UNIVERSITY OF THE PUNJAB

First Semester   2017

Examination: B.S. 4 Years Programme

PAPER: Statistics-I   TIME ALLOWED: 30 mins.
Course Code: STAT-101 / STT-11314   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

OBJECTIVE

Q1. Read the following items carefully and encircle the correct option listed below at each item.

(One mark for each)

i) In an experiment to determine if antibiotics increase the final dressed weight of cattle, the following were measured on each animal in the study.

sex, initial weight, weight gain, grade of meat.
where grade is recorded as (A, B, or C). The scales of measurement of these variables are:
a) Nominal, ratio, interval, nominal   c) Nominal, ratio, ratio, nominal
b) Nominal, ratio, ratio, ordinal   d) Ordinal, ratio, ratio, ordinal

ii) A random sample of 500 households in Vancouver was selected and several variables are recorded for each household. **Which of the following is NOT CORRECT?**
a) Household total income is a ratio scaled variable.
b) Household income (which averages about $35,000) and was rounded to the nearest $100 can be treated as a continuous variable even though it is "discrete".
c) Socioeconomic status was coded as 1=low income, 2=middle income, 3=high income and is an interval scaled variable.
d) The primary language used at home is a nominal scaled variable.

iii) A financial analyst's sample of six companies' book value were $25, $7, $22, $33, $18, $15.
The sample mean and sample standard deviation are (approximately):
a) 20 and 8.9 respectively.   c) 120 and 79.2 respectively.
b) 20 and 8.2 respectively.   d) 120 and 8.9 respectively.

iv) In general, which of the following statements is **FALSE?**
a) The sample mean is more sensitive to extreme values than the median.
b) The sample range is more sensitive to extreme values than the standard deviation.
c) The sample standard deviation is a measure of spread around the sample mean.
d) The sample standard deviation is a measure of central tendency around the median.

v) Earthquake intensities are measured using a device called a seismograph which is designed to be most sensitive for earthquakes with intensities between 4.0 and 9.0 on the open-ended Richter scale. Measurements of nine earthquakes gave the following readings:
    4.5   L   5.5   H   8.7   8.9   6.0   H   5.2
where L indicates that the earthquake had an intensity below 4.0 and a H indicates that the earthquake had an intensity above 9.0. **The median earthquake intensity of the sample is:**
a) Cannot be computed because all of the values are not known
b) 8.70   c) 5.75   d) 6.00

P.T.O.

vi) **Which of the following statements is NOT true?**
a) In a symmetric distribution, the mean and the median are equal.
b) The first quartile is equal to the twenty-fifth percentile.
c) In a symmetric distribution, the median is halfway between the first and the third quartiles.
d) The median is always greater than the mean.

vii) **Which of the following is a function of Statistics?**

a) Condensation   b) Comparison   c) Forecasting   d) All of these

viii) **Which of the following is a limitation of Statistics?**

a) Statistical results can be generalized for the population   c) Statistical laws are exact.
b) Statistics does not study individuals.   d) None of these

ix) **........ is the one, which is collected by the investigator himself for the purpose of a specific inquiry or study. Such data is original in character and is generated by survey conducted by individuals or research institution or any organisation.**

a) Primary Data   b) Secondary Data   c) Categorical Data   d) Spatial Data

x) **Which of the following is not a source of secondary data?**

a) Reports and Official Publications   c) Private Publications
b) Semi-Official Publications   d) Data collected through mailed questionnaire

UNIVERSITY OF THE PUNJAB

PAPER: Statistics-I   **TIME ALLOWED: 2 hrs. & 30 mins.**
Course Code: STAT-101 / STT-11314   **MAX. MARKS: 50**

Attempt this Paper on Separate Answer Sheet provided.

SHORT QUESTIONS

Q2. **Differentiate between the following terms:**   *(4 marks each)*

i)   Descriptive and Inferential Statistics
ii)   Primary and Secondary Data
iii)   Regression and Correlation
iv)   Standard Deviation and Variance
v)   Absolute Dispersion and Relative Dispersion.

SUBJECTIVE

Q3   **(a)**   From a random sample of voters in Rawalpindi, Islamabad, voters   *(06)*
are classified by age group, as shown by the following data.

Age Group	18-24	24-30	30-36	36-42	42-48	48-54	54-60	60-66	66-72
Frequency	17	22	26	35	33	30	32	21	15

Calculate Mean, Median and Mode of the above data.

(b)   Following table gives the birth rates and death rates per thousand of a few countries.   *(04)*
Represent them in a multiple bar chart.

Country	Birth Rate	Death Rate
India	33	24
Japan	30	19
New Zealand	18	8
Pakistan	25	16

Q4

Compute first four Mean moments of the following Data. Also compute coefficient of   *(10)*
skewness and kurtosis.

Weight (grams)	65-84	85-104	105-124	125-144	145-164	165-184	185-204
Frequency	9	10	17	10	5	4	5

Q5.

(a)   The following are the results of height (X) and weight (Y) of 12 persons.   *(06)*

$$\sum x = 766, \quad \sum y = 1700, \quad \sum xy = 109380, \quad \sum x^2 = 49068, \quad \sum y^2 = 246100$$

Estimate a regression line of Y on X from the above data. Also find the correlation coefficient between height and weight.

(b)   Consider the following data   *(04)*

Years	Wheat Price	Wheat Qty.	Rice Price	Rice Qty.
1974	5.3	1410	5.2	350
1975	5.6	1135	5.0	400

Construct Fisher's Ideal Quantity Index Number for 1975.

UNIVERSITY OF THE PUNJAB

Roll No. ....

Second Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Statistics-II
Course Code: STAT-103, STT-12314

TIME ALLOWED: 30 mins.
MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

Note: Attempt all questions. Use of Scientific Calculators and Statistical tables is allowed but exchange of anything i.e. calculators etc. is not allowed.

Section-I

Q.1 Encircle the correct answer in the following. **(10)**

I. P (A) = 0.6, P (B) = 0.5, which of the following statement is true?
   (a) A and B are mutually exclusive
   (b) A and B are not mutually exclusive
   (c) A and B are independent
   (d) A and B are dependent

II. If X is a random variable and a and b are constant then Var (aX + b) is equal to
   (a) $a^2 Var(X)+b$
   (b) $a\, Var(X)$
   (c) $a^2 Var(X)$
   (d) $a\, Var(X)+b$

III. The number of ways in which four looks can be arranged on a shelf is
   (a) 4
   (b) 6
   (c) 24
   (d) 12

IV. The parameters of the probability distribution $p(x) = \overset{n}{\underset{x}{C}}\, p^x q^{n-x}$, such that p + q = 1, are
   (a) x and n
   (b) x and p
   (c) x and q
   (d) n and (p or q)

V. In a Poisson distribution
   (a) The mean and variance are equal
   (b) The mean and S.D. are equal
   (c) The mean is greater than the variance
   (d) None of these

VI. In a normal distribution $N(\mu, \sigma^2)$, the semi-inter quartile range is equal to
   (a) $0.7979\, \sigma$
   (b) $0.6745\, \sigma$
   (c) $0.9544\, \sigma$
   (d) $0.9973\, \sigma$

VII. Mean of the distribution, $f(x) = \dfrac{1}{5}$; $5 \le X \le 10$, is
   (a) 25
   (b) 7.5
   (c) 2.5
   (d) 3.75

VIII. If X is a discrete random variable, then the function f(x) is
   (a) A probability density function
   (b) A probability distribution function
   (c) A cumulative distribution function
   (d) None of these

IX. The mean of the uniform distribution $f(x) = \dfrac{1}{n}$, $x = 1,2,3,\ldots\ldots,n$ is
   (a) $\dfrac{n-1}{2}$
   (b) $\dfrac{n+1}{2}$
   (c) $\dfrac{n^2-1}{12}$
   (d) $\dfrac{n^2+1}{12}$

X. The normal distribution will be less spread out when
   (a) The mean is small
   (b) The median is small
   (c) The mode is small
   (d) The S.D. is small

UNIVERSITY OF THE PUNJAB

PAPER: Statistics-II
Course Code: STAT-103, STT-12314

TIME ALLOWED: 2 hrs. & 30 mins.
MAX. MARKS: 50

Attempt this Paper on Separate Answer Sheet provided.
Note: Attempt all questions. Use of Scientific Calculators and Statistical tables is allowed but exchange of anything i.e. calculators etc. is not allowed.

Section-II

Q.2 Define the following: (20)

    (i)   Compound event

    (ii)   Probability density function

    (iii)   Mathematical Expectation

    (iv)   Binomial distribution

    (v)   Bay's Theorem

    (vi)   Harmonic Mean for probability distribution

    (vii)   Multiplication Law of Probability

    (viii)   Conditional Probability distribution

    (ix)   Cumulative distribution function for discrete variable

    (x)   Area between $\mu - \sigma$ and $\mu + \sigma$ for Normal curve.

Q.3 The probability that an industry will locate in city A is 0.7, the probability that it (06) will locate in city B is 0.4, and the probability that it will locate in either A or B or both is 0.8. What is the probability that the industry will locate
    (a) in both cities
    (b) in neither city

Q.4 A random variable X has the probability density function (06)

$$f(x) = \begin{bmatrix} x & for & 0 < x < 1 \\ 2-x & for & 1 \leq x < 2 \end{bmatrix}$$

Find P (0 < X < 2) and P (X < 1.5)

Q.5 From a lot of 15 missiles,6 are selected at random and fired. If the lot contains 4 (06) defective missiles that will not fire, what is the probability that
    (a) All 6 will fire?
    (b) At most one will not fire?

Q.6 Derive the variance of Poisson distribution. (06)

Q.7 The I.Q. of 1000 applicants to a certain college are approximately normally (06) distributed with a mean of 110 and a standard deviation of 15. If the college requires an IQ of at least 90, how many of the applicants will be rejected on this basis regardless of their other qualifications?

UNIVERSITY OF THE PUNJAB

Roll No. ....

Second Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Business Statistics   TIME ALLOWED: 30 mins.
Course Code: STAT-121 / BUS-12133   MAX. MARKS: 10

Attempt this paper on this Question sheet only.
Note: Attempt all questions. Use of Scientific Calculators and Statistical tables is allowed but exchange of anything i.e. calculators etc. is not allowed.

Section-I

Q.1   Encircle the correct answer in the following.   **(10)**

I.   A cumulative frequency distribution is graphically represented by
    (a) Frequency curve   (b) Histogram
    (c) Ogive   (d) Frequency polygon

II.   The median is larger than the arithmetic mean when
    (a) The distribution is positively skewed   (b) The distribution is negatively skewed
    (c) The distribution is symmetrical   (d) The distribution is extremely skewed

III. Symmetrical distribution will always have skewness equal to
    (a) Zero   (b) Negative
    (c) Positive   (d) Close to zero

IV.   The index number for a base year is always
    (a) Zero   (b) Greater than 100
    (c) Less than 100   (d) None of these

V.   Which of the following statements is incorrect about correlation co-efficient?
    (a) In passes through the means of the data   (b) It is symmetrical with respect to X and Y
    (c) It is independent of origin and scale   (d) It is geometric mean between the two regression co-efficient

VI.   If X is a discrete random variable then the function p(x) is:
    (a) A probability distribution function   (b) A probability density function
    (c) A distribution function   (d) None of these

VII. If X is a random variable and a and b are constants, then Var (a+bx) is equal to
    (a) Var (a) + b Var (X)   (b) b^2 Var (X)
    (c) b Var (X)   (d) None of these

VIII. A chi-square value cannot be negative because
    (a) Difference between expected and observed frequencies are squared   (b) The absolute value of the differences is computed
    (c) A negative value would mean that the observed frequency is less than expected frequency   (d) None of these

IX.   When ANOVA for "k" treatments and "n" observations for each treatment is performed, the degrees of free for error is:
    (a) n-k-1   (b) n-k
    (c) (k-1)   (d) nk-k

X.   The values of the regression Co-efficient b and d in the regression lines
$\hat{Y} = a + bX$, $\hat{X} = c + dY$ are
    (a) Both positive   (b) Both negative
    (c) One is negative and other positive   (d) Both (a) and (b) but not (c)

UNIVERSITY OF THE PUNJAB

Attempt this Paper on Separate Answer Sheet provided.

