

# Statistics

## Advanced Statistical Analysis

Annual Examination 1997

University of Karachi

**Time Allowed: 3**

**hours**

**Maximum Marks: 100**

*Instructions:*

- 1) Attempt any 5 questions,
- 2) Marks are indicated against each question.

**1. (a)** Describe briefly the different types of diagrams generally used for the presentation of a statistical data **(10)**

**(b)** The amount of money collected by each of 300 children, involved in a sponsored walk, was recorded as follows:

Amount collected (in Rs.)	0 – 1	2 – 4	5 – 10	11 – 15	16 – 25
No. of children	20	93	90	58	39

Draw a Histogram to represent these data.

**(10)**

**2. (a)** Explain clearly the meaning, the best method of computing and the use of the following:

(i) Geometric Mean,

(ii) Harmonic Mean

**(10)**

**(b)** Following frequency distribution show the weights (recorded to the nearest kg.) of 100 students of University of Karachi.

**Table: Weights of 100 students at University of Karachi**

<b>Weights (in kgs.)</b>	60 – 62	63 – 65	66 – 68	69 – 71	72 – 74
<b>No. of students</b>	5	18	42	27	8

Based on the given table, calculate:

(i) Arithmetic Mean,

(ii) Standard Deviation.

**(10)**

3. Give answers to the following questions:

(a) If A and B are mutually exclusive events and  $P(A) = 0.3$  and  $P(B) = 0.5$  then find  $P(A \cap B)$ . **(4)**

(b) In how many ways can 5 differently coloured marbles be arranged in a row? **(4)**

(c) A coin is tossed and a dice is rolled once. List the elements of the sample space. **(4)**

(d) In how many ways can a committee of five be chosen of nine people? **(4)**

(e) What is the probability that at least one head appears in 3 tosses of a fair coin? **(4)**

4. (a) Define Standard Deviation. State and prove algebraically the characteristics of standard deviation. **(10)**

(b) Following table shows the distribution of the maximum load in kilo-Newtons supported by certain cables produced by a company:

**(10)**

<b>Maximum load (kn)</b>	<b>No. of Cables</b>
93 – 97	2
98 – 102	5

103 – 107	12
108 – 112	17
113 – 117	14
118 – 122	6
123 – 127	3
128 – 132	1

5. **(a)** Discuss the two techniques by which correlation may be studied. **(10)**

**(b)** The demand for a commodity in relation to its price is given as under:

Price (in Rs.)	40	42	45	47	50	52	60
Demand (in kgs.)	26	36	28	25	20	30	18

Fit the regression line, using the Least Square Method, and predict demand when price is Rs. 48. **(10)**

6. **(a)** Briefly explain the following:

- (i) Kurtosis,
- (ii) Skewness.

**(12)**

**(b)** Ten speakers were ranked by two judges as follows:

<b>Speakers</b>	A	B	C	D	E	F	G	H	I	J
<b>Ranks given by X</b>	4	7	3	5	1	6	2	10	8	9
<b>Ranks given by Y</b>	6	5	2	4	1	9	3	10	8	7

Does this ranking suggests that the judgement is consistent? Support your answer with facts. **(8)**

7. **(a)** Show that Lespeyre's index number formula does not satisfy the (i) time reversal test and (ii) factor reversal

test.

**(8)**

**(b)** Following table shows a country's average wholesale price and production of milk, butter and cheese for the year 1949, 1950 and 1958:

	Price (per kg.)			Quantity produced (in million of kgs)		
	1949	1950	1958	1949	1950	1958
Milk	3.95	3.89	4.13	9675	9717	10436
Butter	19.49	62.2	59.7	117.7	115.5	115.5
Cheese	1949	35.4	38.9	77.93	74.39	82.79

Using the table's data, compute the Laspayre's price index and Paasche's price index for 1958 with 1949 as base year.

**(12)**

8. Write short notes on any two of the following:

- (a) Interpolation,
- (b) Classification of time series movements,
- (c) Relative measure of dispersion,
- (d) Ideal average.

**(20)**

9. **(a)** Define and explain with suitable examples the:

- (i) Mathematical probability, and
- (ii) Statistical probability.

