
<th>Moderate</th>
<th>Short</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>48</td>
<td>40</td>
<td>22</td>
</tr>
<tr>
<td>²</td>
<td>30</td>
<td>24</td>
<td>36</td>
</tr>
</tbody>
</table>

Formulate and test the relevant hypothesis using Chi squares at 1%: [08]

b) A sample of 16 match boxes found to contain 53 match sticks on the average with a standard deviation of 6. Construct a 99% confidence interval estimate for the mean number of sticks. [04]

5. A regression for 15 observations is estimated as under [12]

\[ \hat{Y} = 2.691 - 0.4795 X_1 \]
\[ (0.75) \quad (0.134) \]

Where \( Y \) = Coffee consumption in the US (cups/person/day) 
\( X_1 \) = Retail price of coffee (S/pound) in period t

Figures in parentheses are the standard errors of coefficients. Construct 95% confidence intervals for the coefficient (slope) of regression equation and test its significance.

6. Quarterly sales receipts of a poultry farmer for the second quarter of 2002 through 2nd quarter of 2005 are given in million rupees. Fit a linear trend using least squares and estimate the seasonal component. [12]

Sales: 120 110 115 125 125 125 110 135 130 125 120 140 135

7. Write short notes on any three of the following: [12]

(i) Difference between parameter and statistic.
(ii) Component of a time series.
(iii) Coefficients of correlation and the determination.
(iv) Poisson distribution.

G O O D   L U C K
1. a) Use relevant tables to read determine the values of the following. [8]
   i) $z_{0.05}$  
   ii) $z_{0.05}$  
   iii) $t_{0.05}$  
   iv) $\chi^2_{10.05} < \chi^2 < \chi^2_{20.05}$

b) In a certain 30 minutes interval during peak hours 255 vehicles on the 
   average with a standard deviation of 29 vehicles passed through an 
   underpass. In the next 30 minutes period what is the probability that, [10]
   (i) Less than 300 vehicles will pass through the underpass.
   (ii) More than 350 vehicles will pass through the underpass.
   (iii) Between 200 to 330 vehicles will pass through the underpass.

c) Pakistan and West India are going to play a series of ODI matches. The 
   team that wins three matches wins the series (Assuming no draws). If 
   Pakistan’s probability of winning a single game is 0.65, what is the probability that Pakistan wins the series in the (i) 3rd Game (ii) 
   5th Game. [8]

2. A sample of 500 men and 600 women from degree colleges of Karachi, 
   364 men and 334 women are found married in the staff. [10]
   (i) Test at 1% whether proportion of married men exceeds the 
       proportion of married women.
   (ii) Construct a 99% confidence interval for the difference in the 
       proportion of married men and women.

3. A sample of 20 Lux-buls found to weight 130 grams on the average with a 
   standard deviation of 15 gram, another sample of 22 Rexona-buls found to 
   weigh 140 grams on the average with a standard deviation of 10 grams. [10]
   a) Assuming equal population variance, test at 1% whether Lux-buls 
       weight less than Rexona on the average.
   b) Obtain a 99% confidence interval for the difference between 
       means.
1) Determine the probability of type I error.

2) Calculate the confidence interval for each of the given values of \( \alpha \): 0.05, 0.10, 0.01, 0.001, and 0.0001.

3(a) In a parking plant, a machine packs cartons with tires. A supervisor claims that the machine she is supervising will pack faster than a test that claims, the time it takes each machine to pack the same 48 cartons is recorded. The results, in seconds, are shown in the following table:

<table>
<thead>
<tr>
<th>New machine (s)</th>
<th>43.7</th>
<th>43.8</th>
<th>43.9</th>
<th>44.0</th>
<th>44.1</th>
<th>44.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent machine (s)</td>
<td>41.5</td>
<td>41.6</td>
<td>41.7</td>
<td>41.8</td>
<td>41.9</td>
<td>42.0</td>
</tr>
</tbody>
</table>

Do the data provide sufficient evidence to conclude that, on the average, the new machine packs faster? Perform the required hypothesis test at the 0.05 level of significance.

3(b) According to the Arizona Real Estate Commission, in 1992 only 1.6% of people holding real estate licenses were active in the industry. An independent agency has been asked by the commission to determine whether the percentage of active licenses has increased. A random sample of 100 licensed people reveals that 24 are currently active. Perform the appropriate hypothesis test at the 0.05 level of significance.

b) Suppose we tell you that the percentage of all adult U.S. males who are married exceeds the percentage of all adults U.S. females who are married. Further suppose that to check this claim, you randomly select 500 adult U.S. males and 500 adults U.S. females. You find that 0.62 of the males and 0.61 of the females are married. Does the data provide sufficient evidence, at the 0.05 level of significance, to support the claim? Explain your answer.

An economist is interested in the relationship between the disposable income of a family and the amount of money spent annually on food. For a particular study, the economist takes a random sample of eight middle-income families of the same race (father, mother, two children). The results are shown below, where x indicates disposable income, in thousands of dollars, and y indicates expenditure, in hundreds of dollars.