Note: Attempt all questions. Use of Scientific Calculators and Statistical tables is allowed but exchange of anything i.e. calculators etc. is not allowed.

Section-II

Q.2 Define the following: (20)
- (i)   Chi-square variable
- (ii)   Price Index Number
- (iii)   ANOVA table for one way classification
- (iv)   Probability density function
- (v)   Regression
- (vi)   skewness
- (vii)   Absolute measure of dispersion
- (viii) Consumer price index number
- (ix)   Descriptive statistics
- (x)   Median

Q.3 The following data show the amount of phosphates per load of laundry in   (06)
grams. 48, 47, 42, 42, 41, 29, 34, 31, 29, 30, 26, 29
Find mean, median and mode.

Q.4 A study was made by a retail merchant to determine the relation between   (06)
weekly advertising expenditures (X) and average daily sales (Y)

X:	40	20	25	30	48	40	45	35
Y:	64	67	66	80	77	82	76	71

Find the linear regression equation and estimate average daily sale when advertising cost is 50.

Q.5 The following data represent the number of packages of four brands of   (06)
cigarettes sold by market on five randomly selected days:

Brand					
A	21	35	32	28	14
B	35	12	27	41	19
C	45	60	43	56	37
D	32	29	22	30	23

Perform an analysis of variance and draw conclusion at 0.05 level of significance.

Q.6 In an experiment to study the dependence of Hypertension on smoking habits, (06)
the following data were taken on different persons.

	Non smokers	Moderate smokers	Heavy smokers
Hypertension	22	37	31
Non hypertension	49	27	20

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

Q.7 An inquiry into the budgets of the middle class families in a city gave the (06)
following information.

	Food	Rent	Clothing	Fuel	Misc.
Expenses on:	38%	15%	18%	10%	19%
Price (base year):	150	35	80	30	40
Price (current year):	160	40	60	32	45

What changes in cost of living figures in current year as compared with that of base year are seen?

PAPER: Statistics-III	TIME ALLOWED: 30 mins.
Course Code: STAT-201/STT-21314	MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

Part-1 (OBJECTIVE)

Note: Attempt ALL questions. Cutting and overwriting is not allowed in MCQs
All questions carry equal marks.

Q.1.   Some possible answers are given for each of the followings, encircle   (10x1)
       (or tick) the correct answer.

i)   Population parameter may also be ---------- other than which is calculated
       from population data.
       A) Nominal from population   B) A &C
       C) Characteristics beyond which pop. do not exit.   D) Statistic.

ii)   Sampling distribution is the probability distribution of
       A) population variance.   B) sample.
       C) statistic.   D) parameter.

iii)   Central limit theorem is a -------- theorem
       A) population dependent   B) limiting theorem for sample
       C) limiting theorem for population   D) none of the above

iv)   Maximum likelihood estimator is--------
       A) biased.   B) unbiased.
       C) asymptotically unbiased.   D) A&C

v)   Judgment sampling is -------sampling.
       A) probability   B) non probability
       C) A&B   D) None of the above.

vi)   In the comparison of type-I and type-II errors
       A) Type-I error is more dangerous   B) Type-II error is more dangerous
       C) both are dangerous   D) B&C

vii)   Null hypothesis is designed for possible--------
       A) acceptance   B) rejection
       C) both A&B   D) none of the above

viii) Sum of two chi-square variates is
       A) chi- square   B) Double chi-square
       C) Normal   D) None of the above

ix)   Parameter of chi-square distribution is
       A) mean   B) n
       C) A&B   D) None of the above.

x)   F-distribution is
       A) uni-directional   B) two tails
       C) unimodal   D) A&C

UNIVERSITY OF THE PUNJAB

Third Semester   2017

Examination: B.S. 4 Years Programme

Roll No. ....

PAPER: Statistics-III	**TIME ALLOWED: 2 hrs. & 30 mins.**
Course Code: STAT-201/STT-21314	**MAX. MARKS: 50**

Attempt this Paper on Separate Answer Sheet provided.

Part-2 (SUBJECTIVE)

Note: Attempt ALL questions.

Q.3.   **Write short answers to each of the followings.**   **(2x10)**

 i)   What is the object of sampling?

 ii)   Differentiate between target and sampled population.

 iii)   Differentiate between random and simple random sampling.

 iv)   In what sense unbiasedness differs from consistency?

 v)   What is meant by mean square error of an estimator?

 vi)   What is the purpose of goodness- of- fit test?

 vii)   Why is the z-test usually inappropriate as a test statistic when sample size is small?

 viii)   Under what conditions, the paired t-test is used?

 ix)   What do you mean by level of significance?

 x)   What do you mean by p-value?

Q.4.   Suppose a population of $N = -9$ is stratified into 3 strata with the following measurements: **(06)**

Stratum I	$X_{11} = 1$, $X_{12} = 2$, $X_{13} = 4$
Stratum II	$X_{21} = 6$, $X_{22} = 8$
Stratum III	$X_{31} = 11$, $X_{32} = 15$, $X_{33} = 16$, $X_{34} = 19$

If two measurements are drawn from each stratum for the sample, state how many samples of size 6 could be chosen from this population? List these samples and compute the mean for each sample.

Q.5.   The following table shows the distribution of 14-year-old schoolboy intelligence test markings: **(06)**

I.Q	80-89	90-99	100-109	110-119	120-129	130-139	140-149
Number	30	52	75	109	65	42	27

On the assumption that this group is a random sample, estimate the standard error of the mean and explain its usefulness.

Q.6.   Let the simple linear regression be $Y_i = \hat{\alpha} + \hat{\beta} X_i + e_i$, where each Y_i is normally distributed and X_i's are fixed. Then show (prove) that $\hat{\alpha}$ and $\hat{\beta}$ are unbiased estimators of the parameters α and β.   **(06)**

Q.7.   A form of intelligence test was given to random samples of soldiers and sailors in a certain country. The following results were recorded: **(06)**

	Number in Samples	Mean Score	Sample Standard Deviation
Soldiers	332	12.78	2.43
Sailors	615	12.99	2.48

Assume the populations of scores to be normal. What conclusion should be drawn?

Q.8.   To verify whether a course in statistics improved performance, a similar test was given to 12 participants both before and after the course. The original grades recorded in alphabetical order of the participants were 44, 40, 61, 52, 32, 44, 70, 41, 67, 72, 53 and 72. After the course, the grades were in the same order 53, 38, 69, 57, 46, 39, 73, 48, 73, 74, 60 and 78.   **(06)**

Was the course useful, as measured by performance on the test? Consider these 12 participants as a sample from a population.

PAPER: Elementary Statistics   TIME ALLOWED: 30 mins.
Course Code: STAT-211/GEN-21129   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

OBJECTIVE

Q.1: Each Question has four possible answers. Encircle the correct answer.   (10)

(i) Number of books in a library is an example of:
   (a) discrete variable   (b) qualitative variable
   (c) Attribute   (d) continuous variable

(ii) The median of the following data 5,8,7,3,5,8,4,7 is:

   (a) 5   (b) 4   (c) 7   (d) 6

(iii) A numerical quantity calculated from sample data is called:
   (a) Statistic   (b) parameter
   (c) Constant   (d) None of above

(iv) Which of the following is not a measure of Central Tendency :
   (a) Mean   (b) Standard Deviation
   (c) Median   (d) Mode

(v) Sampling error increases by:
   (a) Increasing sample size   (b) decreasing sample size
   (c) Fixing the size   (d) none

(vi) For a Negatively Skewed distribution:
   (a) Mean, mode, median are not equal (b) Mean, mode, median are equal
   (c) Mean is greater than mode   (d) Mean is less than mode

(vii) Observed data organized into rows and columns is called:
   (a) Classification   (b) Tabulation
   (c) Array   (d) Frequency Distribution

(viii) The standard Deviation of -4, -4, -4, -4, -4 is:
   (a) -4   (b) 0   (c) 4   (d) 16

(ix) The parameters of normal distributions are :
   (a) n and p   (b) μ and σ
   (c) p and q   (d) α and β

(x) A deserving player is not selected in the team is an example of:
   (a) Correct decision   (b) Type –II error   (c) Type –I error (d) none

UNIVERSITY OF THE PUNJAB

Attempt this Paper on Separate Answer Sheet provided.

SUBJECTIVE

Q.2: Write short answers to the following questions: (20 marks)
  (i) Define Statistics.
  (ii) Differentiate between Parameter and Statistics.
  (iii) Define Variance.
  (iv) Write properties of the Arithmatic Mean.
  (v) Define correlation.
  (vi) Define Type I Error and Type II Error.
  (vii) Write down two properties of Standard Deviation.
  (viii) What do you mean by relative dispersion?
  (ix) Describe Probability Sampling.
  (x) Define additional law of probability for mutually exclusive events.

QUESTIONS WITH BRIEF ANSWERS   (30 marks)

Q. No.3: The following frequency distribution gives the ages of 80 workers in a factory:

Ages	19-25	26-32	33-39	40-46	47-53	54-60	Total
Number of Workers	6	16	20	21	15	2	100

Compute Mean , Median and Variance.   (10 Marks)

Q. No.4: The data of heights and weights is given below:

Height (X)	65	64	73	63	66	65	60	70	71	69
Weight (Y)	57	59	68	64	68	72	67	73	69	72

Find Regression Line of Y on X and Correlation Coefficient between X and Y. (10 Marks)

Q. No.5: Prepare a frequency table for the following data taking classes as 21-30, 31-40,....

70, 62, 58, 36, 56, 25, 45, 89, 45, 78, 54, 62, 42, 73, 46, 24, 39, 47, 58, 65, 43, 21, 54, 70, 56, 80, 41, 64, 26, 39, 44, 28, 33, 84, 66, 54, 22, 33, 87, 25, 55, 73, 56, 55, 70, 49, 37, 28, 46, 61, 37, 83, 47, 59, 67, 43, 29, 34, 76.

Also draw the Histogram.   (10 Marks)

UNIVERSITY OF THE PUNJAB

Roll No. ....