**(10)**

(b) A government agency employees 100 clerks – typists, classified by sex and marital status as shown in the table given below:

Sex	Marital Status		Total
	Single	Married	
Male	16	24	40
Female	24	36	60

If an employee is picked at random from 100 employees, what is the probability that he or she is:

- (i) single?
  - (ii) single, given that the employee is male?
- (10)**

-----

**Annual Examination 1998**

1. Teacher in primary schools are classified on the basis of daily wages as under:

Wages (in Rs.)	No. of Teachers
100 – 103.9	10
104 – 107.9	20
108 – 111.9	30
112 – 115.9	35
116 – 119.9	30
120 – 123.9	25
124 – 127.9	20
128 – 131.9	15
132 – 135.9	10

136 – 149.9	3
140 – 143.9	2

(a) Determine the Mean and the Standard Deviation daily wages.

(b) Determine, also the percentage of teachers drawing:

- (i) more than average wage,
- (ii) less than (mean + S.D.) wage,
- (iii) if the management decides to raise the wages 15%, how the consistency of wages would be effected?

2. (a) Define and distinguish between:

- (i) Primary and Secondary data,
- (ii) Census and Sample survey,
- (iii) Partial and Multiple correlation.

(b) A series consisting of 10 observations found to have mean 50 and standard deviation 4, later two observations (48 and 50) were replaced by (45 and 55), determine the new Mean and Variance.

3. Given the following data for 1980 – 1991 on annual disposable personal income (X) and personal consumption expenditure (Y) in billion rupees. Estimate, using an appropriate linear regression, the level of consumption in 1992 if the disposable income is Rs. 100 million. Also determine the value of  $r$ .

X	45	51	56	60	62	67	71	74	77	82	93	95
Y	43	48	55	58	59	62	65	70	73	80	89	91

4. (a) Define and explain the following:

- (i) Probability,
- (ii) Mutually exclusive events,
- (iii) Equality likely events.

**(b)** State and explain with examples or proof the characteristic features of Arithmetic Mean.

5. Referring the data in Question 1 above, determine and interpret:

- (i) the lowest wage of the top 15 percent of the teachers,
- (ii) the highest wage of the bottom 25 percent of the teachers,
- (iii) the Coefficient of Skewness ( $\beta_1$ ).

6. **(a)** Explain the significance of index number in economics.

**(b)** For the data given below construct the Marshal's and Fischer's indices for 1995 with 1990 as base:

Items	Prices		Quantity	
	1990	1995	1990	1995
A	12	15	20	18
B	10	12	10	12
C	15	12	15	12
D	18	20	12	12
E	20	25	10	15

7. **(a)** First two moments for a set of 10 observations about 15 as origin are given as under:

$$\mu'_1 = 1.5 \text{ and } \mu'_2 = 76.5$$

Determine the first two Absolute Moments and Raw Moments about 18 as origin.

**(b)** Let  $X$  be a normally distributed random variable having a mean 55 and a standard deviation 8. Determine the probabilities for the following statements:

- (i)  $P(X > 50)$

- (ii)  $P(X < 53)$
- (iii)  $P(50 < X < 58)$
- (iv)  $P(45 < X < 53)$
- (v)  $P(58 < X < 68)$

8. Given the following information:

$$\log 2 = 0.3010$$

$$\log 3 = 0.4771$$

$$\log 5 = 0.6990$$

$$\log 7 = 0.8451$$

$$\log 10 = 1.0000$$

Using the suitable formula interpolate the values of  $\log 4$ ,  $\log 20$ ,  $\log 28$  and  $\log 40$ .

9. Write short notes on any two of the following:

- (i) Time series analysis,
- (ii) Organisation of sample survey,
- (iii) Method of least square,
- (iv) Rank correlation.

-----  
**Annual Examination 1999**

1. **(a)** Not available.

**(b)** A list was given to 200 candidates for few vacancies in a bank. Marks obtained in the test are presented in the following table:

Marks	No. of Candidates
-------	-------------------



1 – 6	10
7 – 12	30
13 – 18	40
19 – 24	45
25 – 30	35
31 – 36	20
37 – 42	15
43 – 48	5

(a) Determine the Mean and the Standard Deviation.

(b) If the top 10 percent of the candidates are to be considered for employment what is the lowest limit of marks?