<table>
<thead>
<tr>
<th>x (Disposable Income)</th>
<th>y (Expenditure)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>20</td>
</tr>
<tr>
<td>45</td>
<td>25</td>
</tr>
<tr>
<td>50</td>
<td>30</td>
</tr>
<tr>
<td>40</td>
<td>25</td>
</tr>
<tr>
<td>41</td>
<td>26</td>
</tr>
<tr>
<td>42</td>
<td>27</td>
</tr>
<tr>
<td>43</td>
<td>28</td>
</tr>
</tbody>
</table>

a) What is the regression equation for the data?

b) Graph the regression equation and the data points.

c) What does the slope of the regression line represent in terms of disposable income and annual food expenditure?

d) Use the regression equation to predict the annual food expenditure of families with a disposable income of $25,000.

e) Test at the 0.05 level of significance, the hypothesis that the correlation coefficient is zero. Use the regression equation provided for this purpose.

b) Do the data provide sufficient evidence to conclude that the slope of the population regression line is not equal to zero? Use the p-value to test this hypothesis.
A random survey research found that 34% of respondents claimed they had purchased their last car on Black Friday. What is the 95% confidence interval for the proportion of consumers who purchase cars on Black Friday?

\[ \hat{p} = 0.34, n = \text{sample size} \]

\[ \text{Margin of Error} = z_{0.025} \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \]

\[ \text{Confidence Interval} = \hat{p} \pm \text{Margin of Error} \]

\[ \text{Confidence Interval} = 0.34 \pm \text{Margin of Error} \]

\[ 0.34 \pm 0.06 \]

\[ 0.28 \text{ to } 0.40 \]

At the 95% confidence level, the data supports the claim that 34% of consumers purchase cars on Black Friday.
(a) Differentiate between correlation and regression. Give an example.

(b) In planning for an orientation meeting with new business majors, the Administration of the Business School wants to emphasize the importance of doing well in the major courses in order to get better-paying jobs after graduation. To support this point, the chairman plans to show that there is a strong positive correlation between starting salaries for recent graduates from the Department of Business Administration and their grade-point averages in the major courses. Record for seven of last year’s graduates are selected at random and are given in the following Table.

<table>
<thead>
<tr>
<th>Major course (s)</th>
<th>Grade-Point averages in</th>
<th>starting salary (y in thousands of rupees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.58</td>
<td>35.0</td>
<td></td>
</tr>
<tr>
<td>3.27</td>
<td>18.0</td>
<td></td>
</tr>
<tr>
<td>3.83</td>
<td>44.0</td>
<td></td>
</tr>
<tr>
<td>3.50</td>
<td>41.0</td>
<td></td>
</tr>
<tr>
<td>3.33</td>
<td>39.0</td>
<td></td>
</tr>
<tr>
<td>2.50</td>
<td>46.0</td>
<td></td>
</tr>
<tr>
<td>4.00</td>
<td>50.0</td>
<td></td>
</tr>
</tbody>
</table>

a) Plot the data.
b) Find the values of correlation coefficient (r) and interpret the result.
c) Find the least square regression equation.
d) Check whether or not the fit is good.
e) What is the mean starting salary for graduates with grade point averages equal to 3.0?

9. (a) Explain “Seasonal Variations.” Give examples where we find seasonal variations and why? What are the methods of Isolating Seasonal Variations?

(b) The US-Car has recorded sales (in $1,000s) over the last two years of

<table>
<thead>
<tr>
<th>Month 2004</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>Aug</th>
<th>Sept</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>18.3</td>
<td>16.2</td>
<td>19.7</td>
<td>20.3</td>
<td>21.5</td>
<td>22.0</td>
<td>24.7</td>
<td>23.9</td>
<td>26.3</td>
<td>22.0</td>
<td>18.0</td>
<td>19.3</td>
</tr>
</tbody>
</table>

i) Plot the data. Does there appear to be any trend in the data? Any Cyclical or Any Seasonal variation?

ii) Calculate the seasonal indices.

iii) What are the deseasonalised values? How would you interpret them?
Instructions:  
1) Please return the question paper along with your answer script.  
2) Attempt only Five (5) questions. All questions carry equal marks.  
3) You must write your name, Institute Name and Seat No above.

Q1 a) Define Normal distribution and describe its properties.  
b) Given a normal distribution with μ=40 and σ=6, find the probability and shade the area  
i) below 32  
ii) above 27  
iii) between 42 and 51

Q2 a) What is meant by a statistical hypothesis? What are the two types of errors of decision that arise in testing a hypothesis? Briefly explain how a statistical hypothesis is tested.  
b) The financial manager of a large department-store chain selected a random sample of 200 of its credit card customers and found that 136 had incurred an interest charge during the previous year because of an unpaid balance. Compute a 90% confidence interval for the true proportion of credit card customers who incurred an interest charge during the previous year.

Q3 a) Let X₁, ..., Xₙ be a random sample from N(μ₁, σ₁²) and Y₁, ..., Yₙ be a random sample from N(μ₂, σ₂²). Assume that the two samples are independent of each other. Derive an expression for μ₁ - μ₂ at the (1-α)100% level of confidence.  
b) Fifty-eight of 2,000 randomly sampled corporations had their 2005 federal income tax returns audited. In another sample of 2,500 corporations, 61 had their 2004 returns audited. Was the fraction of corporate returns audited in 2005 significantly different from 2004 fractions? Test the appropriate hypothesis at α = 0.05?

Q4 A manufacturer suspects a difference in the equality of spare parts he receives from two suppliers. He obtains the following data on the service life list of parts from two suppliers.  

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Mean</th>
<th>Sample size</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>50</td>
<td>150</td>
<td>10</td>
</tr>
<tr>
<td>B</td>
<td>100</td>
<td>153</td>
<td>5</td>
</tr>
</tbody>
</table>

Test whether the difference between two sample means is statistically significant at 0.01 L.O.S.