Fourth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Statistics-IV   TIME ALLOWED: 30 mins.
Course Code: STAT-203 / STT-22314   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

SECTION-I

Part-1 (OBJECTIVE)

**Note: Attempt ALL questions. Cutting and overwriting is not allowed.
All questions carry equal marks.**

Q.1. Some possible answers are given for each of the followings, encircle   (1x10)
      (or tick) the correct answer.

i) The estimates of the common population variance σ^2 may be obtained by
    partitioning the total variation present in the k-samples (of equal size) taken
    from

    a) N populations.   b) K populations.
    c) K normal populations.   d) N normal populations.

ii) If we have to test that the two variables of classification are independent as
     null hypothesis Reject H_0, if the computed value of

    a) $\chi^2 = \chi^2_{a,(r-1)(c-1)}$ b) $\chi^2 \leq \chi^2_{a,(r-1)(c-1)}$

    c) $\chi^2 \geq \chi^2_{a,(r-1)(c-1)}$ d) both a & c

iii) A goodness-of-fit test is a hypothesis test that is concerned with the
      determination whether results of a sample conform to

    a) real distribution   b) a hypothesized distribution
    c) expected distribution   d) sample distribution

iv) In the regression analysis, the explained variation of the dependent
     variable Y is given by

    a) $\Sigma(Y - \bar{Y})^2$ b) $\Sigma(Y - \hat{Y})^2$
    c) $\Sigma(\hat{Y} - \bar{Y})^2$ d) $\Sigma(Y - \hat{Y})$

v) The coefficient of determination in multiple regression is given by

    a) $R^2_{Y.13} = 1 - (SST / SSE)$.   b) $R^2_{Y.13} = 1 - (SSR / SST)$.

    c) $R^2_{Y.13} = 1 - (SSE / SSR)$.   d) $R^2_{Y.13} = 1 - (SSE / SST)$.

vi) when H_1 is $\beta < \beta_0$. the critical region is

    a) $|t| \geq t_{a/2,(n-2)}$ b) $t \geq t_{a,(n-2)}$

    c) $t \leq -t_{a,(n-2)}$ d) not any of a, b, or c

P.T.O.

vii) In one way ANOVA Total df =

    a) Within df + Between df   b) Within df + error df

    c) both a & b   d) Between df+ error df

viii) While dealing with analysis of variance, E_{xy}^2 / E_{xx} having _____ degree of freedom with the term "$E_{yy} - E_{xy}^2 / E_{xx}$" having_____ degree of freedom by applying the F-statistic.

    a) 1, n-k   b)   1, n-k-1

    c) r-1, n-k   d)   r-1, n-k-1

ix) An experiment is planned to

    a) Get maximum information for minimum expenditure in the minimum possible time

    b) Avoid systematic errors

    c) Ignore spurious effects, if any   d)   all of a,b &c

x) A replication is used

    a) to decrease the experimental error b) to increase precision

    c) both a & b   d)   none of the above

UNIVERSITY OF THE PUNJAB

Fourth Semester - 2017

Examination: B.S. 4 Years Programme

PAPER: Statistics-IV
Course Code: STAT-203 / STT-22314

TIME ALLOWED: 2 hrs. & 30 mins.
MAX. MARKS: 50

Attempt this Paper on Separate Answer Sheet provided.

SECTION - II

Part-2 (SUBJECTIVE)

Q.2. **Write short answers to each of the followings.** **(2x10)**

i) Describe Bartlett's test for homogeneity of variances.

ii) What are the assumptions in the application of χ^2-test of goodness-of-fit?

iii) What are the properties of the sampling distribution of b, the estimate of β?

iv) What is the basic difference between simple regression and multiple regression?

v) What is the difference between $R_{1.23}$ and $r_{12.3}$?

vi) What is meant by Analysis of Variance?

vii) Why we use the LSD test in analysis of variance?

viii) What is the use of Duncan's multiple range test?

ix) What are disadvantages of completely randomized design?

x) What are the basic principles of experimental designs?

Q.3. The weights of a random sample of 10 boxes of a particular brand of cereal are 14.2, 13.7, 14.1, 14.3, 14.1, 13.8, 14.4, 14.8, 13.9 and 14.3. Test the hypothesis that $H_0 : \sigma^2 = 0.02$ against the alternative $H_1 : \sigma^2 < 0.02$, using a 0.01 level of significance. **(06)**

Q.4. Three independent samples gave the following results: **(06)**

Size	Observations
5	34, 40, 47, 60, 84.
9	40, 59, 60, 67, 86, 92, 95, 98, 108.
3	46, 93, 100.

Use Bartlett's test to test the hypothesis of equal variances. Let $\alpha = 0.05$.

Q.5. An instructor of mathematics wished to determine the relationship of grades on a final examination to grades on two quizzes given during the semester. Calling X1, X2 and X3 the grades of a student on the first quiz, second quiz and final examination respectively, he made the following computations for a total of 120 students. **(06)**

$\overline{X}_1 = 6.8$ S1 = 1.0   r12 = 0.60

$\overline{X}_2 = 7.0$ S2 = 0.8   r13 = 0.70

$\overline{X}_3 = 74$ S3 = 9.0   r23 = 0.65

Find the least-squares regression equation of X3 on X1 and X2.

P.T.O.

Q.6. In a feeding experiment of some animals, the following results were **(06)** obtained, the numbers in the table being the gains in weight in pounds. The animals were in groups of 3 each.

Groups		I	II	III	IV
	A	7.0	16.0	10.5	13.5
Rations	B	14.0	15.5	15.0	21.0
	C	8.5	16.5	9.5	13.5

Test the hypothesis of no difference in rations, at $\alpha = 0.05$.

Q.7. Carry out the analysis of variance for the following Latin square. **(06)**

V1 (2.3)	V2 (3.0)	V3 (3.3)	V4 (2.5)
V2 (3.1)	V3 (4.1)	V4 (2.4)	V1 (2.4)
V3 (4.3)	V4 (2.5)	V1 (2.1)	V2 (2.9)
V4 (2.6)	V1 (2.0)	V2 (2.4)	V3 (4.4)

UNIVERSITY OF THE PUNJAB

Fourth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Probability and Statistics	**TIME ALLOWED: 30 mins.**
Course Code: STAT-221 / IT-22407	**MAX. MARKS: 10**

Attempt this Paper on this Question Sheet only.

Part-1 (OBJECTIVE)

Note: Attempt ALL questions. Cutting and overwriting is not allowed.
All questions carry equal marks.

Q.1. Some possible answers are given for each of the followings, (10x1) encircle (or tick) the correct answer.

i) Which of the following values can,t be the probability of an event?
  a) 0.43   b) 0
  c) 1.2   d) 1

ii) The two events A and B are mutually exclusive which of the following statements must be true?
  a) $P(A \cup B) = 0$ b) $P(A) + P(B) = 1$
  c) $P(A \cap B) = 1$ d) Not any of a,b,&c

iii) A random variable can assume ____ value with a given probability.
  a) one   b) many
  c) multiple   d) b and c both

iv) The distribution function of a random variable X, denoted by F(x) is defined as
  a) $F(x) = P(X \leq x)$ b) $F(x) = P(X \geq x)$
  c) $F(x) = P(X = x)$ d) not any of a,b,or c

v) For Normal distribution
  a) mean = median   b) mean= mode
  c) both b& c   d) not any of a,b&c

vi) If $P(A)=1-P(B)$, then A and B are _____ events.
  a) Mutually exclusive events   b) complimentary events
  c) exhaustive events   d) both b&c

vii) The random variable X have the values from $0,1,2,3....\infty$ for
  a) Normal distribution.   b) Binomial distribution.
  c) Hypergeometic distribution.   d) Poisson distribution.

viii) The variance of Bernoulli distribution is
  a) npq   b) pq
  c) np   d) nq

ix) For a Standard Normal probability distribution, the mean and standard deviation arel
  a) 1 and 1 respectively   b) 1 and 0 respectively
  c) 0 and -1 respectively   d) 0 and 1 respectively

x) In a Normal distribution with mean μ and standard deviation Ω, mean deviation is equal to
  a) $1\,\Omega$ b) $0.8\,\Omega$
  c) $0.6745\,\Omega$ d) $2\,\Omega$

UNIVERSITY OF THE PUNJAB

Fourth Semester - 2017

Roll No. ....

PAPER: Probability and Statistics
Course Code: STAT-221 / IT-22407

TIME ALLOWED: 2 hrs. & 30 mins.
MAX. MARKS: 50

Attempt this Paper on Separate Answer Sheet provided.

Part-2 (SUBJECTIVE)

Note: Attempt ALL questions.

Q.2.   **Write short answers to each of the followings.**   (2x10)

   i)   Differentiate between descriptive and inferential Statistics.

   ii)   Differentiate between Joint and Marginal probabilities.

   iii)   What is the difference between stem-and –leaf plot and Histogram

   iv)   Write three axioms of probability.

   v)   Write the properties of a Distribution Function for a continuous variable.

   vi)   Write the properties of a Binomial Probability Experiment.

   vii)   Compare conditional probability and relative probability

   viii)   Let a random variable Z follow the Standard Normal Distribution, find $P(0.6 \leq Z \leq 1.67)$ and also $P(-1.67 \leq Z \leq -0.6)$.

   ix)   Let X follow the Uniform Distribution $f(x)=1/(b-a)$, $a \leq x \leq b$ then prove that mean of X be equal to $(a+b)/2$

   x)   A Binomial Distribution tends to become Normal Distribution; discuss.

Q.3.   In a certain college, 4% of the men and 1% of the women are taller than 6 feet. Furthermore, 60% of the students are women. Now if a student is selected at random and is taller than 6 feet, what is the probability that the student is a woman?   (06)

Q.4.   a) Find the value of k so that the function f(x) defined as follows, may be a density function   (06)

   $f(x) = kx,$ $0 \leq x \leq 2$

   $= 0,$ elsewhere

   Also Find also the probability that both of two sample values will exceed 1.

   b) Compute the distribution function F(x).

Q5.   If X is binomially distributed with mean 3.20 and variance 1.152, find the complete binomial probability distribution.   (06)

Q.6.   A car hire firm has 2 cars, which it hires out day by day. The number of demands for a car on each day is distributed as a Poisson distribution with parameter 1.5. Calculate the proportion of days on which neither car is used, and the proportion of days on which some demand is refused.   (06)

Q.7.   In a normal distribution, 31% of the items are under 45 and 8% are over 64. Find the mean and standard deviation of the distribution.   (06)

UNIVERSITY OF THE PUNJAB

Roll No. ....

Fifth Semester   2017
Examination: B.S. 4 Years Programme

PAPER: Parametric and Nonparametric Tests (Theory)   TIME ALLOWED: 30 mins.
Course Code: STAT-301   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.
OBJECTIVE TYPE

1. The formula for Kruskal Wallis test ios based upon
   (a) Means   (b) Deviations   (c) Ranks   (d) Categories

2. Which of the following tests would you less to assess whether there is a significant difference between the mean ranks of two conditions?
   (a) Kruskal Wallis   (b) Spearman's Rho   (c) Wilcoxon   (d) Friedman

3. Nature of data for sign test is
   a) Nominal or Ordinal   b) Two sets of measurements can be matched
   c) Direction of changing given   d) All of the above

4. Non parametric tests are used on which of the following situations?
   a) Sample Size is quite small   b) Assumptions like normality not satisfied
   c) Data is given in either nominal or ordinal scale   d) All of the above

5. Which test is non sensitive?
   a) None parametric   b) Parametric
   c) Both (a) and (b)   d) None of these

6. The range of chi-square variable is
   a) $0 \leq \chi^2 \leq +\infty$ b) $-\infty \leq \chi^2 \leq 0$
   c) $-\infty \leq \chi^2 \leq +\infty$ d) None of these

7. To find the goodness of fit, the test used is
   a) Student's t-test   b) F-test
   c) Chi-square test   d) None of the above

8. The range of t-Statistic is
   a) $0 \leq t \leq +\infty$ b) $-\infty \leq t \leq +\infty$
   c) $-\infty \leq t \leq 0$ d) None of above

9. A passing student is failed by an examiner, it is an example of
   a) Type I error b) Type II error c) Best decision d) All of the above

10. Homogeneity of several variances can be tested by
    (a) Bartlett's test   (b) Fisher's exact test (c) F test   (d) T-test.

UNIVERSITY OF THE PUNJAB

Attempt this Paper on Separate Answer Sheet provided.
SUBJECTIVE TYPE

Q1.   Write short answers on the following questions

a) What is estimation? Write properties of a good estimator.
b) Write the procedure of a non-parametric test which is equivalent to paired samples T-test.
c) What is the Power of the test
d) Write the procedure of Bartlett test
e) Steps of tests of hypothesis   (4+4+4+4+4)

Q2.   In a certain experiment to compare two types of sheep food A and B, the following results of increase in weights were observed:

Sheep No.	1	2	3	4	5	6	7	8
Food A	49	53	51	52	47	50	52	53
Food B	52	55	52	53	50	54	54	53

a) Assuming that the two samples of sheep are independent, can we conclude that food B is better than food A?

b) Examine the case when the same set of eight sheep were used in both the foods.

