

**Statistics for Economists**  
**2006**

***Total time: 3 hours***

***Maximum marks:***

***100***

*Attempt any four questions. All questions carry equal marks. Use of Scientific Calculator and statistical tables are allowed.*

1. (a) Differentiate between descriptive statistics and inferential statistics.

(b) Prove the  $AM > GM > HM$

Where:

A.M stands for Arithmetic Mean

GM stands for Geometric Mean

HM stands for Harmonic Mean.

(c) You are sales manager for a regional division of a beverage company. The sales goals for you representative have an average of \$768,000 with a standard deviation (SD) of \$240,000. You have been instructed to raise the sales goals of each representative by \$85,000.

(i) What happen to standard deviation and mean?

(ii) Compare the coefficient of variation (CV) before and after the sales goal adjustment. Why does it change (or not change) in this way?

(a) What is extrapolation? Why is it especially troublesome?

(b) (i) What is covariance between X and Y?

(ii) Which is easy to interpret, the covariance or the correlation? Why?

(c) The following sample observations were randomly selected:

X:    4    5    3    6    10

Y: 4 6 5 7 7

(i) Determine the regression equation of Y on X, also interpret the regression coefficients.

(ii) Determine the value of Y when X is 07.

(iii) At 0.05 significance level, can we conclude that there is positive correlation between X and Y.

(a) (i) What is Multi-collinearity?

(ii) What are the harmful effects of extreme Multi-collinearity?

(b) In a Multiple Regression equation K (Number of independent variables) = 05 and  $n = 20$ , the MSE (Mean Square Error) value is 5.10 and SS total is 519.68. At the 0.05 significance level, can we conclude that any of the regression coefficients are not equal to Zero?

(c) Three variables have in pairs simple correlation coefficients given by  $r_{12} = 0.8$ ,  $r_{13} = -0.7$ ,  $r_{23} = -0.9$ . Find Multiple correlation coefficient  $R_{1,23}$  of  $X_1$  on  $X_2$  and  $X_3$ .

4. (a) Differentiate between probability distribution and sampling distribution.

(b) Two coins are tossed. What is the conditional probability that two heads results, given there is at least one head,

(c) What is Binomial Distribution? Just state its properties.

(d) A new project will be declared successful if you achieve a market share of 10% or more in the next two years. Your marketing department has considered all possibilities and decided that it expects the product to attain a market share of 12% in this time. However this number is not certain The Standard Deviation (SD) is forecasted to be 3% indicating the uncertainty in 12% forecast as 3% points. You may assume a Normal Distribution.

(i) Find the probability that the new project is successful.

(ii) Find the probability that the new project fails.

(iii) Find the probability that the new project is widely successful defined as achieving at least a 15% market share.

(iv) To assess the precision of the marketing projections, find the probability that the attained market share falls close to the projected value of 12%, that is, between 11% to 13%.

5. (a) Differentiate between the following

(i) Probability and Non probability sampling.

(ii) Point estimates and confidence intervals.

(iii) Type I Error and Type II Error

(iv) Null Hypothesis and Alternative Hypothesis.

(v) 't' statistic and 'z' statistic.

(b) The following hypothesis are given:

$$H_0: \sigma_1^2 < \sigma_2^2 \quad H_1: \sigma_1^2 > \sigma_2^2$$

A random sample of 05 observations from the first population resulted in a standard deviation (SD) of 12. A random sample of 07 observations from the second population showed a standard deviation of 07. At 01% significance level, is there more variation in First Population?

6. (a) What do you know about indexes? Why do we convert data into indexes?

(b) Suppose that, in the year 2006, in Pakistan, the Nominal rate of interest is 17.5%, the rate of inflation is 5.25%, what is real rate of interest in Pakistan in 2006?

(c) Cally's Software. Inc. is a rapidly growing supplier of computer software to the Sarasota area. Sales for the last five years are given below:

Years:	1996	1997	1998	1999	2000
	1.1	1.5	2.0	2.4	3.1

(i) Determine Logarithmic trend equation.

(ii) Compute Geometric mean rate of increase or decrease from 1996 to 2000.

(iii) Estimate sales for the year 2005.