Q5a) Fuel costs are important to profitability in the airline business. A small regional carrier has been operating three types of aircraft and has collected the following cost data from its 14 planes, expressed as fuel costs (in ten rupees) per available seat mile:  

<table>
<thead>
<tr>
<th>Type</th>
<th>7.3</th>
<th>8.3</th>
<th>7.6</th>
<th>6.8</th>
<th>8.0</th>
<th>8.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type B</td>
<td>7.9</td>
<td>9.5</td>
<td>8.7</td>
<td>8.3</td>
<td>9.6</td>
<td>8.4</td>
</tr>
</tbody>
</table>

a) At the 0.01 level of significance, can we conclude that there is no true difference between plane types in fuel costs?  
b) Construct a 95% confidence interval for the true difference of the two means.

Q6 a) Explain any Three of the following terms:  
i) Null and alternative hypothesis  
ii) Simple and composite hypothesis  
iii) Statistical parameter  
iv) One tail and two tailed test  
b) A new filtering device is installed in a chemical unit. Before its installation, a random sample yielded the following information about the percentage of impurity: Y₁ = 12.5, s₁² = 101.17, and n₁ = 8. After installation, a random sample yielded: Y₂ = 10.2, and s₂² = 94.73.  
c) Can you conclude that the true variances are equal at α = 0.10?  
d) Construct a 95% confidence interval for the mean of the true variances.

Q7 a) Describe the three uses of Chi square test.  
b) In an experiment to study the dependence of hypertension on smoking habits, the following data were taken on 180 individuals:  

<table>
<thead>
<tr>
<th>Smoking status</th>
<th>Non-smokers</th>
<th>Moderate smokers</th>
<th>Heavy smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td>21</td>
<td>36</td>
<td>30</td>
</tr>
<tr>
<td>No hypertension</td>
<td>41</td>
<td>26</td>
<td>19</td>
</tr>
</tbody>
</table>

Test the hypothesis that the presence or absence of hypertension is independent of smoking habits. Use a 0.05 level of significance.
2) Differentiate between correlation and regression? Give an example.

b) In planning for an orientation meeting with new business measures, the Administration of the Business School wants to emphasize the importance of doing well in the major courses in order to get better-paying jobs after graduation. To support this point, the chairman plans to show that there is a strong positive correlation between starting salaries for recent graduates from the Department of Business Administration and their grade-point averages in the major courses. Record for seven of last year’s graduates are selected at random and are given in the following Table

<table>
<thead>
<tr>
<th>Grade-Point averages in Major course (x)</th>
<th>starting salary (y in thousands of rupees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.50</td>
<td>35.0</td>
</tr>
<tr>
<td>3.85</td>
<td>38.0</td>
</tr>
<tr>
<td>3.35</td>
<td>40.0</td>
</tr>
<tr>
<td>3.25</td>
<td>39.0</td>
</tr>
<tr>
<td>4.00</td>
<td>46.0</td>
</tr>
<tr>
<td>3.50</td>
<td>50.0</td>
</tr>
</tbody>
</table>

a) Plot the data.
b) Find the values of correlation coefficient (r) and interpret the result.
c) Find the least square regression equation.
d) Check whether or not the fit is good.
e) What is the mean starting salary for graduates with grade point averages equal to 3.0?

9. a) Explain “Seasonal Variations” Give examples where we find seasonal variations and why? What are the methods of Isolating Seasonal Variations?
b) The US-Cars has recorded sales (in $1,000s) over the last two years of

<table>
<thead>
<tr>
<th>Month 2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>18.7</td>
</tr>
<tr>
<td>Feb</td>
<td>19.3</td>
</tr>
<tr>
<td>March</td>
<td>20.2</td>
</tr>
<tr>
<td>April</td>
<td>21.7</td>
</tr>
<tr>
<td>May</td>
<td>22.4</td>
</tr>
<tr>
<td>June</td>
<td>24.2</td>
</tr>
<tr>
<td>July</td>
<td>24.2</td>
</tr>
<tr>
<td>Aug</td>
<td>25.7</td>
</tr>
<tr>
<td>Sep</td>
<td>21.3</td>
</tr>
<tr>
<td>Oct</td>
<td>19.3</td>
</tr>
<tr>
<td>Nov</td>
<td>22.7</td>
</tr>
<tr>
<td>Dec</td>
<td>19.3</td>
</tr>
</tbody>
</table>

i) Plot the data. Does there appear to be any trend in the data? Any Cyclical or Any Seasonal variation?

ii) Calculate the seasonal indices.

iii) What are the deseasonalised values? How would you interpret them?
Q1 a) What is a Binomial distribution and in which situation it is appropriate to use it?
b) Test for impurities commonly found in drinking water from private wells showed that 70% of all wells in a particular country have impurity A. If a random sample has five wells from the country, what is the probability that:
i) Exactly two will have impurity A? ii) At most two? iii) Fewer than two?
c) A restaurant prepares a tossed salad containing the average 6 vegetables. Find the probability that the salad contains more than 5 vegetables
i) on a given day ii) or 3 of the next 4 days,

Q2a) Define Normal distribution and describe its properties.
b) Given a normal distribution with μ=40 and σ² = 36, find the probability and shade the area
i) be ow 32 ii) above 27 iii) between 47 and 51

Q3 a) What is meant by a statistical hypothesis? Explain the situation in which one sided and two sided tests are used.
b) A new filtering device is installed in a chemical unit. Before its installation, a random sample yielded the following information about the percentage of impurity: \( \bar{x}_1 = 12.5, s_1^2 = 101.17 \) and \( n_1 = 8 \). After installation, a random sample yielded: \( n_2 = 9, \bar{x}_2 = 10.2, \) and \( s_2^2 = 94.73 \). Construct a 95% confidence interval for the difference of the true means.