(8)

Q3.   In an experiment to determine which of three different missile systems is preferable the propellant burning rate is measured. The data after coding, are given below. Use Kruskal Wallis test and a significance level of 0.05 to test the hypothesis that the propellant burning rates are the same for the three missile system.

(10)

Missile System		
1	2	3
24, 16.7, 22.8, 19.8, 18.9	23.2, 19.8, 18.1, 17.6, 20.2, 17.8	18.4, 19.1, 17.3, 17.3, 19.7, 18.9, 18.8, 19.3

Q4. Find the linear regression equation from the following data:

X	65, 50, 55, 65, 55, 70, 65, 70, 55, 70, 50, 55
Y	85, 74, 76, 90, 85, 87, 94, 98, 81, 91, 76, 74

Assuming normality, test the hypothesis

i) $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$;

ii) $H_0 : \alpha = 23$ against $H_1 : \alpha \neq 32$;

at the 0.01 level of significance.

(12)

UNIVERSITY OF THE PUNJAB

Roll No. ....

Fifth Semester   2017
Examination: B.S. 4 Years Programme

PAPER: Design and Analysis of Experiments (Theory)   TIME ALLOWED: 30 mins.
Course Code: STAT-303   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.
OBJECTIVE TYPE

Q1.   Read the following items carefully and encircle the correct option listed below at each item.

(One mark for each)

i)   The basic principles of experimental designs consist of:
   a) Randomization   b) Replication   c) Local Control   d) All of these

ii)   The smallest subdivision of the experimental material is called:
   a) Treatments   b) Experimental Unit   c) Experimental Error   d) None of these

iii)   In a completely randomized design, treatments are assigned to experimental units
   .... at random.
   a) Completely   b) Partially   c) Systematically   d) None of these

iv)   The assumptions under analysis of variance consist of:
   a) Normality and Independence   c) Linearity and Additivity
   b) Both (a) and (b)   d) None of these

v)   The following design provides the maximum number of degrees of freedom for error sum
   of squares:
   a) Completely Randomized Design   c) Completely Randomized Block Design
   b) Latin Square Design   d) None of these

vi)   Multiple comparisons tests are applicable when:
   a) Null Hypothesis about equality of means is rejected
   b) Null Hypothesis about equality of means is accepted
   c) Does not depend upon the rejection or acceptance of Null Hypothesis
   d) None of these

vii)   One can estimate the missing observation through covariance technique by simply
   changing the sign of....
   a) b   b) r   c) Correction Factor   d) None of these

viii)   The efficiency of two experimental designs can simply be measured through ..... of error
   variances.
   a) Addition   b) Subtraction   c) Multiplication   d) Ratio

ix)   Two Latin squares are .... if each letter of one square design occurs exactly
   once with every letter of the other square when they are superimposed.
   a) Orthogonal   b) Factorial Designs   c) Efficient   d) None of these

x)   A contrast is ....combination of treatments.
   a) Linear   b) Exponential   c) Quadratic   d) None of these

PAPER: Design and Analysis of Experiments (Theory)   TIME ALLOWED: 2 hrs. & 30 mins.
Course Code: STAT-303   MAX. MARKS: 50

Attempt this Paper on Separate Answer Sheet provided.
SUBJECTIVE TYPE
SHORT QUESTIONS

Q2.   Explain the following: *(6,3,4,3,4)*

- i)   Random Effects, Fixed Effects and Mixed Effects Models
- ii)   Analysis of Covariance
- iii)   Principles of Experimental Design
- iv)   Assumptions under Analysis of Variance
- v)   Orthogonal Polynomials.

SUBJECTIVE

Q3. a) In an experiment 'k' treatments and 'r' blocks are selected at random from a large number of treatments and blocks. Develop expected mean squares by clearly indicating the assumptions used.

b) Given the following ANOVA for a CR design for four treatments: (6+4)

S.O.V	d.f.	SS
Treatments	3	1.1986
Error	36	1.0323

Test the significance of difference between treatment means by using Duncan's Multiple Range Test when treatment means for four treatments were 1.464, 1.195, 1.325, and 1.66.

Q4. a) Seven treatments arranged in six randomized complete blocks gave the following sum of squares and products

S.O.V.	YY	XY	XX
Blocks	1200	600	200
Treatments	800	300	100
Error	1400	700	600

- i)   Is the regression of Y on X significant at 0.05 level of significance?
- ii)   Construct ANOVA and write the inference.

b) The analysis of Variance for a RCB design produced the table shown below: (6+4)

S.O.V	d.f.	SS	MS	F-Ratio
Treatments	3	28.2	-	
Blocks	5	-	13.80	
Error	-	34.1	-	

Complete the ANOVA table and test the significance of difference among the treatment means.

Q5. a) Derive formula for estimating N missing observations in a LS Design when values are missing in different rows, different columns and different treatments.

b) In an experiment to examine the effects of row spacing on the yield of wheat, 8 row spacing were used and 6 blocks of an experiment were used. The sum of squares for Total, Blocks and Treatments were 2195.48, 617.86 and 1283.65 respectively. Find the relative efficiency of this design with the design in which blocks are ignored.   *(7+3)*

UNIVERSITY OF THE PUNJAB

Roll No. ....

Fifth Semester   2017

Examination: B.S. 4 Years Programme

PAPER: Sampling Techniques (Theory)   TIME ALLOWED: 30 mins.
Course Code: STAT-305   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

OBJECTIVE

Q. No. 1: Encircle the correct option.

1. When the sample survey becomes a census survey, the sampling error becomes ----------
   A. Zero   B. One
   C. Constant   D. None of the above

2. For a population that is not normally distributed, the distribution of the sample means will ------
   A. be negatively skewed   B. approach the normal distribution.
   C. be positively skewed.   D. never approaches normal distribution.

3. The process of drawing a sample from a population is known as --------------.
   A. Census   B. Survey
   C. Sampling   D. None of above

4. Suppose we select every fifth invoice in a file. What type of sampling is this?
   A. Simple Random   B. Cluster
   C. Stratified   D. Systematic

5. Which of the following is not an example of random sampling method:
   A. Simple random Sampling   B. Stratified Sampling
   C. Cluster Sampling   D. Convenience Sampling

6. All possible samples of size 'n' are selected from a population and the mean of each sample is determined. What is the mean of the sample means?
   A. Exactly the same as the population mean   B. Larger than the population mean
   C. Smaller than the population mean   D. Cannot be estimated in advance

7. Sampling in which sampling unit can be repeated more than once is called --------------------
   A. Sampling with replacement   B. Sampling without replacement
   C. Both (A) & (B)   D. None of the above

8. ------------ is the standard deviation of a sampling distribution.
   A. Sample standard deviation   B. Replication error
   C. Meta error   D. Standard error

9. A complete list of all the units in the population is called --------------
   A. Sampling unit   B. Population unit
   C. Sampling frame   D. None of the above

10. If a sample is drawn from each stratum minimizing the probabilities of non-preferred samples, it is known as:
    A. Selection with proportional allocation   B. Controlled selection
    C. Haphazard selection   D. None of the above

UNIVERSITY OF THE PUNJAB

Fifth Semester   2017
Examination: B.S. 4 Years Programme

PAPER: Sampling Techniques (Theory)   TIME ALLOWED: 2 hrs. & 30 mins.
Course Code: STAT-305   MAX. MARKS: 50

Attempt this Paper on Separate Answer Sheet provided.
SUBJECTIVE

Attempt all questions.

Q. No. 2. Write a short note on the following: (5 each)
- i. Inverse Sampling
- ii. Method of controlled selection in Stratified Random Sampling
- iii. Reasons of Stratification
- iv. Application of Systematic Sampling

Q. No. 3. If the loss function due to an error in $\bar{y}$ is $d(\bar{y} - \bar{Y})^2$ and the cost function is   (07)
$C = C_0 + C_1 n$, then show that the most economical value of 'n' in simple
random sampling, ignoring finite population correction is $\sqrt{\frac{dS^2}{C_1}}$.

Q. No. 4. Prove that sample proportion is an unbiased estimator of population   (07)
proportion with variance

$$V_{pst} = \frac{1}{N^2} \sum \frac{N_h^2 (N_h - n_h)}{N_h - 1} \frac{P_h Q_h}{n_h}$$

Q. No. 5. If the terms in $\frac{1}{N_h}$ are not ignored then prove that   (09)

$$V_{ran} = V_{prop} + \frac{1-f}{n(N-1)} \left[\sum_{i=1}^{L} N_h (\bar{Y}_h - \bar{Y})^2 - \frac{1}{N} \sum_{i=1}^{L} (N - N_h) S_h^2 \right]$$

Q. No. 6. Show that the mean of a systematic sample is more precise than the mean of   (07)
a simple random sample if and only if $S_{wsy}^2 > S^2$
where S_{wsy}^2 and S^2 are variance within systematic samples and population
variance respectively.

PAPER: Probability Theory (Theory)   TIME ALLOWED: 30 mins.
Course Code: STAT-307   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.
OBJECTIVE TYPE

Q.1   Tick the correct answer for the following multiple choice questions.

I.   The distribution having the moment generating function $\dfrac{1}{(3-2e^t)}$ can be classified as:

   a. Negative binomial distribution
   b. Geometric distribution
   c. Binomial distribution
   d. None of the above

II.   The probability is p=0.80 that a patient with a certain disease will be successfully treated with a new medical treatment. Suppose that the treatment is used on 40 patients. What is the "expected value" of the number of patients who are successfully treated?

   a. 0.40
   b. 20
   c. 8
   d. 32

III.   Suppose that the probability of event A is 0.2 and the probability of event B is 0.4. Also suppose that the two events are independent. Then $P(A/B)$ is:

   a. $P(A) = 0.2$
   b. $P(A)/P(B) = 0.2/0.4 = 1/2$
   c. $P(A)P(B) = (0.2)(0.4) = 0.08$
   d. None of the above

IV.   A coin is tossed three times. What is the probability that it lands on heads exactly one time?

   a. 0.125
   b. 0.250
   c. 0.333
   d. 0.375

V.   Whenever P(success) $< \frac{1}{2}$ and the number of trials is fewer than ten, the shape of a given distribution will be:

   a. Symmetrical but only if n is large
   b. Positively skewed towards the left
   c. Positively skewed towards the right
   d. Symmetrical

VI.   The moment generating function of a random X is, $M_{x(t)} = \frac{2}{5} + \frac{1}{3}e^{2t} + \frac{4}{15}e^{3t}$. The expected value of X is

P.T.O.

a. 22/15

b. 9/5

c. 17/15

d. 11/5

e.

VII. The probability that a leap year will have 53 Sundays is

a. 1/7

b. 2/7

c. 2/53

d. 52/53

VIII. In the outcomes of a discrete random variable follow a Negative binomial distribution, then their:

a. Mean equals the variance of that discrete random variable

b. Mean less than the variance of that discrete random variable

c. Variance less than the mean of that discrete random variable.

d. Both a and c

IX. Given $P(A) = 0.4$, $P(B) = 0.3$, $P(B/A) = 0.2$. What are $P(A$ and $B)$ and $P(A$ or $B) =$?

a. $P(A$ and $B)=0.12$, $P(A$ or $B)=0.58$

b. $P(A$ and $B)=0.08$, $P(A$ or $B)=0.62$

c. $P(A$ and $B)=0.12$, $P(A$ or $B)=0.62$

d. $P(A$ and $B)=0.08$, $P(A$ or $B)=0.70$

X. Naveen is a high school basket ball player. She is a 70% free throw shooter. That means her probability of making a free throw is 0.70. What is the probability that Naveen makes her first free throw on her fifth shot?

a. $5 (0.70) (0.30)^4$

b. $(0.70)(0.30)^4$

c. $5(0.70)(0.30)^5$

d. None of the above

UNIVERSITY OF THE PUNJAB

Fifth Semester   2017
Examination: B.S. 4 Years Programme   Roll No. ....