7. Write a note on any three of the following:

(a) Exponential Smoothing

- (b) Histograms and Histograms by Hand
- (c) Types of Measurements Scales.
- (d) P-value in Hypothesis Testing
- (e) Chi-Squared test for independence.

### 2007

1. a. Briefly explain various types of Measurement Error.
- b. State the Properties of Arithmetic Mean.
- c. The following returns were realized on an Investment over a 5-year period.

Year	1	2	3	4	5
Rate of Return	0.10	0.22	0.06	-0.05	0.20

- i. Compute the Mean, and Median of the returns.
  - ii. Also find Geometric Mean.
  - iii. Which one of the three statistics computed above test describes the return over the 5-year period? Give reasons.
2. (a) What is the difference between “Standardized Variable” and “Coefficient of Variation.”
  - (b) (i) A friend Ahmed Bilal calculates a Standard Deviation and reports that it is  $-0.5$ . How do you know he has made a serious calculation error?
  - (ii) Create a sample of five numbers whose Mean is 10 and whose Standard Deviation is Zero.
  - (c) Draw the Box Plot and identify Outlines, if any, of the following set of data:  
9, 28, 15, 21, 12, 22, 29, 20, 23, 31, 11, 19, 24, 16, 13.
3. (a) Differentiate between:
    - (i) Correlation Coefficient and Coefficient of Determination.
    - (ii) The ‘t’ Statistics and the “F” statistics.
    - (iii) Confidence Interval and Prediction Interval Estimate.
    - (iv) Extrapolation and Interpolation.

(v) Level of Significance and P - Value.

(b) A simple Regression  $TR = a + bQ$  is not able to explain 21% variation in TR. What is Coefficient of Correlation and Coefficient of Determination between TR and Q.

(c) On the first Statistics Examination, the Coefficient of Determination between Hours Studied and Grade Earned was 0.80. The Standard Error of Estimate was 10. There were 20 students in the class. Develop an ANOVA table.

4. (a) In a Multiple Regression Model, what would you suspect is the problem if the  $R^2$  is larger and significant, but non of the "X" variable has a "t" test that is significant.

(b) "Standard Error of Estimate" cannot be considered as an Absolute Criterion for evaluating Model S Utility. Discuss.

(c) (i) Three Variables have in Paris Simple Correlation Coefficient, given by:

$$r_{12} = 0.8, \quad r_{13} = -0.7, \quad r_{23} = -0.9,$$

Find Multiple Correlation Coefficients  $R_{1,23}$  and  $R_{2,13}$ .

(ii) What are the Assumptions of Multiple Linear Regression Model?

(d) Write a brief note on "Normal Distribution."

5. a. Determine and Differentiate the followings:

(i) Type I and type II Errors.

(ii) Interval Estimate and Point Estimate.

(iii) Probability and Non Probability Sampling.

(iv) One sided and two sided Confidence Interval.

b. A federal agency responsible for enforcing laws governing Weights and Measures routinely inspects packages to determine whether the weight of contents is at least as great as that advertised on the package. A random sample of 18 containers whose packaging slates that the contents weight 08 ounces was drawn. Can we conclude that on average the containers are mislabeled?

(Use  $\alpha = 0.10$ )

7.80, 7.91, 7.93, 7.99, 7.94, 7.75, 7.95, 7.79, 8.06, 7.82, 7.89, 7.92, 7.87, 7.92, 8.05, 7.91

6. Write a Comprehensive Note on:

- a. Chi-Square Test
- b. The “F” Distribution.

7. a. What do you know about Indexes? Why do we convert data into Indexes?

b. i) Suppose Consumer Price Index (CPI) for June, 2008, is 100 (1999-00 = 100). What is a purchasing Power of a dollar?

ii) Suppose that, in the year 2008, in Pakistan, the Rate of Inflation is 10.6% and Nominal Rate of Interest is 2.0%. What is the Real Rate of Interest in Pakistan in the year 2008’?

(iii) N.B Books is a rapidly growing book store in Johar Town, Lahore. Sales for the last five years are given below:

Year	Sale (\$ in million)
2003	1.0
2004	1.5
2005	2.5
2006	3.0
2007	3.1

Determine Logarithmic Trend Equation.