Q4 A manufacturer suspects a difference in the equality of spare parts he receives from two suppliers. He obtains the following data on the service life of parts from two suppliers.

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Mean</th>
<th>Sample size</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>56</td>
<td>28</td>
<td>25</td>
</tr>
<tr>
<td>B</td>
<td>100</td>
<td>15</td>
<td>49</td>
</tr>
</tbody>
</table>

Test whether the difference between two sample means is statistically significant at \( \alpha = 0.05 \). Also construct a 95% confidence interval.

Q5a) The Murad Perfume Company recently developed a new fragrance that they plan to market under the name "Heavenly." A number of market studies indicate that Heavenly has very good market potential. The sales department at Murad's is particularly interested in whether there is a difference in the proportions of younger and older women who could purchase Heavenly if it were marketed. There are two independent populations, a population consisting of the younger women and a population consisting of older women. Each sampled woman will be asked to smell Heavenly and indicate whether she likes the fragrance well enough to purchase a bottle. A random sample of 100 young women revealed 60 liked the Heavenly fragrance well enough to purchase it. Similarly, a sample of 200 older women revealed 100 liked the fragrance well enough to make a purchase. Test at 5% level of significance.

Q6 a) Explain any three of the following terms:
i) Null and alternative Hypotheses  ii) Simple and Composite Hypotheses
iii) Statistic and Parameter iv) One tailed and two tailed test
b) A manufacturer of car batteries claims that the life of his batteries has a standard deviation equal to 0.9 year. If a random sample of 10 of these batteries have a standard deviation of 1.2 years, do you think that \( \sigma > 0.9 \) year? Use a 0.05 level of significance.
Q7 Low-involvement consumer purchases of items such as floor wax, dishwashing detergent, and ballpens are a challenge to the creativity of marketers trying to generate enthusiasm for their brands. On the other hand, consumers may pay little attention to advertising for their products; on the other hand, the brand they choose is often the first brand that comes to mind in connection with that product category. Thus a brand’s share of the market in its product class is often directly related to its “share of mind,” the consumer’s degree of brands awareness. The data in the table represent market share, Y, and brand awareness index value, X, for 5 low-involvement consumer brands.

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>25</td>
</tr>
<tr>
<td>4.0</td>
<td>40</td>
</tr>
<tr>
<td>3.6</td>
<td>35</td>
</tr>
<tr>
<td>7.5</td>
<td>60</td>
</tr>
<tr>
<td>8.2</td>
<td>70</td>
</tr>
</tbody>
</table>

a) Identify the predictor and response variable.
b) Find the regression equation for the above data by the method of least square.
c) Use the regression equation to predict the market share when the brand awareness index value is 6.0.
d) What is the coefficient of Correlation and Coefficient of Determination calculate for the above data? Interpret your result.
KARACHI UNIVERSITY BUSINESS SCHOOL
UNIVERSITY OF KARACHI

FINAL EXAMINATION DECEMBER-2007: AFFILIATED COLLEGES
ADVANCED BUSINESS STATISTICS BA(P)-412
BBA-IV

Time Allowed: 3 Hours
Dated: 26/12/2007
Max: Marks: 60

INSTRUCTIONS: Attempt only five (5) questions. Questions carry equal marks.

1.a) Use relevant tables to read/ determine the values of the following.
   i) \( (20.24)_{(9.9)} \)
   ii) \( 32.375 \)
   iii) \( 32.85 \)
   iv) \( P(14, 0.95 < \chi^2 < 15, 0.98) = \) __________

b) In a certain 30 minutes interval during peak hours 255 vehicles on the average with a standard deviation of 50 vehicles passed through an underpass. In the next 30 minutes period, what is the probability that:
   (i) Less than 300 vehicles will pass through the underpass.
   (ii) More than 350 vehicles will pass through the underpass.
   (iii) Between 200 to 330 vehicles will pass through the underpass.

   

2. A sample of 500 men and 600 women from degree colleges of Karachi, 364 men and 354 women are found married in the staff.
   (i) Test at 1% whether proportion of married men exceeds the proportion of married women.
   (ii) Construct a 99% confidence interval for the difference in the proportion of married men and women.

3. A sample of 20 Lux-bars found to weight 130 grams on the average with a standard deviation of 15 gram, another sample of 22 Rexona-bars found to weight 140 grams on the average with a standard deviation of 10 grams.
   a) Assuming equal population variance, test at 1% whether Lux-bars weight less than Rexona on the average.
   b) Obtain a 99% confidence interval for the difference between means.
It is generally believed that stature of children depends on the statures of their parent. To test this proposition a survey of 200 children revealed the following information:

<table>
<thead>
<tr>
<th>Children's Stature</th>
<th>Tall</th>
<th>Moderate</th>
<th>Short</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tall</td>
<td>48</td>
<td>40</td>
<td>22</td>
</tr>
<tr>
<td>Short</td>
<td>30</td>
<td>24</td>
<td>36</td>
</tr>
</tbody>
</table>

Formulate and test the relevant hypothesis using Chi squares at 1%.

b) A sample of 16 match boxes found to contain 53 match sticks on the average with a standard deviation of 6. Construct a 99% confidence interval estimate for the mean number of sticks.

5. A regression for 15 observations is estimated as under

\[ \hat{Y} = 2.691 - 0.4795 X_c \]

\[ (0.73) \quad (0.134) \]

Where \( Y \) = Coffee consumption in the US (cups/person/day)
\( X_c \) = retail price of coffee ($/pound) in period 1.

Figures in parentheses are the standard errors of coefficients. Construct 95% confidence intervals for the coefficient (slope) of regression equation and test its significance.