(c) Approximately what percent of candidates earned more than  $(\bar{x} + 2\sigma)$  marks

2. For the data given in Question 1 (b) above, determine:

(a) First three Raw Moments about 21.5 as origin,

(b) First three True Moments,

(c) The above value of  $(\beta_1)$  and comment about skewness of the data drawing a rough sketch of the distribution.

3. Given below is a demand schedule, where X is the price per unit in rupees and Y is the quantity of a good in thousand units:

X	10	12	15	17	20	21	25	30
Y	50	45	42	40	39	37	35	34

(a) Estimate the demand function and predict the demand when price is Rs. 32 per unit.

(b) Estimate the average revenue function and predict the average revenue when 55,000 units were demanded.

(c) Determine the Correlation Coefficient between X and Y.

4. Distinguish between:

(a) Discrete and Continuous variables,

(b) Primary and Secondary data,

(c) Time series and Cross section data,

(d) Measures of location and Measures of dispersion,

(e) Sample survey and Census.

5. (a) A set of 10 observations found to have a mean 55 and variance 15. Later, on checking, it is discovered that 2 observations 45 and 55 were mistakenly recorded. While the correct observations were 40 and 60. Determine the correct Mean and Variance.

(b) State and explain the characteristics of Mean and Standard Deviation.

6. (a) State and explain the steps involved in the construction of index number.

(b) The group indices and corresponding weights obtained from a household income and expenditure survey are given as under. Construct the case of Living Index:

Group	Index	Weight
Food	115	0.20
Rent	110	0.25
Utilities	125	0.15
Education	130	0.20
Clothing	135	0.10

7. Given below are the population of a town in millions during the past five census. Interpolate the population of the town during 1991:

<b>Years</b>	1951	1961	1971	1981	1998
<b>Population (in millions)</b>	2.35	2.47	2.59	2.72	3.00

8. (a) Let  $X$  be a normally distributed random variable having a mean 57 and standard deviation 8, determine the following probabilities:

- (i)  $P(X > 50)$
- (ii)  $P(X > 62)$
- (iii)  $P(50 < X < 67)$
- (iv)  $P(62 < X < 70)$
- (v)  $P(X = 60)$

(b) Three digit number are formulated using the digits 0, 1, 2, 3, 4, 5. No digit repeated in the same number. Find the probability of:

- (i) Even numbers,
- (ii) Odd numbers,
- (iii) A number greater 450.

9. (a) Write a short note on Time Series Analysis.

(b) Fit a Second Degree Trend and calculate the Trend Values for the following data:

<b>Years</b>	1980	1981	1982	1983	1984	1985	1986	1987
1988								
<b>Sales (million Rs.)</b>	25	28	33	38	45	53	64	77

1. Weekly wages in US\$ for 250 high school teachers of a small town in Minnesota are given as under:

<b>Wages (US\$)</b>	<b>No. of Teachers</b>
Less than 200	15
201 – 225	28
226 – 250	45
251 – 300	55
301 – 350	52
351 – 375	30
376 & above	25

- (a) Determine the Mode and Median wages.
- (b) If 35% of the teachers are exempted from tax, what is the lowest limit of wages to be taxed.
- (c) Determine the variability of the data using a suitable Measure of Dispersion.
- (d) Approximately what percent of teachers draw more than \$300 per week.
2. First three moments about the 14 as origin are 2, 20 and 104 respectively determine:
- (a) First three Moments about the Mean, and the Mean,
- (b) First two Moments about:
- (i) 15 as origin,
- (ii) Zero as origin.
- (c) Comment about the skewness of the data.

3. Quarterly sales receipts of an ice-cream factory for 3<sup>rd</sup> quarter of 1994 through 3<sup>rd</sup> quarter of 1997 are given in million rupees. Fit a Parabolic Trend and estimate the Seasonal Component:

<b>Sales (million Rs.)</b>			110	115	120	125	125	120	110
135	130	125	120	140	135				

4. (a) What are the desirable characteristics of a good Sample. Describe an appropriate Sampling Methodology if you are to draw a sample from the Karachi City.
- (b) Construct Marshal's and Fischer's indices for 2000 which base 1995 from the following:

Quantity	Items	1995		2000
		Price	Quantity	Price
12	A	25	10	30
20	B	20	15	22
15	C	10	12	15
10	D	12	8	15
25	E	15	20	18

5. (a) Given below the consumption expenditure (X) and family's income (Y) in US\$ for 8 households:

X	180	270	260	190	250	240	270	230
Y	200	300	300	220	290	280	280	250

Construct an appropriate Regression and predict the consumption of a family with an income \$310. Also determine the Correlation Coefficient.