PAPER: Probability Theory (Theory)   **TIME ALLOWED: 2 hrs. & 30 mins.**
Course Code: STAT-307   **MAX. MARKS: 50**

Attempt this Paper on Separate Answer Sheet provided.
SUBJECTIVE TYPE

Q.2   Write short answer on the following questions.

a)   Explain what is meant by a Random Experiment, a Sample Space and an Event.

b)   If $P(A) = 0.5, P(B) = 0.3$, and $P(A\ or\ B) = 0.65$, are events A and B are independent?

c)   Explain how negative binomial distribution can be distinguished from binomial distribution.

d)   A coin is known to be unbalanced in such a way that heads only comes up 0.4 0f the time.

   (i)   What is the probability that first head appears on 4th toss?

   (ii)   How many toss would it take, on average, to flip two heads?

e)   What is a geometric experiment and what are its properties?

(20)

Q.3.a) Box A contains 5 green and 7 red balls. Box B contains 3 green, 3 red and 6 yellow balls. A box is selected at random and a ball is drawn at random from it. What is the probability that the ball drawn is green?

b)   There are four candidates for the office of the Highway Commissioner. The respective probabilities that they will be selected are 0.3, 0.2, 0.4, and 0.1and the probabilities for a   project approval are 0.35, 0.85, 0.45 and 0.15, depending on which of the 4 candidates are selected. What is the probability of the project getting approved?

(5+5)

Q.4   Compute mean, variance, coefficient of skewness and kurtosis of negative binomial distribution.

(10)

Q.5.a) Derive the Poisson distribution as the limiting form of the binomial distribution, stating clearly the assumptions you make.

b)   Three people each toss a coin and the odd man pays for the coffee. If the coins all show heads or all show tail, they are tossed again. What is the probability that a decision is reached in five tosses or fewer?

(5+5)

UNIVERSITY OF THE PUNJAB

Roll No. ....

PAPER: Statistical Computer Packages   **TIME ALLOWED: 30 mins.**
Course Code: STAT-309   **MAX. MARKS: 10**

Attempt this Paper on this Question Sheet only.
OBJECTIVE TYPE

Q1.   Read the following items carefully and encircle the correct option listed below at each item.

(One mark for each)

i)   **What is the advantage of using SPSS over calculating statistics by hand?**
a) It equips you with a useful transferable skill.
b) It reduces the chance of making errors in your calculations.
c) Many researchers use SPSS as it is a recognized software package.
d) All of the above.

ii)   **In SPSS, what is the "Data Viewer"?**
a) A table summarizing the frequencies of data for one variable.
b) A spreadsheet into which data can be entered.
c) A dialog box that allows you to choose a statistical test
d) A screen in which variables can be defined and labeled

iii) **How is a variable name different from a variable label?**
a) It is shorter and less detailed.   c) It is longer and more detailed.
b) It is abstract and unspecific.   d) It refers to codes rather than variables.

iv)   **How would you use the drop-down menus in SPSS to generate a frequency table?**
a) Open the Output Viewer and click: Save As ➔ Pie Chart
b) Click on: Analyze; Descriptive Statistics ➔ Frequencies
c) Click on: Graphs; Frequencies ➔ Pearson
d) Open the Variable Viewer and recode the value labels

v)   **In which sub-dialog box can the Chi Square test be found?**
a) Frequencies ➔ Percentages   c) Crosstabs ➔ Statistics
b) Bivariate ➔ Pearson   d) Gender ➔ Female

vi)   **If the null hypothesis is true, then $p < 0.05$ means:**
a) The obtained result is not due to chance.
b) The obtained result is a fairly important effect.
c) The obtained result is likely to occur by chance 95% of the time.
d) The obtained result is likely to occur by chance less than 5 times in a hundred.

vii) **If you achieved a p-value of 0.04 on a two-tailed test, what would the equivalent one-tailed p-value be?**
a) 0.2   b)0.002   c)0.06   d)0.02

viii)   **Of the below non-parametric tests, which relies on the calculation of ranks?**
a) Mann Whitney   b) Spearman's Rho
b) Wilcoxon sign rank test.   d) All of the above

ix)   **One-way ANOVA is used when:**
a) analyzing the difference between more than two population means
b) analyzing the results of a two tailed tests
c) analyzing the results from a large samples
d) analyzing the results between two population means

x)   **Which of the following tests are used to test for normality?**
a) A t-test and an ANOVA test   c) An Empirical CDF test and an F-test
b) A Chi-Square test and a Lilliefors test   d) A Quantile-Quantile plot and a p-value test

UNIVERSITY OF THE PUNJAB

Fifth Semester   2017
Examination: B.S. 4 Years Programme

PAPER: Statistical Computer Packages
Course Code: STAT-309

TIME ALLOWED: 2 hrs. & 30 mins.
MAX. MARKS: 50

Attempt this Paper on Separate Answer Sheet provided.
SUBJECTIVE TYPE

SHORT QUESTIONS

**Q2.   Define the following:**   *(4 marks each)*

a)   Enlist various graphical and numerical statistical method to assess the normality of a data
b)   Note on EDA (Explanatory Data Analysis) tools in SPSS
c)   Give detail about non-parametric tests which are equivalent to independent and paired samples T-test.
d)   Categorize the scales of the measurement (nominal, ordinal, interval and ratio) regarding metric and non-metric variables.
e)   Write short note on SPSS DATA editor window.

SUBJECTIVE

Note: You are required to justify your answers in this section by stating hypothesis, statistics, reference of graphs etc. on your answer sheet.

Q3: a) The following output generated through SPSS from a data containing the Beginning salaries of employees in a company.

		Statistic	Std. Error
Mean		17016.09	361.510
95% Confidence Interval for Mean	Lower Bound	16305.72	
	Upper Bound	17726.45	
5% Trimmed Mean		16041.71	
Median		15000.00	
Variance		61946945	
Std. Deviation		7870.64	
Minimum		9000	
Maximum		79980	
Range		70980	
Interquartile Range		5168	
Skewness		2.853	.112
Kurtosis		12.390	.224
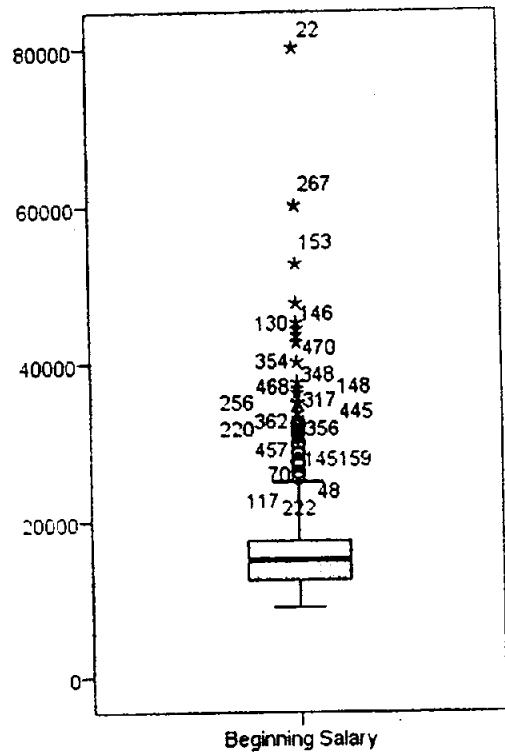
Beginning Salary

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	Df	Sig.
Beginning Salary	.252	474	.000	.715	474	.000

a. Lilliefors Significance Correction

Study carefully the above output and answer the following questions: (2+2+2)

i) **What are the P-values of Kolmogorov-Smirnov and Shapiro-Wilk tests?**
ii) **Does this data set follow the normal distribution?**
iii) **Set null and alternative hypothesis and accept or reject the normality of the data at 1% level of significance**

b) **A sample of individuals applying for factory positions at Davis Enterprises revealed the following scores on an eye perception test (X) and a mechanical aptitude test (Y):**

Spearman's rho	Eye Perception	Correlation Coefficient	1.000	-.191
		Sig. (2-tailed)		.596
		N	10	10
	Mechanical Aptitude	Correlation Coefficient	-.191	1.000
		Sig. (2-tailed)	.596	.
		N	10	10

i) **Compute the coefficient of rank correlation.**
ii) **At the 0.05 significance level can we conclude that the correlation in the population is different from 0?**   (2+2)

Q4: a) The annual report of an institute gave the particulars classification of their students smoking habits (Never Smoked, Past Smoker, and Current Smoker) and their class ranks (freshman, sophomore, junior, senior). The following output is generated from the recorded data using SPSS.

			Class Rank				
			Freshman	Sophomore	Junior	Senior	
Smoking Habit	Never Smoked	Count	108	68	60	47	283
		Expected Count	102.2	65.4	69.1	46.3	283.0
		% within Smoking Habit	38.2%	24.0%	21.2%	16.6%	100.0%
		% within Class Rank	77.7%	76.4%	63.8%	74.6%	73.5%
		% of Total	28.1%	17.7%	15.6%	12.2%	73.5%
	Past Smoker	Count	10	4	18	4	36
		Expected Count	13.0	8.3	8.8	5.9	36.0
		% within Smoking Habit	27.8%	11.1%	50.0%	11.1%	100.0%
		% within Class Rank	7.2%	4.5%	19.1%	6.3%	9.4%
		% of Total	2.6%	1.0%	4.7%	1.0%	9.4%
	Current Smoker	Count	21	17	16	12	66
		Expected Count	23.8	15.3	16.1	10.8	66.0
		% within Smoking Habit	31.8%	25.8%	24.2%	18.2%	100.0%
		% within Class Rank	15.1%	19.1%	17.0%	19.0%	17.1%
		% of Total	5.5%	4.4%	4.2%	3.1%	17.1%
Total		Count	139	89	94	63	385
		Expected Count	139.0	89.0	94.0	63.0	385.0
		% within Smoking Habit	36.1%	23.1%	24.4%	16.4%	100.0%
		% within Class Rank	100.0%	100.0%	100.0%	100.0%	100.0%
		% of Total	36.1%	23.1%	24.4%	16.4%	100.0%

		Value	Approx. Sig
Nominal by Nominal	Contingency Coefficient	.197	.017
N of Valid Cases		385	

i) State the null and alternative hypothesis for the association between class rank and smoking habits of the students
ii) What percentage of senior rank students who never smoked
iii) How many degree of freedom are there?
iv) What is the decision regarding null hypothesis?   (1+1+2+2)

b) We find some depressed people and check that they are all equivalently depressed to begin with. Then we allocate each person randomly to one of three physical exercise groups: no exercise; 20 minutes of jogging per day; or 60 minutes of jogging per day. The following output of allocated groups is generated through SPSS

Levene Statistic	df1	df2	Sig.
1.708	2	21	.205

	Exercise Group
Chi-Square	7.290
df	2
Asymp. Sig.	.026

Kruskal Wallis Test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1306.083	2	653.042	4.414	.025
Within Groups	3107.250	21	147.964		
Total	4413.333	23			

i) Test the homogeneity of exercise groups
ii) Can we conclude that there is no significant difference among the exercise groups   (2+2)

Q5. The administrator of a new paralegal program at Seagate Technical College want to estimate the grade point average in the new program. He thought that high school GPA, the verbal score on the Scholastic Aptitude Test (SAT) and the mathematics score on the SAT would be good predictors of paralegal GPA. The following output is generated from the 9 students recorded data using SPSS:

Correlations

		High School GPA	SAT Verbal	SAT Math	Paralegal GPA
High School GPA		1	.609	.636	.911
	Sig. (2-tailed)		.082	.066	.001
	N	9	9	9	9
SAT Verbal		.609	1	.599	.616
	Sig. (2-tailed)	.082		.088	.077
	N	9	9	9	9
SAT Math		.636	.599	1	.487
	Sig. (2-tailed)	.066	.088		.183
	N	9	9	9	9
Paralegal GPA		.911	.616	.487	1
	Sig. (2-tailed)	.001	.077	.183	
	N	9	9	9	9

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.928[a]	.861	.778	.37513	1.974

a. Predictors: (Constant), High School GPA, SAT Verbal, SAT Math
b. Dependent Variable: Paralegal GPA

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.359	3	1.453	10.326	.014^b
	Residual	.704	5	.141		
	Total	5.063	8			

a. Dependent Variable: Paralegal GPA

b. Predictors: (Constant), High School GPA, SAT Verbal, SAT Math

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-.411	.782		-.525	.622
	SAT Verbal	.002	.002	.170	.759	.482
	SAT Math	-.002	.002	-.215	-.935	.393
	High School GPA	1.201	.296	.945	4.065	.010

a. Dependent Variable: Paralegal GPA

i) Compute the value of coefficient of determination from the above output.

ii) Which independent variable has the strongest correlation with the dependent variable and also test its significance.

iii) Build a regression model using paralegal GPA as dependent variable and all others as independent variables and also conduct a test of hypothesis for the individual regression coefficients.

iv) What Paralegal GPA would you estimate for a student when High School GPA score is 2.81, SAT verbal 320 and SAT Math is zero.   (2+2+4+2)

UNIVERSITY OF THE PUNJAB

Roll No. ....

