

**KARACHI UNIVERSITY BUSINESS SCHOOL**  
 University of Karachi  
 FINAL EXAMINATION, DECEMBER 2009; AFFILIATED COLLEGES  
**ADVANCE BUSINESS STATISTICS: BA (M) - 601**  
**MBA - III**

$$SP = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{(n_1 + n_2 - 2)}$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{SP \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

Date: January 12, 2010

Max Time: 3 Hours  
 Max Marks: 60

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

- Q1 a) Define Normal distribution and describe its properties.  
 b) Sketch normal distribution with  
 i) means -3, 0 and 5 respectively ii) variances 1, 5 and 7 respectively.

Q2 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) 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:  $\bar{x}_1 = 32.42$ ;  $s_1 = 6.08$ ;  $\bar{x}_2 = 33.72$ ;  $s_2 = 6.75$ . 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.

Handwritten notes:  
 $\mu_1 = -3, \mu_2 = 0, \mu_3 = 5$   
 $\sigma_1^2 = 1, \sigma_2^2 = 5, \sigma_3^2 = 7$   
 $H_0: \mu \leq \mu_0$   
 $H_a: \mu > \mu_0$

Q3 a) Distinguish between simple and stratified random sampling schemes. Give at least two practical examples where the two sampling schemes are used.

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.

Q4 Following are the data on rate of return for 10 randomly selected firm from Textile sector and 10 firms from Energy Sector.

Textile	15	11	10	14	18	08	05	16	07	15
Energy	15	13	13	10	16	16	12	14	09	19

- a) Can we conclude, at  $\alpha = 5\%$  that investment in the two sectors is equally risky by comparing variability of return.  
 b) Based on the result of part (i) test that the expected return is same for the both sectors.

Q5 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.

Handwritten notes:  
 $S^2 = \frac{n \sum x^2 - (\sum x)^2}{n(n-1)}$   
 $\hat{p} \pm Z_{\alpha/2} \sqrt{\hat{p}(1-\hat{p})}$   
 $E(\chi^2) = np \cdot \chi^2 = \frac{(n-1)S^2}{\sigma^2}$

Q6 a) State and explain three important applications of  $\chi^2$  statistic.

b) A brand manager of CD Drive is concerned that her brand's share may be evenly distributed throughout the country. In 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 results:

	Region	S	P	B	NWF
Status of brand					
Purchase the brand		40	55	45	50
Do not purchase		60	45	55	50

Handwritten notes:  
 $\frac{(o-e)^2}{e}$   
 $E(\chi^2) = E(X^2) - [E(X)]^2$

Using a 5% significance level, test whether brand status and region are independent or dependent.

Q. 7(a) Define the following terms:

- (i) Simple linear Regression  
 (ii) Partial and Multiple Regression.
- (b) A multiple regression was carried out to relate  $Y$ =tensile strength to a synthetic fiber specimen, the variable  $X_1$ =percent cotton and  $X_2$ =drying time. The data set consisted of  $n=12$  observations. The estimated coefficients are  $\hat{\beta}_0 = 180, \hat{\beta}_1 = 1, \hat{\beta}_2 = 10.5$ . Calculate tensile strength when percent cotton=15 and drying time=3.5. IF sums of squares were  $SST=1210.3, SSE=117.4$ , compute the coefficient of multiple determination.

**KARACHI UNIVERSITY BUSINESS SCHOOL**  
**UNIVERSITY OF KARACHI**  
**FINAL EXAMINATION, JUNE & JULY 2009: AFFILIATED 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 32    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) Statistic 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:  $\bar{x}_1=32.42$ ;  $s_1 = 6.08$ ;  $\bar{x}_2 =33.72$ ;  $s_2 = 6.75$ . 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** Two floppy manufacturers are competing for our business. To guarantee a ready supply of floppies, we desire to assign half of our purchases to each firm. However, as a preliminary check, we want assurance that their products are not markedly different in quality. One of the primary measures of quality is the easy accessibility to the data saved in the floppies. Ten cases of floppies from each vendor are purchased, and each case is tested to assess the performance of the respective floppies. The test have given  $x_1=1210$ ,  $s_1^2 = 2550$ ,  $x_2=1175$ , and  $s_2^2 = 3600$ .