- (b) Assign ranks to the values of X and Y in part (a) above and determine the Rank Correlation.

6. (a) Given below are the values of some function corresponding to given values of  $x$ . Interpolate, using an appropriate method, the value of function for  $x = 8$ .

$x$	5	10	15	20	25
$f(x)$	6.859	4.587	4.073	3.850	3.725

(b) Draw the graph line for the data given in part (a) above, and estimate  $f(x)$  for  $x = 8$ .

7. (a) State the advantages and limitations of diagrammatic and graphic presentation of statistical data.

(b) Construct Histogram, Frequency Polygon and Less Than Ogive for the data given below:

Marks	1 – 4	5 – 8	9 – 12	13 – 16	17 – 20	21 – 24	25 – 28
Frequency	2	5	8	10	6	3	1

8. (a) Find the probability of randomly selected 8 apples and receiving 25% bad from a basket containing 15 apples of which 5 are spoiled.

(b) Pakistan and New Zealand are going to play a series of one-day international. The team that wins 3 games wins the series (ties not allowed). If Pakistan's probability of winning a single is 0.6, what is the probability that Pakistan wins the series?

9. Write short notes on the following:

- (a) Secondary data,
- (b) Organisation of a sample survey.

-----

**Annual Examination 2001**

1. Management of Kentucky Friend Chicken operating in Karachi pays their lower grade employees in US\$ on monthly basis as under:

<b>Wages</b>	<b>No. of Workers</b>
--------------	-----------------------

120 – 135	10
135 – 150	22
150 – 165	25
165 – 180	32
180 – 195	48
195 – 210	33
210 – 225	22
225 – 240	16
240 – 255	12

(a) Determine the Mean and Variance of wages.

**(10)**

(b) Determine the Geometric Mean of wages.

**(7)**

(c) If workers drawing more than  $(X=\sigma)$  wages are charged to tax, what is the lowest limit of wage to be taxed?

**(3)**

2. For 10 observations given the following:

$$\begin{aligned}\sum x &= 140 \\ \sum x^2 &= 2118 \\ \sum x^3 &= 34130\end{aligned}$$

Determine the following:

(a) Three Absolute Moments,

**(3)**

(b) Three True Moments,

**(6)**

(c) First two Raw Moments about 15 as origin, **(8)**

(d) What are the values for the Mean and Variance. **(3)**

3. KESC bills its consumers on progressive rate. For a sample of five households the units consumed (x) and the amount of bill in Rs. (y) are given below:

x	300	500	600	700	1000
y	350	1200	1500	2000	3500

(a) Find an equation to estimate the amount of bill as a function of units consumed. **(8)**

(b) Determine the simple correlation. **(5)**

(c) Estimate the amount of bill for a consumer which consumed 550 units. **(2)**

(d) State and prove the characteristic feature of Correlation Coefficient with respect to the addition and multiplication by a constant. **(5)**

4. **(a)** Describe Sampling Methodologies for:

(i) A small population like that of a steel mill,

(ii) A big population like that of the Karachi city,

(iii) A very big population like that of Pakistan. **(12)**

**(b)** Discuss the relative merits and demerits of a Census and Sample Survey. **(8)**

5. **(a)** A shipment of 15 monitors contain 4 defective monitors. In how many ways a cyber café purchases 5 of these monitors and



receives at most 2 defective monitors.  
)

(10)

(b) In a single throw of an unbalanced dice the probability an even number to occur is  $\frac{5}{8}$ . If the dice is rolled 6 times, what is the probability of getting an odd number atleast 4 times.

(10)

6. (a) What is an Ideal Index Number? Explain how the base of an index number series can be shifted to a more recent period?