Fifth Semester   2017
Examination: B.S. 4 Years Programme

PAPER: Applied Statistics
Course Code: STAT-321

TIME ALLOWED: 30 mins.
MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

OBJECTIVE

NOTE: There is no choice. Cutting and overwriting is not allowed.

Section A

Q1: Choose the best answer.   **(10 Marks)**

1.For a distribution Q1=18 and Q3 = 27, Semi Inter quartile range for this distribution is _____

   a) 4
   b) 9
   c) 4.5
   d) None of the above

2. The probability of a sure event is :
   a) 1
   b) 0
   c) – 1
   d) 0.50

3. Changing the value of a score in a distribution will always change the value of _____.

   a) Mean
   b) Median
   c) Mode
   d) All of the above

4.Which of following is not a continuous variable?

   a) Weight of students
   b) Onset age of specific disorder
   c) Number of episodes
   d) GPA of M.Phil students

5. Low Birth weight data is best shown by:

   a) Simple Bar chart
   b) Pie chart
   c) Histogram
   d) Multiple bar chart

                                                              P.T.O.

6. In how many ways can 9 people be seated on a sofa when 6 seats are available?

   a) $^9C_6$
   b) $^6C_9$
   c) $^9P_6$
   d) $^6P_9$

7. If $X = 65$, $M = 59$, $z = -2$, the value of σ will be

   a) 4
   b) 3
   c) 2
   d) 5

8. The arrangement of data according to some common characteristic is called:

   a) Tabulation
   b) Frequency Distribution
   c) Classification
   d) Histogram

9 A specific part out of whole set of possibilities is known as

   a) Proportion
   b) Probability
   c) Unit
   d) All of the above

10. If a sample has variance of 16, the sample standard deviation will be :

   a) 4
   b) 8
   c) 256
   d) None of the above

UNIVERSITY OF THE PUNJAB

PAPER: Applied Statistics   **TIME ALLOWED: 2 hrs. & 30 mins.**
Course Code: STAT-321   **MAX. MARKS: 50**

Attempt this Paper on Separate Answer Sheet provided.

Section B

Q2: Give short answers.   **(10 X 2= 20 Marks)**

1. Differentiate between descriptive and inferential statistics
2. Describe two laws of probability.
3. Explain what happens to standard deviation when a constant of 2 is added with each score.
4. Write down properties of normal distribution.
5. Score of students is distributed with mean 50 and SD=20, find the z-score of student whose actual score is 40.
6. What do you mean by skewness and how it be measured
7. What is cumulative frequency distribution, give an example.
8. Define mutually exclusive and equally likely events in context of probability.
9. What do you mean by Bar-chart and how is it constructed?
10. What is Shepherd's Correction?

Section C   (30 Marks)

Q3: Calculate the Median and Mode for the following data.   **(10 Marks)**

Classes	f
30-35	2
36-41	4
42-47	8
48-53	17
54-59	12
60-65	6

Q.4: What do you mean by scales of measurement, discuss its four types in details.

(10 Marks)

Q5: Assume a normal distribution for each question and solve the following. (10 Marks)

   a) What is the probability of obtaining a z- score less than 0.25?
   b) Find p (X > 70) when $\mu = 50$ and $\sigma = 6$

UNIVERSITY OF THE PUNJAB

Attempt this Paper on Separate Answer Sheet provided.

SHORT QUESTIONS

Q2.　Differentiate between the following terms:　　　　*(9, 6, 5)*

　i)　Fixed Effect, Random Effect and Mixed Effect Model
　ii)　Complete and Partial Confounding
　iii)　Balanced Incomplete Block Design and Partially Balanced Incomplete Block Design

SUBJECTIVE

Q3.　Given the following ANOVA table:　　　　　　　　　(08)

SOV	d.f	Mean Square
A	2	5341.86
B	2	19558.36
AB	4	2403.44
Error	27	675.21
Total	35	

Assuming completely randomized model, interpret the effects of A, B and AB, when

　i)　The levels of both factors are fixed
　ii)　The levels of both factors are random
　iii)　Levels of A are fixed but levels of B are random

Q4.　Suppose that 4 factors, temperature (A), pressure (B), concentration of reactant (c), and stirring rate (D) are studied in a pilot plant to determine their effect on product filtration rate. Suppose now that the $2^4 = 16$ treatment combinations cannot all be run on the same day. The experimenter can run 8 treatment combinations in one day, so a 2^4 design confounded in 2 blocks seems appropriate. It is logical to confound the highest interaction ABCD with blocks.　　(08)

Block 1	Block 2
(1) = 45	a = 71
ab= 65	b = 48
ac= 60	c = 68
bc = 80	d = 43
ad = 100	abc =65
bd =45	bcd = 70
cd = 75	acd = 86
abcd = 96	abd = 104

Q5.　Complete the following table　　　　　　　　　(06)

S.O.V	d.f.	SS	MS	F-Ratio
Block	3	28.44	-	-
A	-	4.92	1.64	-
Error (a)	-	-	-	
Subtotal	-	40.85		
B	-	29.30	-	-
AB	9	-	-	-
Error (b)	-	86.65	-	

Complete the ANOVA table and test the significance of difference among the treatment means.

Q6.　An engineer is studying the mileage performance of 5 types of gasoline additive. In the road test he wishes to use cars as blocks, however, because of a time constraint, he must use an incomplete block design. He runs the balanced design with the 5 blocks that follow. Analyze the data from this experiment and draw conclusions.　　(08)

Additive	Car				
	1	2	3	4	5
1	-	17	14	13	12
2	14	14	-	13	10
3	12	-	13	12	11
4	13	11	11	12	-
5	11	12	10	-	10

UNIVERSITY OF THE PUNJAB

Roll No. ....

Sixth Semester - 2017

Examination: B.S. 4 Years Programme

PAPER: Advanced Experimental Design (Theory)   TIME ALLOWED: 30 mins.

Course Code: STAT-310   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

OBJECTIVE

Q1. Read the following items carefully and encircle the correct option listed below at each item.

(One mark for each)

i) An experiment is called a .... experiment if the treatments consist of all possible combinations of several factors:

a) Latin Square   b) Factorial   c) Latin Square   d) Quasi-Latin Square

ii) A .... effect of a factor is an average change in the factor produced by changing the level of the factor :

a) Main   b) Random   c) Mixed   d) None of these

iii) The joint effect of two or more factors is called:

a) Main Effect   b) Fixed Effect   c) Interaction   d) Random Effect

iv) In confounding, blocks may be constructed through:

a) Sign Table   b) Contrast Method   c) Principle Block Method   d) All of these

v) In .... Confounding, the effect(s) are confounded in each replicate.

a) Complete   b) Partial   c) Mixed   d) None of these

vi) In a design of four factors confounded in 4-blocks, the number of interactions confounded with blocks are:

a) 2   b) 3   c) 4   d) None of these

vii) In Fractional Replication Design, the accuracy of the estimates of the effects ....

a) improves   b) Reduce   c) Remains Constant   d) None of these

viii) When the factors were analyzed according to their importance in the design, the following design is useful:

a) Factorial   b) Fractional Replication   c) Split Plot   d) Incomplete Block Design

ix) In........... Design, blocks do not contain full set of treatments in each and every block.

a) Factorial   b) Fractional Replication   c) Split Plot   d) Incomplete Block

x) In PBIB design, pair of treatments i and j that occur together in the same block are called:

a) First Associates   c) Second Associates

b) Parameters of first kind   d) None of these

PAPER: Advanced Sampling Techniques (Theory)   **TIME ALLOWED: 30 mins.**
Course Code: STAT-312   **MAX. MARKS: 10**

Attempt this Paper on this Question Sheet only.

Q. No. 1: Encircle the correct option.

1. Cluster sampling, stratified sampling and systematic sampling are types of-----------.
   A. random sampling   B. non random sampling
   C. direct sampling   D. indirect sampling

2. ---------- estimator is unbiased when sample size is large.
   A. BLUE   B. Regression   C. Ratio   D. None of above

3. In general the -------------- has a bias of order $\frac{1}{n}$.
   A. Unbiased estimate   B. Ratio estimate
   C. Regression estimate   D. None of the above

4. Like the ratio estimate, the linear regression estimate is designed to increase ----------------- by the use of an auxiliary variate x_i that is correlated with y_i.
   A. Precision   B. Sample size
   C. Variability   D. Confidence width

5. If regression line y on x is linear and passes through origin, then bias of ratio estimate is---------
   A. Low   B. High
   C. zero   D. None of the above

6. The regression estimate is more precise than the ratio estimate unless --------------
   A. $B = R$ B. $B < R$
   C. $B > R$ D. None of the above

7. The non-availability of respondents at home is the type of -------------in surveys.
   A. Sampling Error   B. Standard Error
   C. Response Error   D. Non-response Error

8. In -------------- number of units at each stage are different.
   A. Simple random sampling   B. Stratified random sampling
   C. Systematic sampling   D. Multistage sampling

9. People who are available, volunteer, or can be easily recruited are used in the sampling method called --------------.
   A. Area   B. Probability proportional to size
   C. Multi-stage   D. Convenience

10. The correlation co-efficient between elements in the same ------------- is called Intracluster correlation.
    A. cluster   B. stratum
    C. sample   D. None of the above

UNIVERSITY OF THE PUNJAB

Attempt this Paper on Separate Answer Sheet provided.

Attempt all questions.

Q. No. 2. Write a short note on the following: (5 each)

   i.   Non-response Bias
   ii. Combined Ratio Estimator
   iii. Two-stage Sampling
   iv. Double Sampling.

Q. No. 3. Under the model $y_i = \beta x_i + \varepsilon_i$ where $\varepsilon_i's$ are independent of the x_i and   (08)
$x_i > 0$, show that the ratio estimator $\hat{y}_R = \frac{\bar{y}}{\bar{x}}X$ is Best Linear Unbiased
Estimator (BLUE) for any sample, random or not, selected solely according to
the values of x_i.

Q. No. 4. For simple random sampling in which b_0 is a pre assigned constant, show that   (07)
the linear regression estimate $\bar{y}_{lr} = \bar{y} + b_0(\bar{X} - \bar{x})$ is an unbiased estimate of
$\bar{Y}$ with variance

$$V(\bar{y}_{lr}) = \frac{1-f}{n}(S_y^2 - 2b_0 S_{yx} + b_0^2 S_x^2)$$

Q. No. 5. If S_b^2 is the variance between units in the population and $S_w^2 = AM^g$ $(g > 0)$ (07)
is the variance between elements that lie in the same unit, find the optimum
value of M, the size of the unit.