- i) Can you conclude that the true means are equal at  $\alpha = 0.05$  ?
- ii) Can you conclude that the true variances are also equal at  $\alpha = 0.05$  ?
- ✓iii) 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).

Subjects	1	2	3	4	5	6	7	8	9	10	11	12
Before	201	231	221	260	228	237	326	235	240	267	284	201
After	200	236	216	233	224	216	296	195	207	247	210	209

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

Degree	Low	Medium	High
Arts	20	100	190
Commerce	20	120	70
Science	60	120	100

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

Quantity produced (000):	1	2	4	5	7	9	13
Manufacturing costs (£000):	5	5.9	6.5	7.5	8	9.5	10.8

- Calculate the correlation coefficient for the association between quantity produced and manufacturing costs. Test it for significance at 5%.
- Plot the data on a scatter diagram.
- Calculate the regression line for predicting manufacturing costs from quantity produced, interpret its coefficients and add it to the scatter diagram
- Estimate how much of the variation in costs is accounted for by the model.
- ~~Predict the manufacturing costs of an eighth text-book which has an~~

**KARACHI UNIVERSITY BUSINESS SCHOOL**

**UNIVERSITY OF KARACHI**

**FINAL EXAMINATION: AFFILIATED COLLEGES**

**Advance Business Statistics: BA (M) – 601**

**MBA – III**

Time : 3 Hours

Max. Marks: 60

Date: January 22, 2009

Instructions : (i) Attempt any five (5) questions.

(ii) All questions carry equal marks.

- ✓ Q.1 (a) Differentiate between Parameters and Statistic(s). (3)
- (b) The mean and variance of a normal random variable  $x$  are 50 and 25 respectively. Find  $P(37.5 \leq x \leq 62.5)$   $Z_1$  and  $Z_2$  (9)
- ✗ Q.2 (a) Define central limit theorem. (3)
- (b) What is the probability of drawing a random sample with a mean of 30 or more from a population with mean of 28? The sample size is 100 and the population variance is 81. (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. (3)
- (b) Given two random samples of size  $n_1 = 9$  and  $n_2 = 16$  from two independent normal populations with  $\bar{x}_1 = 64$ ;  $\bar{x}_2 = 52$ ;  $s_1^2 = 36$ ; and  $s_2^2 = 25$ . Find a 95% confidence interval for  $(\mu_1 - \mu_2)$ . Assume  $\sigma_1 = \sigma_2$ . (9)
- ✓ Q.4 (a) Define Type – I and Type – II errors. (3)
- (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  $\bar{x} = 11.2$  minutes and  $s = 1.8$  minutes. At 0.05 level of significance, does this constitute evidence that the figure claimed is too low? (9)
- Q.5 (a) Define Pooled Variance. (3)
- (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 50 bars, cut on machine 2, gives a mean of 55.5 inches with standard deviation of 0.12 inches. At the 0.05 level of significance, are the machines cutting bars of equal length? (9)

continue page – 2

Q.6 (a) Define Null and Alternate hypotheses. (3)

(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 drawn from the population shows the following preferences:

Brand	A	B	C	D
Number of persons	192	342	44	22

Do these data provide sufficient evidence to reject the research's belief at  $\alpha = 0.01$  (9)

Q.7 In terms of Yule's notation if  $r_{12} = 0.8$ ,  $r_{13} = -0.4$ ,  $r_{23} = -0.56$  then find the value of (i)  $r_{12.3}$  (ii)  $R_{1.23}^2$  (12)

Q.8 (a) What is a time series? Describe its components. (3)

(b) Determine a least square equation of trend line fitted to the following data. Estimate the value for the year 2010. (9)