- (i) Compute Geometric Mean Rate of Change from 2003 to 2007.
- (ii) Estimate sales for the year 2010.

## 2009

1 (a) Differentiate between

- (i) Univariate and Bivariate Data
- (ii) Qualitative and Quantitative Data
- (iii) Descriptive and Inferential Statistics
- (iv) Measure of Central Tendency and Measure of Dispersion

(v) Box Plot and Detailed Box Plot

(b) At the end of the term, university students often complete questionnaires about their courses. Suppose that in one university, students were asked the following

(i) Rate the course (highly relevant, relevant, irrelevant)

(ii) Rate the professor (very effective, effective, not too effective, not at all effective)

(iii) What was your mid-term grade (A, B, C, D)?

Determine the type of data each question produces.

(c) State and explain the properties of Arithmetic Mean (AM)

2. (a) Differentiate between Absolute Measures of Dispersion and Relative Measures of Dispersion.

(b) Briefly explain the properties of Standard Deviation and Variance.

(c) In a study of the mileage of automobiles manufactured in 2008, the mean number of miles per gallon was 27.5 and the median was 26.8. The smallest value in the study was 12.70 miles per gallon and largest value was 50.20. The first and the third quartiles were 17.95 and 35.45 miles per gallon respectively. Develop a Box Plot and comment on the distribution. Is it a Symmetrical Distribution?

3. (a) Distinguish between Causation and Correlation.

(b) What is the difference between Interpretation and Extrapolation?

(c) Pakistan International Airlines (PIA) selected a random sample of 25 flights and found that the correlation between the number of passengers and the total weight, in pounds, of luggage stored in the luggage compartment is 0.94. Using the 0.05 significance level, can we conclude that there is positive association between the two variables?

4. (a) Write a brief note on Multi-collinearity.

(b) How can you differentiate between Confidence Interval and Prediction Interval?

(c) What are the different shortcomings of Coefficient of Determination? How can these shortcomings be resolved?

(d) Given the following ANOVA table,

source	DF	SS	MS	F
Regression	1	1000		
Error				
Total	14	1500		

(i) Complete the ANOVA

(ii) Determine the Standard Error of Estimation and Interpret it.

(i) Assuming the direct relationship between variable, what is the Coefficient of Correlation and Coefficient of Determination?

5. (a) Briefly explain the difference between the following terms

(1) Chi-squared test and z-test

(ii) Null and Alternative Hypotheses

(iii) Sampling and Non-Sampling Error

(iv) Point Estimation and Interval Estimation

(v) Sample Frame and Sample Design

(b) A fast food franchiser is considering building a restaurant at a certain location, Based on financial analysis, a site is acceptable only if the number of pedestrians passing the location averages more than 100 per hour, the number of pedestrians observed for each of 40 hours was recorded and  $\bar{X} = 105.7$ ,  $\sigma = 12$ . Can we conclude that the site is acceptable (set your own significance level).

6. (a) State the different applications of index numbers in an economy.

(b) What is Prediction and Forecasting?

(c) Suppose that a random sample of 100 observations was drawn from a population, after which the Mean and Standard Deviation were calculated. Each observation was standardized and the number of observations in each of the intervals below was counted. Can we infer at the 5% significance level that the data were not drawn from a normal population?

Interval	Frequency
$Z \leq -1.5$	10
$-1.5 < Z \leq -0.5$	18
$-0.5 < Z \leq 0.5$	48
$0.5 < Z \leq 1.5$	16
$Z > 1.5$	8

7. Write a note on the following:

- (i) Skewness and Kurtosis
- (ii) Sampling and its different types
- (iii) Components of a time Series

### 2010

1. a) Differentiate between the following:

- (i) Biased and Unbiased Error
- (ii) Time Series and Cross Sectional Data
- (iii) Parameter and Statistic
- (iv) Quantitative and Qualitative Data
- (v) Bar Chart and Pie Chart

b) State the advantages and disadvantages of Mean, Median and Mode.

c) Consider a sample with data values of 10, 20, 12, 17 and 16, Prove that  $AM > GM > HM$ , where  $AM =$  Arithmetic Mean,  $GM =$  Geometric Mean and  $HM =$  Harmonic Mean

2. a) Differentiate between Mean Deviation and Standard Deviation.

b) Differentiate between Absolute and Relative Measures of dispersion.

c) The mean marks of 40 Economics Students in a University is 44 while when we calculate the relative measure of variability, it was 18.2. Calculate Standard Deviation and Variance for these 40 students. What is the unit of measurements of these Measures of Dispersion?

d) The national average for the verbal portion of the College Board's Scholastic Aptitude Test (SAT) is 507. The College Board periodically rescales the test scores such that the standard deviation is approximately 100. Answer the following questions using a Bell Shaped Distribution and the Empirical Rule for the verbal test scores.