6. Quarterly sales receipts of a poultry farmer for the second quarter of 2002 through 2005 are given in million rupees. Fit a linear trend using least squares and estimate the seasonal component.

Sales: 120 110 115 125 120 110 135 130 125 120 140 135

7. Write short notes on any three of the following:

- Difference between parameter and statistic.
- Components of a time series.
- Coefficients of correlation and the determination.
- Poisson distribution.

Good luck!
Name ___________________________ Regular/Repeater: ___________ Seat No: ___________

KARACHI UNIVERSITY Business School, Affiliated Colleges
BBA IV: BA(P) 412 Advanced Business Statistics
Dated, 12th December 2006 Max Time: Three Hours

Instructions:
1) Please return the question paper along with your answer script.
2) Attempt only Five (5) questions. All questions carry equal marks.
3) Write your name, Institute Name and Seat No above.

Q1 a) What is Binomial distribution and in which situation it is appropriate to use?

b) Test for impurities commonly found in drinking water from private wells showed that 70% of all wells in a particular country have impurity A. If a random sample has five wells from the country, what is the probability that:
   i) Exactly two will have impurity A?
   ii) At most two?
   iii) Fewer than two?

c) A restaurant prepares a tossed salad containing on the average 6 vegetables. Find the probability that the salad contains more than 5 vegetables
   i) on a given day
   ii) on 3 of the next 4 days;

Q2a) Define: Normal distribution and describe its properties.

b) Given a normal distribution with \( \mu = 48 \) and \( \sigma = 36 \), find the probability and shade the area
   i) below 32
   ii) above 72
   iii) between 42 and 51

Q3 a) What is meant by a statistical hypothesis? Explain the situation in which one sided and two sided tests are used.

b) A new filtering device is installed in a chemical unit. Before its installation, a random sample yielded the following information about the percentage of impurity: \( \bar{y} = 12.5 \), \( s_y^2 = 101.17 \), and \( n = 8 \). After installation, a random sample yielded: \( \bar{y} = 10.2 \), and \( s_y^2 = 94.73 \). Construct a 95% confidence interval for the difference of the two means.

Q4 A manufacturer suspects a difference in the equality of spare parts he receives from two suppliers. He obtains the following data on the service list of parts from two suppliers.

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Mean</th>
<th>Sample size</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>50</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>B</td>
<td>100</td>
<td>15</td>
<td>40</td>
</tr>
</tbody>
</table>

Test whether the difference between two sample means is statistically significant at \( \alpha = 0.05 \). Also construct a 95% confidence interval.

Q5a) The Murad Perfume Company recently developed a new fragrance that they plan to market under the name 'Heavenly.' A number of market studies indicate that Heavenly has very good market potential. The sales department at Murad is particularly interested in whether there is a difference in the proportion of younger and older women who would purchase Heavenly if it were marketed. There are two independent populations, a population consisting of the younger women and a population consisting of older women. Each sampled woman will be asked to smell Heavenly and indicate whether she likes the fragrance well enough to purchase a bottle. A random sample of 100 young women revealed 60 liked the Heavenly fragrance well enough to purchase it. Similarly, a sample of 200 older women revealed 100 liked the fragrance well enough to make a purchase. Test at 5% level of significance.

b) Construct a 95% confidence interval for the tree difference of the two proportions.

Q6 a) Explain any three of the following terms:
   i) Null and alternative Hypotheses
   ii) Simple and Composite Hypotheses
   iii) Statistical and Parameter

b) A manufacturer of car batteries claims that the life of his batteries has a mean deviation equal to 0.9 year. If a random sample of 10 of these batteries shows a mean deviation of 1.2 years, do you think that \( \sigma > 0.9 \) year? Use a 0.05 level of significance.
4.a) It is generally believed that stature of children depends on the stature of their parent. To test this proposition a survey of 200 children revealed the following information.

<table>
<thead>
<tr>
<th>Child's Stature</th>
<th>Parent's Stature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tall</td>
<td>48</td>
</tr>
<tr>
<td>Short</td>
<td>30</td>
</tr>
<tr>
<td>Tall</td>
<td>40</td>
</tr>
<tr>
<td>Short</td>
<td>24</td>
</tr>
</tbody>
</table>

Formulate and test the relevant hypothesis using Chi squares at 1%.

b) A sample of 16 match boxes found to contain 53 match sticks on the average with a standard deviation of 5. Construct a 99% confidence interval estimate for the mean number of sticks.

5. A regression for 15 observations is estimated as under

\[ Y = 2.691 - 0.4795 X, \]
\[ (0.75) \quad (0.134) \]

Where \( Y \) = Coffee consumption in the US (cup/person/day)
\( X \) = Retail price of $/pound in period t.

Figures in parentheses are the standard errors of coefficients.
Construct 95% confidence intervals for the coefficient (slope) of regression equation and test its significance.

6. Quarterly sales receipts of a poultry farmer for the second quarter of 2002 through 2nd Quarter of 2005 are given in million rupees. Fit a linear trend using least squares and estimate the seasonal component.

Sales: 120 110 115 125 120 110 135 130 125 120 140 135

7. Write short notes on any three of the following:

- (i) Difference between parameter and statistic.
- (ii) Components of a time series.
- (iii) Coefficients of correlation and the determination.
- (iv) Poisson distribution.

GOOD LUCK
Name: ____________________________ Regular/Repeater: ________________ Seat No: ________________

Korachi University Business School, Affiliated Colleges
BBA:IV, RA-II 412 Advanced Business Statistics

Date: 12th December 2006 Max Time: Three Hours

Instructions:
1) Please return the question paper along with your answer script.
2) Attempt only Five (5) questions. All questions carry equal marks
3) Write your own hand, Institute Name and Seat No above.