Years	Values
1999	32
2000	41
2001	43
2002	55
2003	50
2004	82
2005	76
2006	79
2007	88

xxx////xxx////xxx////xxx////xxx////xxx////xxx////xxx////xxx////xxx

**KARACHI UNIVERSITY BUSINESS SCHOOL**  
**UNIVERSITY OF KARACHI**  
**FINAL EXAMINATION, JUNE - 2008: AFFILIATED COLLEGES**  
**Advanced Business Statistics: BA (P) - 412**  
BBA - IV

Time: 3 Hours

Date: June 16, 2008

Max. Marks: 60

Instructions: (1) Attempt any five (5) questions.

(2) Use of scientific calculator is allowed.

- Q.No.1 (a) Define Random Variable. Differentiate between discrete and continuous random variables. (4)  
 (b) It is known that 6% of the production of a manufacturer is defective. A random sample of 10 items is selected from the production. The number of defective items in the sample is the binomial random variable  $x$ .  
 Find (i)  $P(x \leq 1)$  (ii)  $E(x)$  (iii)  $V(x)$

- Q.No.2 (a) Define (i) Sampling distribution of sample mean and (4)  
 (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 11?  $\rightarrow P(\bar{x} > 10.5)$

- Q.No.3 (a) In Karachi city a simple random sample of 400 shoppers revealed that 250 prefer reduction sales when purchasing clothing. Construct 95% confidence interval for the true proportion of all shoppers who prefer reduction sales when purchasing clothing. (4)  
 (b) Given two random samples of size  $n_1 = 9$  and  $n_2 = 16$  from two independent normal populations with  
 $\bar{x}_1 = 64$ ;  $\bar{x}_2 = 52$ ,  $S_1^2 = 36$  and  $S_2^2 = 25$ .  
 Find a 95% confidence interval for  $(\mu_1 - \mu_2)$ . Assume  $\sigma_1 = \sigma_2$ . (8)

- Q.No.4 (a) Define the following terms: (4)  
 (i) Null and Alternate Hypothesis  
 (ii) 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 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? (8)

- Q.No.5 (a) A market researcher believes that in a certain population the proportion of persons preferring brands A, B, C and D of toothpaste are 0.30, 0.60, 0.08 and 0.02 respectively. A simple random sample of 600 persons drawn from the population shows the following preferences: (6)

Brand	A	B	C	D
No. of persons	192	342	44	22

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

Please Turn Over

$a = b \times 1.4$

Handwritten notes and calculations on the right side of the page, including  $P(\bar{x} > 10.5)$ ,  $n=400$ ,  $S^2=36$ ,  $S^2=25$ , and  $Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ .

X 2

4.a)

It is generally believed that statures of children depends on the stature of their parent. To test this proposition a survey of 200 children revealed the following information.

Children's Stature	Parent's Stature		
	Tall	Moderate	Short
Tall	48	40	22
Short	30	24	36

$(o_{ij} - c_{ij})^2$

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

5

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.

$\mu = 53$  [04]  
 $\sigma = 6$   
 $n = 16$  [12]

5. A regression for 15 observations is estimated as under

$\hat{Y}_t = 2.691 - 0.4795 X_t$   
 (0.75) (0.134)

Where Y = Coffee consumption in the US (cups/person/day)  
 $X_t$  = retail price of coffee (\$/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 2<sup>nd</sup> Quarter of 2005 are given in million rupees. Fit a linear trend using least squares and estimate the Seasonal component.