(8)

(b) Construct Paasche's and Marshal's indices for 2001 with 2000 as base from the following data:(12)

Item	Price		Quantity	
	2000	2001	2000	2001
A	15	18	10	8
B	10	12	8	10
C	20	25	5	4
D	22	25	12	15
E	25	30	6	10

7. Fit a Straight-line Trend to the data given below for the period 1991 to 2001. Also isolate the Seasonal Component using a multiplicative model:

(20)

220   225   232   247   249   250   254   258   265   275   300

8. (a) Given below are the values of a function for some specified values of x, interpolate the value of the function for x = 11:

(12)

x      10      12      15      20      23

$f(x)$     1000    1728    3375    8000    12167

(b) Given:

$x$     10    12    18    20    23    25    28

Determine the following:

(i)  $\sum x^2$ ,

(ii)  $\sum (x - \bar{x})^2$ ,

(iii)  $\sum (x + 5)^2$ ,

(iv)  $\sum (3x^2 - 5x + 10)$ .

(8)

9. Write short notes on any two of the following:

(a) Skewness,

(b) Sheppard's correction,

(c) Principle of least square.

(20)

-----  
**Annual Examination 2002**

1. In a medical study to determine the normal limits of blood cholesterol some 220 adult healthy males were examined with respect to their level of cholesterol. Its distribution is given as under:

Level of Cholesterol (mg.)	No. of Persons
Less than 100	18
100 – 110	22
110 – 115	28

115 – 120	30
120 – 130	35
130 – 145	43
145 – 165	24
165 – 175	12
175 and above	8

(a) Determine the Median and Mode level of cholesterol. **(10)**

(b) Determine the Quartile Deviation (Q.D.) of the data. **(6)**

(c) If median + 2Q.D. are considered to be the normal limits, what are the normal limits and what percent of persons have normal level of cholesterol approximately. **(4)**

2. Marks obtained by a group of 10 students out of 50 are given as under:

40 42 38 45 50 47 36 35 48 45

(a) Determine the first three Raw Moments about 45 as origin. **(7)**

(b) Determine the first three True Moments and comment about Skewness. **(7)**

(c) From the raw moments obtained in (a) above shift the origin of the first two to 40. **(6)**

3. Proficiency in economics (measured by marks obtained in economics) is thought to be influenced by the proficiency in mathematics (measured by marks obtained in mathematics). A group of 10 students obtain marks as follows:

<b>Marks in maths</b>	50	55	58	65	55	65	66	59	65	70
<b>Marks in eco.</b>	75	82	78	85	77	82	88	78	84	90

(a) Construct an appropriate Regression, determine the value of  $r^2$  and interpret your result. **(12)**

(b) Determine the Spearman's Rank Correlation for the Coefficient for the above data. **(8)**

4. **(a)** Explain, using examples, the effect of a change in the units of measurement and scale on the following:

(i) Arithmetic mean, **(4)**

(ii) Standard deviation, **(4)**

(iii) Coefficient of variation. **(2)**

**(b)** What is the lowest limit of cholesterol of the:

(i) top 15% of persons,

(ii) next 12% of persons for the data of Question 1 above. **(10)**

5. **(a)** Two dices are rolled simultaneously and the product of their outcome is recorded:

(i) List the elements of the sample space,

(ii) Find the probability of getting perfect square,

(iii) Find the probability of a number multiple of 6. **(5)**

**(b)** A coin biased so that the probability of a head is  $\frac{2}{3}$ . Find the probability of exactly three heads in 5 tosses of that coin.

**(5)**

(c) Find the probability of three digit even number greater than 450, from the digits 0, 1, 2, 3, 4, 5 and 6 (no digit repeat in the same number).

(5)

(d) From the letters of the word 'STATISTICS', find the probability of all words starting and finishing with the letters.

(5)

6. (a) Given below is a schedule of prices and the profit earned at each of these prices of a product. Interpolate, using a suitable method, the profit from a firm when the price is Rs.

18:

(12)

Price (Rs.)	10	13	15	20
Profit (million Rs.)	35	45	48	55

(b) Not available.

7. (a) Construct Fisher's and Marshal's indices for 2002 with 2001 as base from the following:

(10)

Items	Price		Quantity	
	2001	2002	2001	2002
A	20	15	30	25
B	25	30	10	8
C	18	20	20	25
D	30	25	25	30
E	15	18	22	28

(b) Construct 4-year moving averages centred for the following sales in million rupees for the period 1992 to 2002: (10)

220 225 232 247 250 255 258 265 278 300 325

8. Write short notes on any two of the following:

**(20)**

- (a) Time series analysis,
- (b) Organisation of sample survey,
- (c) Sources of secondary data.

-----