Q1 a) What is Binomial distribution and in which situation is it appropriate to use?
   b) Test for impurities commonly found in drinking water from private wells showed that 70% of all wells in a particular country have impurity A. If a random sample has five wells from the country, what is the probability that:
      i) Exactly two will have impurity A? ii) At most two? iii) Fewer than two?
   c) A restaurant prepares a tossed salad containing on the average 6 vegetables. Find the probability that the salad contains more than 5 vegetables
      i) on a given day ii) on 3 of the next 4 days;

Q2a) Define Normal distribution and describe its properties.
   b) Given a normal distribution with \( \mu = 40 \) and \( \sigma^2 = 36 \), find the probability and shade the area
      i) below 32 ii) above 27 iii) between 42 and 51

Q3 a) What is meant by a statistical hypothesis? Explain the situation in which one sided and two sided tests are used.
   b) A new filtering device is installed in a chemical unit. Before its installation, a random sample yielded the following information about the percentage of impurity: \( \bar{y} = 12.5, s_x^2 = 101.17 \), and \( n = 8 \). After installation, a random sample yielded: \( t_0 = 9, s_x^2 = 10.2 \), and \( s_x^2 = 94.73 \). Construct a 95% confidence interval for the difference of the true means.

Q4 A manufacturer suspects a difference in the equality of spurt parts he receives from two suppliers. He obtains the following data on the service list of parts from two suppliers.

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Mean</th>
<th>Sample Size</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10</td>
<td>28</td>
<td>15</td>
</tr>
<tr>
<td>B</td>
<td>20</td>
<td>15</td>
<td>49</td>
</tr>
</tbody>
</table>

Test whether the difference between two sample means is statistically significant at \( \alpha = 0.05 \). Also construct a 95% confidence interval.

Q5a) The Murad Perfumes Company recently developed a new fragrance that they plan to market under the name "Heavenly". A number of market studies indicate that Heavenly has very good market potential. The sales department at Murad is particularly interested in whether there is a difference in the proportion of younger and older women who could purchase Heavenly if it were marketed. These are two independent populations, a population consisting of the younger women and a population consisting of older women. Each sampled woman will be asked to smell Heavenly and indicate whether she likes the fragrance well enough to purchase it. A random sample of 100 young women revealed 20 liked the Heavenly fragrance well enough to purchase it. Similarly, a sample of 200 older women revealed 100 liked the fragrance well enough to make a purchase. Test at 5% level of significance.

b) Construct a 95% confidence interval for the true difference of the two proportions.

Q6 a) Explain any three of the following terms:
   i) Null and Alternative Hypotheses ii) Simple and Composite Hypotheses
   iii) Statistic and Parameter iv) One-tailed and two-tailed test

b) A manufacturer of car batteries claims that the life of his batteries has a standard deviation equal to 0.9 year. If a random sample of 10 of these batteries have a standard deviation of 1.2 years, do you think that \( \sqrt[3]{\sigma} = 0.9 \) year? Use a 0.05 level of significance.
Low-involvement consumer purchases of items such as floor wax, dishwashing detergent, and shampoo pans are a challenge to the creativity of marketers trying to generate enthusiasm for their brands. On the other hand, consumers may pay little attention to advertising for their products. On the other hand, the brand they choose is often the first brand that comes to mind in connection with that product category. Thus a brand's share of the market in its product class is often directly related to its "share of mind" - the consumer's degree of brand awareness. The data in the table represent market share, Y, and brand awareness index value, X, for 5 low-involvement consumer brands.

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>25</td>
</tr>
<tr>
<td>4.0</td>
<td>40</td>
</tr>
<tr>
<td>5.0</td>
<td>55</td>
</tr>
<tr>
<td>7.5</td>
<td>60</td>
</tr>
<tr>
<td>3.0</td>
<td>70</td>
</tr>
</tbody>
</table>

a) Identify the predictor and response variable.
b) Find the regression equation for the above data by the method of least squares.
c) Use the regression equation to predict the market share when the brand awareness index value is 6.0.
d) What is the coefficient of Correlation and Coefficient of Determination calculate for the above data? Interpret your result.
Instructions:
1) Please return the question paper along with your answer script.
2) Attempt only Five (5) questions. All questions carry equal marks.

Q1 a) What is Binomial distributions and in which situations it is appropriate to use?

b) Test for impurities commonly found in drinking water from private wells showed that 30% of all wells in a particular country have impurity A. If a random sample has five wells from the county, what is the probability that:
   i) Exactly two will have impurity A? ii) At most two? iii) Fewer than two?

c) A restaurant prepares a tossed salad containing on the average 5 vegetables. Find the probability that the salad contains more than 5 vegetables on a given day i) on 3 of the next 4 days;

Q2a) Define Normal distribution and describe its properties.

b) Given a normal distribution with μ=40 and σ=6, find the probabilities and shade the areas below 32 i) above 27 ii) between 42 and 51

Q3 a) What is meant by a statistical hypothesis? Explain the situation in which one sided and two sided tests are used.

b) Fifty-eight of 2,000 randomly sampled corporations had their 2005 federal income tax returns audited. In another sample of 2,500 corporations, 61 had their 2004 returns audited. Was the fraction of corporate returns audited in 2005 significantly different from 2004 fraction? Test the appropriate hypothesis at α = 0.05.

Q4 a) A manufacturer suspects a difference in the equality of spare parts he receives from two suppliers. He obtains the following data on the service list of parts from two suppliers.