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

$\frac{\sum x^2}{n}$  [12]

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

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

GOOD LUCK

2 4

X

K  
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

Max: Marks: 60

Dated: 26/12/2007

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

1.a) Use relevant tables to read/ determine the values of the following. [04]

i)  $(20,24)F_{0.99}$       ii)  $25\chi^2_{0.975}$       iii)  $33t_{0.05}$

iv)  $P[15\chi^2_{0.95} < \chi^2 < 15\chi^2_{0.05}] =$  \_\_\_\_\_

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. [04]

- (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 Indies 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 with W. Indies is 0.65, what is the probability that Pakistan wins the series in the (i) 3<sup>rd</sup> Game (ii) 5<sup>th</sup> Game. [04]

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. [06]  
(ii) Construct a 99% confidence interval for the difference in the proportion of married men and women. [06]

3. A sample of 20 Lux-bars found to weigh 130 grams on the average with a standard deviation of 15 gram. another sample of 22 Rexona-bars found to weigh 140 grams on the average with a standard deviation of 10 grams.

- a) Assuming equal population variance, test at 1% whether Lux-bars weigh less than Rexona on the average. [06]  
b) Obtain a 99% confidence interval for the difference between means. [06]



where  $\mu$  is the mean amount spent for maintenance and repairs per residential property owner.

- Determine the probability of type-I error.
- Determine the probability of type-II error for each of the given values of  $\mu = 306, 331, 360, 381, 431$  and  $482$ .
- What will happen in (ii) if we increase a sample size from 40 to 70?

Q4a). In a packing plant, a machine packs cartons with jars. A salesperson claims that the machine she is selling will pack faster. To test that claim, the times it takes each machine to pack 10 cartons are recorded. The results, in seconds, are shown in the following table:

	New machine $\mu_1$	$\sigma_1$	Present machine $\mu_2$	$\sigma_2$
	42.0	42.7	41.0	43.6
	41.1	43.8	41.8	43.3
	42.4	42.5	42.8	43.5
	43.2	43.1	42.3	41.7
	41.8	44.0	42.7	44.1

Do the data provide sufficient evidence to conclude that, on the average, the new machine packs faster? Perform the required hypothesis test at the 5% level of significance.

b). Determine a 95% Confidence Interval for the difference  $\mu_1 - \mu_2$ .

Q5a). According to the Arizona Real Estate Commission, in 1982 only 10% of people holding real-estate licenses were active in the industry. An independent agency has been asked by the commission to determine whether this percentage is higher. A random sample of 150 licensed people reveals that 24 are currently active. Perform the appropriate hypothesis test at the 1% significance level.

b). Suppose we tell you that the percentage of adult U.S. males who are married exceeds that percentage of adult U.S. females who are married. Further suppose that to check this claim, you randomly select 550 adult U.S. males and 575 adult U.S. females. You find that 367 of the males and 453 of the females are married. Do your data provide sufficient evidence, at the 5% significance level, to support your claim? Explain your answer.

Q6). 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 preliminary study, the economist takes a random sample of eight middle-income families of the same size (father, mother, two children). The results are as follows: where  $x$  denotes disposable income, in thousands of dollars, and  $y$  denotes food expenditure, in hundreds of dollars.

X	30	36	27	20	16	24	19	25
y	55	60	47	40	37	26	39	43

- Determine the regression equation for the data.
- Graph the regression equation and the data points.
- What does the slope of the regression line represent in terms of disposable income and annual food expenditure?
- Use the regression equation to predict the annual food expenditure of a family with a disposable income of \$25,000.

Q7a). Consider the data in Q. No. (6). Compute SST, SSR and SSE and the Coefficient of determination. State how useful the regression equation appears to be for making predictions.

b). Do the data provide sufficient evidence to conclude that the slope of the population regression line is not zero and hence that disposable income is a useful predictor of food expenditure for middle income families of the

12

DEPARTMENT OF BUSINESS ADMINISTRATION  
University of Karachi  
Advanced Business Statistics

Dated: June 16<sup>th</sup> 2005

BBAVI

Course # 312

Time 3 Hrs

Max Marks: 60

Note: Attempt any 5 Questions.

Q1. The city airport commission is considering the establishment of limitations on noise pollution around a local airport. At the present time, the noise level per jet takeoff in one neighborhood near the airport is approximately normally distributed with mean of 100 decibels and a standard deviation of 6 decibels.