(i) What percentage of the students have a SAT verbal score greater than 607?

(ii) What percentage of the students have a SAT verbal score between 307 and 607?

(iii) What percentage of the students have a SAT verbal score between 407 and 507?

3. a) The commercial division of a real estate firm conducted a study in Lahore to determine the extent of the relationship between annual gross rents (Rs. 1,000s) and the selling price (Rs. 1,000s) for apartment buildings. Data were collected on several properties sold and information is as following of Excel output:

ANOVA				
Source	df	Sum of Squares (SS)	Mean Squares (MS)	F test
Regression	1	41587.3		
Residual	7			
Total	8	51984.1		
		Coefficient	Standard Error	t-stat
Intercept		20.000	3.2213	6.21
Annual Gross Rents		7.210	1.3626	5.29

(i) Complete ANOVA table

(ii) How many apartment buildings were in the sample?

(iii) Write the Estimated Regression Equation and interpret the regression coefficient.

(iv) Use the F test to determine whether the selling price is related to annual gross rents,  $\alpha = 0.05$ .

b) How can you differentiate between Multiple and Partial Correlation Coefficients?

c) Make a clear note on Interpolation and Extrapolation.

4. a) What a Standard Error of Estimate? Explain it.

b) What are different assumptions of a Linear Regression Model?

c) From a large group of students, 10 students were taken as a sample. Their final marks in Economics (X), Statistics (Y) and Mathematics (Z) gave the relationship as following:

$$r_{YZ} = 0.2354, r_{XY} = 0.96250 \text{ and } r_{XZ} = 0.0360$$

(i) Compute Correlation Matrix.

(ii) Detect and interpret Multi-collinearity between the variables.

(iii) Compute Partial and Multiple Correlation Coefficients.

5 a) State the difference between the following:

(i) Sampling and Non-Sampling Error

(ii) Permutation and Combination Formula

(iii) Discrete and Continuous Random Variable

(iv) Probability and Non Probability Sampling

(v) Population and Sample

b) Gasoline prices reached record high level in 16 cities during 2009. Two of the affected cities were Lahore and Karachi. The Pakistan Automobile Association reported a sample mean price of Rs. 2.04 per gallon in Lahore and a sample mean price of Rs. 1.72 per gallon in Karachi. Use a sample size of 40 for the Lahore data and a sample size of 35 for the Karachi data. Assume that prior studies indicating a population standard deviation of 0.10 in Lahore and 0.08 in Karachi are reasonable.

(i) What is a point estimate of the difference between the population mean prices per gallon in Lahore and Karachi?

(ii) At 95% confidence, what is the margin of error?

(iii) What is the 95% confidence interval estimate of the difference between the population mean prices per gallon in the two cities?

6. a) Explain the different steps in constructing Price Index Numbers.

b) Write a short note on CPI versus GDP Deflator. Where we use

CPI and where we recommend GDP Deflator.

c) The take-home pay of Ahmed Bilal and the CPI for 2003 and 2009 are:

<b>Year</b>	<b>Take Home Pay</b>	<b>CPI (2000-01 = 100)</b>
<b>2003</b>	<b>Rs. 25,000</b>	<b>152.4</b>
<b>2009</b>	<b>Rs. 41,200</b>	<b>172.2</b>

- (i) What was Ahmed Bilal's real income in 2003?
- (ii) What was Ahmed Bilal's real income in 2009?
- (iii) Interpret your findings.

7. Write a short note on any three of the following.

- (i) Properties of Normal Distribution
  - (ii) Non-Linear Regression Models
  - (iii) Four Factors of a Time Series
  - (iv) Box Plot and Detailed Box Plot
-