<table>
<thead>
<tr>
<th>Supplier</th>
<th>Mean</th>
<th>Sample Size</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>20</td>
<td>100</td>
<td>150</td>
</tr>
<tr>
<td>B</td>
<td>50</td>
<td>135</td>
<td>10</td>
</tr>
</tbody>
</table>

Test whether the difference between two sample means is statistically significant at 0.01 L.O.S.

Q5a) Fuel costs are important to profitability in the airline business. A small regional carrier has been operating three types of aircraft and has collected the following cost data from its 14 planes. Express all fuel costs in ten rupees per available seat mile.

Type A: 7.3 8.3 7.6 6.8 9.0
Type B: 7.9 9.5 8.7 8.3 9.6 8.4

a) At the 0.01 level of significance, can we conclude that there is no true difference between plant types in fuel costs?

b) Construct a 95% confidence interval for the true difference of the two means.

Q6 a) Explain any three of the following terms:
   i) Null and Alternative Hypothesis
   ii) Simple and Composite Hypothesis
   iii) Statistical and Parameter

b) A new filtering device is installed in a chemical unit. Before its installation, a random sample yielded the following information about the percentage of dirt: Y=12.5, s^2=101.17, and n=10. After installation, a random sample yielded; n=10, Y=10.2, and s^2=94.71.

Can you conclude that the true variances are equal at α = 0.10?

Q7 a) Describe the three uses of Chi-square Test.

b) In an experiment to study the dependence of hypertension on smoking habits, the following data were taken on 180 individuals:

<table>
<thead>
<tr>
<th></th>
<th>Non-smokers</th>
<th>Moderate smokers</th>
<th>Heavy smokers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td>21</td>
<td>36</td>
<td>30</td>
</tr>
<tr>
<td>No hypertension</td>
<td>42</td>
<td>20</td>
<td>19</td>
</tr>
</tbody>
</table>

Test the hypothesis that the presence or absence of hypertension is independent of smoking habits. Use a 0.05 level of significance.
b) In planning for an orientation meeting with new business students, the Administration of the Business School wants to emphasize the importance of doing well in the major courses in order to get better-paying jobs after graduation. To support this point, the chairman plans to show that there is a strong positive correlation between starting salaries for recent graduates from the Department of Business Administration and their grade-point averages in the major courses. Record for seven of last year's graduates are selected at random and are given in the following Table.

<table>
<thead>
<tr>
<th>Major course (x)</th>
<th>Grade-Point averages in (y in thousands of pesos)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.58</td>
<td>55.0</td>
</tr>
<tr>
<td>2.27</td>
<td>38.9</td>
</tr>
<tr>
<td>3.85</td>
<td>44.0</td>
</tr>
<tr>
<td>3.50</td>
<td>41.0</td>
</tr>
<tr>
<td>3.35</td>
<td>39.0</td>
</tr>
<tr>
<td>3.50</td>
<td>46.6</td>
</tr>
<tr>
<td>4.00</td>
<td>50.0</td>
</tr>
</tbody>
</table>

a) Plot the data.
b) Find the values of correlation coefficient (r) and interpret the results.
c) Find the least square regression equation.
d) Check whether or not the fit is good.
e) What is the mean starting salary for graduates with grade-point averages equal to 3.0?

4. a) Explain "Seasonal Variations." Give examples where you find seasonal variations and why? What are the methods of isolating Seasonal Variations?
b) The US-Cars has recorded sales (in $'000s) over the last two years of

<table>
<thead>
<tr>
<th>Month</th>
<th>2004</th>
<th>2005</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>123</td>
<td>181</td>
</tr>
<tr>
<td>Feb</td>
<td>167</td>
<td>152</td>
</tr>
<tr>
<td>March</td>
<td>197</td>
<td>203</td>
</tr>
<tr>
<td>April</td>
<td>202</td>
<td>21.3</td>
</tr>
<tr>
<td>May</td>
<td>21.7</td>
<td>22.3</td>
</tr>
<tr>
<td>June</td>
<td>23.1</td>
<td>24.7</td>
</tr>
<tr>
<td>July</td>
<td>24.2</td>
<td>23.9</td>
</tr>
<tr>
<td>Aug</td>
<td>25.7</td>
<td>36.3</td>
</tr>
<tr>
<td>Sept</td>
<td>21.2</td>
<td>20.0</td>
</tr>
<tr>
<td>Oct</td>
<td>19.3</td>
<td>18.6</td>
</tr>
<tr>
<td>Nov</td>
<td>12.7</td>
<td>9.7</td>
</tr>
<tr>
<td>Dec</td>
<td>15.3</td>
<td>17.5</td>
</tr>
</tbody>
</table>

i) Plot the data. Does there appear to be any trend in the data? Any Cyclic or Any Seasonal variation?
ii) Calculate the seasonal indices.
iii) What are the deseasonalized values? How would you interpret them?
Q1: (a) Explain what is meant by the statement “We are 95% confident that our interval estimate contains μ." Explain the difference between an interval estimator and a point estimator for μ.

(i) When choosing a product to purchase, what do you consider most: price or quality? In a poll of 2000 American adults conducted by Roper Starch Worldwide, 64% claim they mainly base their buying decisions on price (Tampa Tribune, Oct. 31, 1993).

(ii) Construct a 99% confidence interval for the true percentage of American adults who base their buying decisions more on price than on quality.

(iii) Interpret the interval.

(iv) How would the width of the confidence interval, part (i), change if the confidence coefficient was decreased from .99 to .95?

Q2: (a) (i) What is the central limit theorem?

(ii) Does the central limit theorem specify the individual cases follow a normal distribution?