- i) What is the probability that a randomly selected jet will generate a noise level greater than 108 decibels?
- ii) What is the probability that a randomly selected jet will generate a noise level of exactly 100 decibels?
- iii) Compute the noise level correspond to less than 35% jet generate a noise level?

b). Consider a population of size  $N = 500$  with a mean  $\mu = 200$  and a S.D  $\sigma = 40$ . Assume that a simple random sample of size  $n = 100$  will be selected from this population.

- (i) What is the value of the standard error of the mean?
- (ii) What is the probability of selecting a simple random sample that provides a value of  $\bar{x}$  that is within  $\pm 5$  of the population mean  $\mu$ ?

Q2. A national survey research firm has past data that indicate that the interview time for a consumer telephone survey has a standard deviation of 6 minutes.

- i) How large a sample should be taken if the firm desires a 98% probability of estimating the mean interview time to within 2 minutes or less?
- ii) Assume that the simple random sample you recommended in (i) is taken and that the mean interview time for the sample is 32 minutes. What is the 98% confidence interval estimate for the mean interview time for the population of interviews?

b). The manufacturer of a new model car, called the Orion, claims that a typical car gets 27 mpg. An independent consumer group is skeptical of this claim and thinks the mean gas mileage of all Orions may very well be less than 27 mpg. To try to justify its contention, the consumer group conducts mileage tests on 30 randomly selected Orions and obtains the following data:

27.5	25.1	25.1	24.6	26.0	25.3
25.3	23.6	26.0	25.4	26.1	26.0
24.1	26.6	25.8	26.4	23.4	23.8
22.6	25.4	25.1	26.6	29.0	24.8
23.3	25.1	26.2	29.6	26.3	23.9

At the 5% significance level, do the data support the consumer group's conjecture? Also determine the P-value.

Q3. a) How you define type-I error and type-II error. Why we concern these two error in our decision making process?

b) According to the Census bureau publication Construction Report, the mean expenditure per residential property owner for maintenance and repairs in 1983 was \$280. Suppose we want to perform a hypothesis test to decide whether the mean amount spent for maintenance and repairs per residential property owner has increased over the 1983 mean of \$280. In a) the null and alternative hypotheses are:

8a)  
b)

Differentiate between correlation and regression? Give an example.

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

Grade-Point averages in Major course (x) starting salary (y in thousands of rupees)

2.58	35.0
3.27	38.0
3.85	44.0
3.50	41.0
3.33	39.0
3.50	46.0
4.00	50.0

- Plot the data.
  - Find the values of correlation coefficient ( $r$ ) and interpret the result.
  - Find the least square regression equation.
  - Check whether or not the fit is good.
  - What is the mean starting salary for graduates with grade point averages equal to 3.0?
- 9.
- Explain "Seasonal Variations" Give examples where we find seasonal variations and why? What are the methods of Isolating Seasonal Variations?
  - The US-Cars has recorded sales (in \$1,000s) over the last two years of

Month	2004	2005
Jan	17.2	18.1
Feb	18.7	19.2
March	19.7	20.3
April	20.2	21.5
May	21.7	22.0
June	23.1	24.7
July	24.2	23.9
Aug	25.7	26.2
Sept	21.2	22.0
Oct	19.3	18.0
Nov	22.7	19.7
Dec	19.3	17.3

- Plot the data. Does there appear to be any trend in the data? Any Cyclical or Any Seasonal variation?
- Calculate the seasonal indices.
- 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  $\mu=40$  and  $\sigma=6$ , 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  $X_1, \dots, X_n$  be a random sample from  $N(\mu_1, \sigma_1^2)$  and  $Y_1, \dots, Y_n$  be a random sample from  $N(\mu_2, \sigma_2^2)$ . Assume that the two samples are independent of each other. Derive an expression for  $\mu_1 - \mu_2$  at the  $(1 - \alpha)100\%$  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  $\alpha = 0.05$ ?
- 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 list of parts from two suppliers.