Q. No. 6. i. If the cluster size varies from cluster to cluster, explain the two   (08)
            methods for estimating the population total Y that are based on a
            simple random sample of n clusters.
        ii. Explain briefly how would you select a PPS sample using Hansen-
            Hurwitz scheme.

UNIVERSITY OF THE PUNJAB

Roll No. ....

Sixth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Multivariate Techniques (Theory)   TIME ALLOWED: 30 mins.
Course Code: STAT-314   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

OBJECTIVE

Q1. Read the following items carefully and encircle the correct option listed below at each item.

(One mark for each)

i. **The multivariate normal distribution is an extension of the:**
   a) uniform distribution   c) normal distribution   e) exponential distribution
   b) chi-square distribution   d) multi-normal distribution

ii. **The off-diagonal elements in the variance-covariance matrix contain:**
   a) regression parameters   c) means   e) covariance estimates
   b) variance estimates   d) none of the above

iii. **Which of the following refers to multivariate analysis?**
   a) a group of statistical techniques used when there are two or more measurements on each element and the variables are analyzed simultaneously
   b) technique that is appropriate when one or more variables can be identified as dependent variables and the remaining as independent variables
   c) a group of techniques in which the whole set of interdependent relationships are examined
   d) technique used to measure the effect of the independent variables on the test units
   e) none of the above

iv. **Generally speaking, multivariate analysis requires:**
   a) fairly large samples.   c) less concern over meeting assumptions than do univariate tests.
   b) small samples.   d) sampling from a population that is not normally distributed.

v. **If X denotes a $(p \times 1)$ column vector of random variables follows a MVN distribution with mean vector μ and variance covariance matrix Σ, then $(X-\mu)' \Sigma^{-1} (X-\mu)$ follows**
   a) Uniform Distribution   c) Multivariate Normal Distribution   e) Chi-Square Distribution
   b) Wishart Distribution   d) None of the above.

vi. **Transpose of a rectangular matrix is a**
   a) Rectangular Matrix.   c) Square Matrix   e) Diagonal Matrix
   b) None of the Above   d) Scalar Matrix

vii. **A process in which a square matrix can be *factored* into the product of matrices derived from its eigenvectors is called?**
   a) Multivariate Analysis   c) Positive Definite Matrix   e) Semi-Positive Definite Matrix
   b) Decomposition of a Matrix   d) None of These

viii. **Which of the following method(s) can be used for decomposition a matrix?**
   a) Spectral   b) Cholesky   c) Singular   d) All of these   e) None of these

ix. $(A \pm B)^T =$
   a) $A^T + B^T$ b) $A^T - B^T$ c) $A^T \pm B^T$ d) $(AB)^T$ e) $B^T A^T$

x. **Any nonzero vector x is said to be a ...... of a matrix A, If there exist a number λ such that $Ax = \lambda x$;?**
   a) Characteristics Vector   b) Characteristic Root   c) Column Vector   d) Unit Vector

UNIVERSITY OF THE PUNJAB

Attempt this Paper on Separate Answer Sheet provided.

SHORT QUESTIONS

Q2. Define the following: (5 marks each)

   i)   Singular-Value Decomposition of a Matrix
   ii)   Eigen-Values and Eigen-Vectors
   iii) Positive Definite and Semi-positive Definite Matrices
   iv)   Wishart Distribution

SUBJECTIVE

Q3. If a random sample of size n is taken from a multivariate normal population of random vector X with mean vector μ and covariance matrix Σ. Find the m.g.f of multivariate normal distribution. *(08)*

Q4. Let $W \sim W_p(f, \Sigma, M)$. If C is any $(p \times q)$ matrix of constants, then show that
$$C'WC \sim W_q(f, C'\Sigma C, MC).$$ *(06)*

Q5. Show that the matrix for the following quadratic form is positive-definite:
$$3x_1^2 + 2x_2^2 - 2\sqrt{2}\,x_1 x_2$$ *(08)*

Q6. The random vector $X' = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 \end{bmatrix}$ has a Multivariate Normal distribution with mean vector μ and covariance matrix Σ given by:
$$\mu = \begin{bmatrix} 4 \\ 0 \\ -3 \\ 7 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 7 & -1 & 0 & 3 \\ & 8 & -2 & 6 \\ & & 12 & 9 \\ & & & 3 \end{bmatrix}$$

Suppose $Y_1' = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ and $Y_2' = \begin{bmatrix} X_3 & X_4 \end{bmatrix}$ are the sub-vectors of X then find

   a) $E\left(Y_1 / Y_2\right)$ b) $Cov\left(Y_1 / Y_2\right)$ *(08)*

UNIVERSITY OF THE PUNJAB

Sixth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Probability Distributions (Theory)   TIME ALLOWED: 2 hrs. & 30 mins.
Course Code: STAT-316   MAX. MARKS: 50

Attempt this Paper on Separate Answer Sheet provided.

SUBJECTIVE

Q.2 Attempt the following questions.

a) Find the mean deviation of the Exponential distribution.   (3)
b) Describe the relationship between the Gamma distribution and Exponential
   distribution.   (2)
c) Find the mean and variance of the rectangular distribution on (-2, 2)   (3)
d) Show that $t^2_{(n)} = F_{(1,n)}$ (3)
e) Find moment generating function of the Student-t distribution.   (3)
f) If $X \sim N(\mu, \sigma^2)$ then find the mean and variance of $Y = aX + b$. (3)
g) Define distribution function and its properties.   (3)

Q.3.a) If "X" follows two parameter Gamma distribution with density function

$$f(x) = \frac{x^{\alpha-1}e^{-\left(\frac{x}{\beta}\right)}}{\Gamma\alpha\,\beta^\alpha} \;;\quad 0 < x < \infty, \alpha > 0, \beta > 0.$$

Derive its moment generating function and use it to find first four cumulants and γ_1
and γ_2.

b) Show that marginal density of "X" in Bivariate Normal distribution follows
univariate Normal distribution.

(7+3)

Q.4 Sate and derive the probability density function of t-distribution.

(10)

Q.5.a) If X has an Exponential distribution given by

$$f(x) = \frac{1}{2}e^{-\left(\frac{x}{2}\right)} \qquad ; 0 \le x \le \infty.$$

What are the mean, variance and m.g.f. of X? Also calculate

$$P(X < 3) \ and \ P(X > 5 | X > 2).$$

b) Five observations x_1, x_2, x_3, x_4, x_5 are drawn at random from the distribution

$$f(x) = 5e^{-5x} \;; 0 \le x \le \infty.$$

What is the distribution of (i) The smallest observation (ii) The largest observation?

(7+3)

UNIVERSITY OF THE PUNJAB

Attempt this Paper on this Question Sheet only.

Objective

Q.1   Tick the correct answer for the following multiple choice questions.

I.   The family of parametric distributions, for which the mean and variance does not exist,

   a.   Normal distribution
   b.   Pareto distribution
   c.   Cauchy distribution
   d.   Weibull distribution

II.   A family of parametric distributions for which moment generating function does not exist is:

   a.   Student t-distribution
   b.   Pareto distribution
   c.   F-distribution
   d.   All the above

III.   For a normal curve, the Q.D., M.D., and S.D. are in the ratio :

   a.   5:6:7
   b.   10:12:15
   c.   2:3:4
   d.   None of the above

IV.   If $X \sim N(8, 64)$, the standard normal deviate Z will be?

   a. $Z = (X - 64)/8$
   b. $Z = (X - 8)/64$
   c. $Z = (X - 8)/8$
   d. $Z = (8 - X)/8$

V.   If $X \sim Expo(5)$ the probability density function X is :

   a. $5e^{-5x}$ for $x > 0$
   b. e^{-5x} for $x > 0$
   c. $5e^{-x}$ for $x > 0$
   d. $\frac{1}{5}e^{-5x}$ for $x > 0$

VI.   The characteristic function of the Cauchy distribution $X \sim C(\alpha, \beta)$ is:

   a. $e^{i\alpha t - \beta t}$
   b. $e^{\alpha t - i\beta t}$
   c. $e^{i(\alpha t - \beta t)}$
   d. $e^{i\alpha t - \beta |t|)}$

(P.T.O.)

VII. If $X \sim N(0,1)$ and $Y \sim (\chi^2/n)$ of the variate $X/\sqrt{y}$ follows:

a. Cauchy's distribution
b. Fisher's F-distribution
c. Student's t-distribution
d. None of the above

VIII. If the sample size $n = 2$, the student's t distribution reduces to :

a. Normal distribution
b. F-distribution
c. Student's t-distribution
d. Cauchy's distribution

IX. The relationship between the mean and variance of χ^2 with n d.f. is:

a. Mean=2 variance
b. 2Mean= variance
c. Mean= variance
d. None of the above

X. Chi square distribution curve in respect of symmetry is:

a. Negatively skew
b. Symmetrical
c. Positively skew
d. None of the above

UNIVERSITY OF THE PUNJAB

Roll No. ....

Sixth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: FORTRAN Computer Language **TIME ALLOWED: 30 mins.**
Course Code: STAT-318 **MAX. MARKS: 10**

Attempt this Paper on this Question Sheet only.
Objective

Q.1 **Tick the correct answer for the following multiple choice questions.**

I. Fortran 77 contains ....intrinsic data types that are built automatically into the language:
a. 2
b. 4
c. 6
d. 8

II. Column 7 through 72 store the Fortran ....
a. Continuation character
b. Comment line
c. Statement labels
d. Program statements

III. The counted loop executes a .... number of lines and the variables controlling the loop cannot be .... during the loop execution.
a. Fixed, predetermined
b. Variable, altered
c. Altered, variable
d. Predetermined, altered

IV. The program assigns the starting value to the' DO' statement is called:
a. LCV
b. Stop
c. Step
d. None of these

V. Micro computer is also called as :
a. Personal computer
b. Laptop computer
c. Note book computer
d. All the above

VI. Machine language consists of......
a. 00
b. 0 & 1
c. 1 & 1
d. 1 & 2

VII. In a computer the raw facts are :
a. Programs
b. Commands
c. Data
d. Attributes

VIII. The computer devices that transforms the input data to the binary language are:
a. Processors
b. Translator devices
c. Output devices
d. Inter faces

IX. Random access memory is also known as :
a. Non volatile
b. Volatile
c. Permanent
d. Secondary

X. One byte is group of ....
a. 2 Bit
b. 4 Bit
c. 6 Bit
d. 8 Bit

UNIVERSITY OF THE PUNJAB

Sixth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: FORTRAN Computer Language
Course Code: STAT-318

TIME ALLOWED: 2 hrs. & 30 mins.
MAX. MARKS: 50

Attempt this Paper on Separate Answer Sheet provided.

SUBJECTIVE

(5x4=20)

Q.2 Write short answer on the following questions.

a) Write a FTN expression corresponding to mathematical expression. **(5)**
  (i) $\sqrt{|ab|}$ (ii) $(a + 3b)(c + d)$ (iii) $Sin^{-1}\sqrt{tan^2|a|}$

b) Indicate which of the following valid declaration / assignment statements are. For those statements that are invalid, provide the reason. **(5)**
  (i) Y+2=X (ii) Dist= Y2-Y1 (iii) NUM=NUM+1 (iv) REAL XYZ

c) Find error if any and give reason. **(5)**
  (i) PRINT I,J,K (ii) 10 GOTO 10 (iii) IF(X=Y) GOTO 10 0
  (iv) X*Y=Z (v) DISTANCE=VELOC*TIME (vi) 10FORMAT (13)

d) Suppose J, K and L contains 10, 20 and 30 respectively. Find the value of each logical expression. **(5)**
  (i) 2*J.EQ.K. AND. K. LE.L (ii) K.EQ.10. AND..NOT.J.LT.L-15

Q.3.a) Write a FTN program to fit straight line equation by the method of least squares?
b) Write a FTN program to find the roots of the quadratic equation.