(iii) How do you interpret the idea that the average has a normal distribution?

(iv) What is the mean of a sum of independent observations of a random variable? What is the standard deviation?

(v) What is the mean of an average of independent observations of a random variable? What is the standard deviation?

(b) You have a factory with 40 production machines that are essentially identical, each producing at a mean daily rate of 100 products with a standard deviation of 15. You may assume that they produce independently of one another. Consider the average daily production per machine tomorrow, which is a random variable.

(i) Find the mean of his random variable. Compare it to the mean for a single machine.

(ii) Find the standard deviation of this random variable. Compare it to the standard deviation for a single machine?

(iii) What is the approximate probability distribution of this random variable? How do you know?

(iv) Find the (approximate) probability that your average daily production per machine will be more than 102 products tomorrow.

(v) Find the (approximate) probability that your average daily production per machine will be between 97 and 103 products tomorrow.

Q3: (a) The chlorine level of water emerging from the South Water treatment plant and at the Fire Station (Fenton Zone 13) was measured overall 12-month period, with the result shown in the accompanying table. Estimate the mean difference in monthly chlorine content between the two locations using a 90% confidence interval.

<table>
<thead>
<tr>
<th>Month</th>
<th>Jan</th>
<th>Feb</th>
<th>Mar</th>
<th>Apr</th>
<th>May</th>
<th>Jun</th>
<th>Jul</th>
<th>Aug</th>
<th>Sept</th>
<th>Oct</th>
<th>Nov</th>
<th>Dec</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Plant</td>
<td>2.0</td>
<td>2.0</td>
<td>2.1</td>
<td>1.9</td>
<td>1.7</td>
<td>1.8</td>
<td>1.7</td>
<td>1.9</td>
<td>2.0</td>
<td>2.0</td>
<td>2.1</td>
<td>2.2</td>
</tr>
<tr>
<td>Fire Station</td>
<td>2.2</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
<td>1.9</td>
<td>1.8</td>
<td>1.7</td>
<td>1.9</td>
<td>1.8</td>
<td>1.8</td>
<td>2.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

(b) Social Science Quarterly (Sept, 1993) reported on study of gender differences among workers in computer software industry. Questionnaires were administered to sample of 298 female and 268 males who were employed in full-time in software-related jobs. In the female sample, 89 had professional occupations (e.g. programmers, analysts, computer scientists) and 209 had nonprofessional jobs (computer and peripheral equipment operators). In contrast the male sample included 150 professionals and 114 nonprofessional. Use a 90% confidence interval to compare the proportions of male and female software workers who hold professional positions.
Q4: (a) Find the sample size needed to estimate a population mean \( \mu \) and cost \( T \) with 90% confidence if \( N = 2000, d = 6 \) and \( \sigma^2 = 256 \) for simple random sample.

(b) Find the sample size needed to estimate a population proportion with 90% confidence if \( N = 1000, d = .15 \) and \( \alpha = .04 \) for simple random sample.

Q5: (a) In what conditions would you prefer stratified random sampling over simple random sampling?

(b) A large firm knows that 500 accounts are wholesalers and 300 are retail. An auditor wishes to sample 100 of these accounts in order to estimate the average amount of accounts receivables for the firm. A simple random sample turn out to contain 70 wholesale accounts and 30 retail accounts. The following data (in dollars) is obtained.

<table>
<thead>
<tr>
<th>Wholesale</th>
<th>Retail</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_i = 750 )</td>
<td>( x_i = 30 )</td>
</tr>
<tr>
<td>( x_i = 520 )</td>
<td>( x_i = 286 )</td>
</tr>
<tr>
<td>( x_i = 210 )</td>
<td>( x_i = 90 )</td>
</tr>
</tbody>
</table>

Find an approximate 95% confidence interval for the population mean and population total.

Q6: In quality control applications of hypothesis testing, the null and alternative hypothesis are frequently specified as:

\( H_0: \) The production process is performing satisfactorily
\( H_1: \) The process is performing in an unsatisfactorily manner

Accordingly, \( \alpha \) is sometimes referred to producer's risk, while \( \beta \) is called the consumer's risk (Montgomery,1991). An injection molding produces plastic golf tees. The process is designed to produce tees with a mean weight of 0.255 ounce. To investigate whether the injection molders is operating satisfactorily, 40 tees are randomly sampled from the past hour's production. Their weights (in ounce) are listed below:

- 247
- 250
- 248
- 251
- 249
- 253
- 255
- 254

Do the data provide sufficient evidence to conclude that the process is not operating satisfactorily? Test using \( \alpha = 0.03 \)

Is the conclusion of this problem explain why it makes sense to call \( \alpha \) the producer's risk and \( \beta \) the consumer's risk.

Q7: (a) What \( p \)-value statement is associated with each of the following outcomes of a hypothesis test?

(i) Not significant.
(ii) Significant.
(iii) Highly significant.
(iv) Very highly significant.

A management team was asked to solve 10 different quality control problems, commonly encountered in their work. A second management team was asked to solve the same problems. Solution times in minutes required by each team is shown below. Calculate and interpret the 90% confidence interval for the difference between the population mean times required for the two teams. What can you conclude about the relative problem solving ability of the two teams?

<table>
<thead>
<tr>
<th>Problem</th>
<th>Team 1</th>
<th>Team 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>26</td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>21</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>19</td>
<td>31</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>25</td>
<td>35</td>
</tr>
<tr>
<td>8</td>
<td>18</td>
<td>28</td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td>27</td>
</tr>
<tr>
<td>10</td>
<td>20</td>
<td>26</td>
</tr>
</tbody>
</table>

Best of Luck