Supplier	Mean	Sample size	S.D.
A	50	150	10
B	100	153	5

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:
- Type A** 7.3 8.3 7.6 6.8 8.0
- 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 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 Hypotheses
  - ii) Simple and Composite Hypotheses
  - iii) Statistic and Parameter
  - iv) One tailed 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:  $\bar{y}_1 = 12.5$ ,  $s_1^2 = 101.17$ , and  $n_1 = 8$ . After installation, a random sample yielded;  $n_2 = 9$ ,  $\bar{y}_2 = 10.2$ , and  $s_2^2 = 94.73$ .
- i) Can you conclude that the true variances are equal at  $\alpha = 0.10$ ?
  - ii) Construct a 95% confidence interval for the ratio 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:

	Non-smokers	Moderate smokers	Heavy smokers
Hypertension	21	36	30
No hypertension	48	26	19

Test the hypothesis that the presence or absence of hypertension is independent of smoking habits. Use a 0.05 level of significance.

Please Turn Over

- 7) 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

Grade-Point averages in Major course (x)	starting salary (y in thousands of rupees)
2.58	35.0
3.27	38.0
3.85	44.0
3.50	41.0
3.33	39.0
3.50	46.0
4.00	50.0

- Plot the data.
- Find the values of correlation coefficient ( $r$ ) and interpret the result.
- Find the least square regression equation.
- Check whether or not the fit is good.
- 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

Month	2004	2005
Jan	17.2	18.1
Feb	18.7	19.2
March	19.7	20.3
April	20.2	21.5
May	21.7	22.0
June	23.1	24.7
July	24.2	23.9
Aug	25.7	26.2
Sept	21.2	22.0
Oct	19.3	18.0
Nov	22.7	19.7
Dec	19.3	17.3

- Plot the data. Does there appear to be any trend in the data? Any Cyclical or Any Seasonal variation?
- Calculate the seasonal indices.
- What are the deseasonalised values? How would you interpret them?

Name \_\_\_\_\_ Regular/Repeater: \_\_\_\_\_ Seat No: \_\_\_\_\_

**Karachi University Business School, Affiliated Colleges**

**BBA-IV: BA(P) 412 Advanced Business Statistics**

**Dated: 12<sup>th</sup> December'2006**

**Max Time: Three Hours**

X

- 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  $\mu=40$  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:  $\bar{y}_1 = 12.5$ ,  $s_1^2 = 101.17$ , and  $n_1 = 8$ . After installation, a random sample yielded;  $n_2 = 9$ ,  $\bar{y}_2 = 10.2$ , 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 list of parts from two suppliers.

Supplier	Mean	Sample size	Variance
A	50	28	25
B	100	15	49

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 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 women 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 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  $\sigma > 0.9$  year? Use a 0.05 level of significance.

**Please Turn Over**

Q7

Low-involvement consumer purchases of items such as floor wax, dishwashing detergent, and ballpoint pens 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 minds," 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.

X	Y
2.5	25
4.0	40
5.0	55
7.5	60
8.0	70

- Identify the predictor and response variable.
- Find the regression equation for the above data by the method of least square.
- Use the regression equation to predict the market share when the brand awareness index value is 6.0.
- 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. [04]
- i)  $(20,24)F_{0.99}$       ii)  $25\chi^2_{0.975}$       iii)  $33t_{0.05}$
- iv)  $P[15\chi^2_{0.95} < \chi^2 < 15\chi^2_{0.05}] = \underline{\hspace{2cm}}$
- 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. [04]
- (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 Indies 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 with W. Indies is 0.65, what is the probability that Pakistan wins the series in the (i) 3<sup>rd</sup> Game (ii) 5<sup>th</sup> Game. [04]
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. [06]  
(ii) Construct a 99% confidence interval for the difference in the proportion of married men and women. [06]
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 weigh 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. [06]  
b) Obtain a 99% confidence interval for the difference between means. [06]



- 4.a) It is generally believed that statures of children depends on the statures of their parent. To test this proposition a survey of 200 children revealed the following information.

<u>Children Stature</u>	<u>Parent's Stature</u>		
	<u>Tall</u>	<u>Moderate</u>	<u>Short</u>
<u>Tall</u>	48	40	22
<u>Short</u>	30	24	36

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]

$$\hat{Y}_t = 2.691 - 0.4795 X_t$$

(0.75) (0.134)

Where Y = Coffee consumption in the US (cups/person/day)  
 $X_t$  = retail price of coffee (\$ 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 2<sup>nd</sup>. Quarter of 2005 are given in million rupees. Fit a linear trend using least squares and estimate the Seasonal component. [12]

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

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

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

GOOD LUCK

Name \_\_\_\_\_ Regular/Repeater: \_\_\_\_\_ Seat No: \_\_\_\_\_

**Karachi University Business School, Affiliated Colleges**

**BBA-IV: BA(P) 412 Advanced Business Statistics**

**Dated: 12<sup>th</sup> 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  $\mu=40$  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:  $\bar{y}_1 = 12.5$ ,  $s_1^2 = 101.17$ , and  $n_1 = 8$ . After installation, a random sample yielded;  $n_2 = 9$ ,  $\bar{y}_2 = 10.2$ , 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 list of parts from two suppliers.

Supplier	Mean	Sample size	Variance
A	50	28	25
B	100	15	49

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 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 women 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 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  $\sigma > 0.9$  year? Use a 0.05 level of significance.

- 4.a) It is generally believed that statures of children depends on the statures of their parent. To test this proposition a survey of 200 children revealed the following information.

<u>Children Stature</u>	<u>Parent's Stature</u>		
	<u>Tall</u>	<u>Moderate</u>	<u>Short</u>
<u>Tall</u>	48	40	22
<u>Short</u>	30	24	36

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]

$$\hat{Y}_t = 2.691 - 0.4795 X_t$$

(0.75) (0.134)

Where Y = Coffee consumption in the US (cups/person/day)  
 $X_t$  = retail price of coffee (\$ 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 2<sup>nd</sup> Quarter of 2005 are given in million rupees. Fit a linear trend using least squares and estimate the Seasonal component. [12]

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

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

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

GOOD LUCK

Name \_\_\_\_\_ Regular/Repeater: \_\_\_\_\_ Seat No: \_\_\_\_\_

**Karachi University Business School, Affiliated Colleges**

**BBA-IV: BA(P) 412 Advanced Business Statistics**

**Dated:** 12<sup>th</sup> 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  $\mu=40$  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:  $\bar{y}_1 = 12.5$ ,  $s_1^2 = 101.17$ , and  $n_1 = 8$ . After installation, a random sample yielded;  $n_2 = 9$ ,  $\bar{y}_2 = 10.2$ , 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 list of parts from two suppliers.

Supplier	Mean	Sample size	Variance
A	50	28	25
B	100	15	49

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 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 women 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 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  $\sigma > 0.9$  year? Use a 0.05 level of significance.

- Q7 Low-involvement consumer purchases of items such as floor wax, dishwashing detergent, and ballpoint pens 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 minds" 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.

X	Y
2.5	25
4.0	40
5.0	55
7.5	60
8.0	70

- Identify the predictor and response variable.
- Find the regression equation for the above data by the method of least square
- Use the regression equation to predict the market share when the brand awareness index value is 6.0.
- 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

930

- Q1 a) What is Binomial distributions and in which situation 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 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 5 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  $\sigma=6$ , 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) 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  $\alpha = 0.05$ ?

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.

Supplier	Mean	Sample size	Standard Deviation
A	50	150	10
B	100	153	5

*done*

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:

Type A	7.3	8.3	7.6	6.8	8.0	
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 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 Hypotheses ii) Simple and Composite Hypotheses  
 iii) Statistic and Parameter iv) One tailed 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:  $\bar{y}_1 = 12.5$ ,  $s_1^2 = 101.17$ , and  $n_1 = 8$ . After installation, a random sample yielded;  $n_2 = 9$ ,  $\bar{y}_2 = 10.2$ , and  $s_2^2 = 94.73$ . Can you conclude that the true variances are equal at  $\alpha = 0.10$ ?

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:

	Non-smokers	Moderate smokers	Heavy smokers
Hypertension	21	36	30
No hypertension	48	26	19

*done*

Test the hypothesis that the presence or absence of hypertension is independent of smoking habits. Use a 0.05 level of significance.

- 8a) 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

Grade-Point averages in Major course (x)	starting salary (y in thousands of rupees)
2.58	35.0
3.27	38.0
3.85	44.0
3.50	41.0
3.33	39.0
3.50	46.0
4.00	50.0

- 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

Month	2004	2005
Jan	17.2	18.1
Feb	18.7	19.2
March	19.7	20.3
April	20.2	21.5
May	21.7	22.0
June	23.1	24.7
July	24.2	23.9
Aug	25.7	26.2
Sept	21.2	22.0
Oct	19.3	18.0
Nov	22.7	19.7
Dec	19.3	17.3

- 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?

*Handwritten:* M. JEM

**UNIVERSITY OF KARACHI**  
**DEPARTMENT OF BUSINESS ADMINISTRATION**  
**ADVANCE STATISTICS**  
**TERMINAL EXAMINATION - 2002 (Dec)**

Time Allowed: 3 Hours

Max Marks: 60

Note: Attempt any Six Questions. All questions carry equal Marks.

Q1 : (a) Explain what is meant by the statement "We are 95% confident that our interval estimate contains  $\mu$ ." Explain the difference between an interval estimator and a point estimator for  $\mu$ .

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).

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

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.

Location	Month											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
South Plant	2.0	2.0	2.1	1.9	1.7	1.8	1.7	1.9	2.0	2.0	2.1	2.2
Fire Station	2.2	2.2	2.1	2.0	1.9	1.9	1.8	1.7	1.9	1.9	1.8	2.0

(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 nonprofessionals 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 total  $T$  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 96% confidence if  $N = 800$ ,  $d = .15$  and  $\pi = 0.40$  for simple random sample.

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

(b) A large firm knows that of 500 accounts 200 are wholesales and 300 are retail. An auditor wishes to sample  $n = 100$  of these accounts in order to estimate the average amount of accounts receivables for the firms. A simple random sample turn out to contain 70 wholesale accounts and 30 retail accounts. The following data (in dollars) is obtained.

Wholesale	Retail
$n_1 = 70$	$n_2 = 30$
$y_1 = 520$	$y_2 = 280$
$S_1 = 210$	$S_2 = 90$

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

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

- $H_0$  : The production process is performing satisfactorily  
 $H_a$  : The process is performing in an unsatisfactorily manner

Accordingly,  $\alpha$  is sometimes referred as producer's risk, while  $\beta$  is called the consumer's risk (Montgomery, 1991). An injection molder produces plastic golf tees. The process is designed to produce tees with a mean weight of .250 ounce. To investigate whether the injection molder is operating satisfactorily, 40 tees are randomly sampled from the last hour's production. Their weights (in ounce) are listed below:

247	.252	.248	.251	.248	.252	.250	.252
253	.253	.254	.253	.256	.254	.253	.252
253	.250	.253	.254	.256	.251	.256	.251
249	.254	.256	.253	.255	.255	.253	.255
251	.253	.251	.251	.253	.251	.254	.253

- (a) Do the data provide sufficient evidence to conclude that the process is not operating satisfactorily? Test using  $\alpha = 0.05$   
 (b) In the context 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.

(b) 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?

Problem	Team 1	Team 2
1	12	25
2	15	26
3	14	21
4	21	23
5	19	31
6	12	19
7	25	35
8	18	28
9	17	27
10	20	26