(6+4)

Q.4.a) Write a FTN program to find the product of two matrices.
b) What are the data types in FTN and explain implicit and explicit data typing?

(6+4)

Q.5.a) Define the LOOPS and its types with structure in FTN ?

b) Write a FTN program to calculate the first n terms of the Fibonacci series?

(4+6)

UNIVERSITY OF THE PUNJAB

Roll No. ....

Sixth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Advance Statistics

Course Code: STAT-323

TIME ALLOWED: 30 mins.

MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

SECTION A (10 Marks)

Q1: Encircle the correct option:

1. A chi-square test of independence is used to evaluate a relationship between two variables. If one variable is classified into 5 categories and other in 3 categories, the chi square statistic will have :
   a. df = 5
   b. df = 6
   c. df = 8
   d. Cannot compute

2. A sample of n = 5 has $s^2 = 20$. What is the value for the estimated standard error?
   a. 3
   b. 2
   c. 9
   d. None of the above

3. If $SS_{between} = 10$, $SS_A = 5$, $SS_B = 4$, SS_{AXB} would be:
   a. 4
   b. 1
   c. 2
   d. None of the above

4. A researcher is conducting an experiment to evaluate a treatment that is expected to decrease the scores for individuals in population which is known to have $\mu = 60$. Which of the following is correct statement for the alternate hypothesis?
   a. $\mu > 60$
   b. $\mu < 60$
   c. $\mu \leq 60$
   d. $\mu \geq 60$

5. An independent measures research study compares 3 treatment conditions using a sample of n = 5 in each treatment. What will be the value of *df* for F-ratio.
   a. 2, 14
   b. 1, 15
   c. 2, 12
   d. None of the above

6. Kruskal Wallis Test can be viewed as an ordinal data alternative to
   a. Single Sample t-Test
   b. Independent measure t-Test
   c. Repeated measure t-Test
   d. Single Factor ANOVA

(P.T.O.)

7. A residual in the regression table is defined as
    a. $Y - \hat{Y}$
    b. Type 1 Error
    c. Sum of Square Error
    d. Regression Sum of Square
8. In a study, first sample had 6 participants and second sample had 9 participants, the *df* for an independent samples *t* test would be:
    a. 15
    b. 13
    c. 12
    d. None of the above
9. The study indicated that the stress had an effect on academic performance when actually it did not is an example of _____.
    a. Type I error
    b. Type II error
    c. Correct Decision
    d. Power of a test
10. A _____ correlation is used to measure the relationship between two variables when both variables are ranked.
    a. Pearson
    b. Spearman
    c. Point bi-serial
    d. None of the above

UNIVERSITY OF THE PUNJAB

Sixth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Advance Statistics
Course Code: STAT-323

TIME ALLOWED: 2 hrs. & 30 mins.
MAX. MARKS: 50

Attempt this Paper on Separate Answer Sheet provided.

SECTION B (5x4= 20 Marks)

Q2: Give short answers

1. Differentiate between critical region and acceptance region in hypothesis testing.
   [04]
2. What is Type-I and Type-II Error
   [04]
3. Describe assumptions of Independent samples t test.
   [04]
4. What are the parametric tests, give examples.
   [04]
5. What is estimated standard error and how is it calculated?
   [04]

SECTION C (3x10= 30 Marks)

Attempt any 3 questions

1. A sample of 6 participants was taken to assess the effect of Therapy A for depression. Depressions scores were assessed before and after therapy. Is there any significant difference in depression scores before and after therapy? Test with $\alpha = 0.05$ [10]

Before	22	20	22	21	20	16
After	11	14	7	14	6	12

2. It is suggested that personality type is related to heart disease. The observed frequencies for all the categories are as follows:

	No Heart Disease	Heart Disease
Type A	15	11
Type B	20	6

Is there a relationship between personality and heart disease? Test with $\alpha = .05$. [10]

3. Following data shows scores of self-esteem and academic performance. Is there any relationship between self-esteem and academic performance? Test with $\alpha = 0.05$ [10]

Self esteem	Academic performance
14	18
13	19
15	17
14	18
13	18
15	19
21	18

4. For the following data, Use Regression Equation to find predicted Y Values for each X value in the data.   [10]

X	Y
42	81
44	83
46	84
47	82
48	81
49	84

PAPER: Statistical Inference-I (Theory)   TIME ALLOWED: 30 mins.
Course Code: STAT-401   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

Q. No. 1: Encircle the correct option.

1. Cramer Rao Inequality gives ------------------------ of the variance of an unbiased estimator.
    A. Unique bound   B. Upper bound
    C. Lower bound   D. None of the above

2. Expected loss is called ------------------.
    A. Error   B. Variance
    C. Risk   D. None of the above

3. The statistic whose conditional distribution is independent of parameter is called -----------
    statistics.
    A. Unbiased   B. Consistent
    C. Efficient   D. Sufficient

4. The estimator which has the variance equal to -------------------- is most efficient estimator.
    A. Mean Square Error   B. Cramer's Rao lower bound
    C. Population mean   D. None of the above

5. The better estimator which also has minimum risk is called ---------------------
    A. Unbiased estimator   B. Admissible estimator.
    C. Consistent estimator   D. Efficient estimator

6. Some statisticians use Mean Square Error for the replacement of variances for the possible
    comparison of ---------------- estimators.
    A. Biased   B. Consistent
    C. Efficient   D. Sufficient

7. If $Eff(T_1/T_2) < 1$, then -------------------- is more efficient.
    A. T_1 B. T_2
    C. both T_1 and T_2 D. None of above

8. If the minimum variance estimator exists, it would essentially be ----------------------
    A. Zero   B. Greater than one
    C. Negative   D. Unique

9. The mean square consistency is also called ----------------------------
    A. Strong   B. Weak
    C. Simple   D. None of the above

10. If the expected value of an estimator is equal to its ------------, the estimator is called unbiased.
    A. Parameter   B. Statistic
    C. Estimate   D. None of the above

UNIVERSITY OF THE PUNJAB

PAPER: Statistical Inference-I (Theory)   **TIME ALLOWED: 2 hrs. & 30 mins.**
Course Code: STAT-401   **MAX. MARKS: 50**

Attempt this Paper on Separate Answer Sheet provided.

Attempt all questions.

Q. No. 2. Write a short note on the following: **(5 each)**

    i.   Relative Efficiency

    ii.   Risk Function

    iii. Drawbacks of Factorization Theorem

    iv. Exponential Family of Distributions

Q. No. 3. Let X has mean p and variance pq and $S_n = X_1 + X_2 + \cdots + X_n \sim B(n,p)$.   **(08)**

    a. Prove that $\frac{S_n}{n}$ is an unbiased estimator of p. show that $E[1 - \frac{S_n}{n}]$ is equal to $1 - p = q$. But $\frac{S_n}{n}(1 - \frac{S_n}{n})$ is not an unbiased estimator of pq. Find the amount of bias?

    b. Prove that $\frac{S_n}{n}(1 - \frac{S_n}{n})$ is asymptotically unbiased estimator of pq.

    c. Prove that $\frac{S_n}{n}$ is consistent in mean square.

    d. Prove that $\frac{S_n}{n}(1 - \frac{S_n}{n})$ is consistent estimator of pq.

Q. No. 4. Given that the pdf of uniform distribution as   **(07)**

$$f(x; \theta) = \frac{1}{\theta} \quad 0 < x < \theta, \quad -\infty < \theta < +\infty$$

Show that $Y_n = \max(X_i)$ is a sufficient statistic.

Q. No. 5. Given that $f(x; \theta) = \frac{1}{\theta} \quad 0 < x < \theta, \theta > 0$ **(07)**

$$= 0 \text{ elsewhere.}$$

Prove that the C.R.L.B is less efficient than the estimator $(\frac{n+1}{n})y_n$, where y_n is the largest value of a random sample of size n and y_n is not unbiased. Suppose its variance is

$$E(y_n^2) - [E(y_n)]^2 .$$

Q. No. 6. State and prove Rao Blackwell theorem.   **(08)**

UNIVERSITY OF THE PUNJAB

Attempt this Paper on Separate Answer Sheet provided.

Q.2. Discuss the following.   **(20)**

   (i) The properties of OLS estimators

   (ii) Stepwise regression

   (iii) CHOW test

   (iv) Functions of Econometrics

   (v) Extraneous information's and its use.

Q.3. For GLR model $\underline{Y} = X_1\underline{\beta}_1 + X_2\underline{\beta}_2 + \underline{u}$, obtain least squares estimates of $\underline{\beta}_1$ and $\underline{\beta}_2$ **(06)**

Q.4. For GLR model $\underline{Y} = X\underline{\beta} + \underline{\epsilon}$, show that ML estimator of β is BLUE.   **(06)**

Q.5. The following calculations are made from 10 sets of observations on Y, X_1 and X_2 **(06)**
which follow

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + u$$

$\Sigma Y = 20, \qquad \Sigma X_1 = 30, \qquad \Sigma X_2 = 40, \qquad \Sigma X^2 = 88.2, \quad \Sigma X_1^2 = 92, \quad \Sigma X_2^2 = 163,$

$\Sigma X_1 Y = 59, \qquad \Sigma X_2 Y = 88, \qquad \Sigma X_1 X_2 = 119$

Test the significance of each regression Co-efficient.

Q.6. Consider the partitioned model $\underline{Y} = X_1\underline{\beta}_1 + X_2\underline{\beta}_2 + \underline{u}$ let b_1 and b_2 be the OLS   **(06)**
estimators of β_1 and β_2 respectively. Suppose that $E\underline{u} = X_2\lambda$. Check the
unbiasedness of b_1 and b_2.

Q.7. Consider SLR $Y = \alpha + \beta X + \epsilon$, Show that $\bar{Y}$ and LS estimator of β are independent.   **(06)**

UNIVERSITY OF THE PUNJAB

Roll No. ....

Seventh Semester 2017
Examination: B.S. 4 Years Programme

PAPER: Basic Econometrics (Theory)　　　　　　　**TIME ALLOWED: 30 mins.**
Course Code: STAT-403　　　　　　　　　　　　　**MAX. MARKS: 10**

Attempt this Paper on this Question Sheet only.

Q.1. State whether the following statements are true, false or uncertain.　　**(10)**

(i)　Ridge Regression estimate is one of the remedial measures when regressors are collinear.

(ii)　In regression analysis, total sum of squares fallows F-distribution.

(iii)　OLS and ML estimators of error variance in a regression model are identical.

(iv)　In general linear regression, regressors are fixed variables for the estimation of parameters by least squares method.

(v)　For comparison of two simple linear regressions students t-statistics is applied for small sample sizes.

(vi)　In general linear regression, variance of the dependent variable is equal to error variance.

(vii)　By the use of extraneous information in linear regression, estimators become more precise.

(viii)　In, $Y = \alpha + \beta X + \epsilon$, If the slope co-efficient β is zero, the intercept parameter is estimated by $\bar{Y}$.

(ix)　Exact models are used in Econometric research.

(x)　In, $Y = \alpha + \beta X + \epsilon$, ML estimators of α, β and error variance are BLUE.

UNIVERSITY OF THE PUNJAB

Roll No. ....

Seventh Semester 2017
Examination: B.S. 4 Years Programme

PAPER: C++ Computer Programming Language   TIME ALLOWED: 30 mins.
Course Code: STAT-405   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

Question1. Objective Part

     a. Fill in the Blanks

         i. Let

            int variable = 2;

            cout<< variable%10;   output will be _____

         ii. Let

            int variable = 5;

            if(variable > 4)

                 cout<< "i am good today";

            else

                 cout<<"i am sick today"

            output will be _____

         iii. Int i = 5;

            while(i > 0){

                 cout<<"wao";

                 i -- ;

            }

            cout<<i;

            How many time "Wao will be printed"   output will be _____

            After loop what is value of i   output will be _____

         iv. If we have a function prototype as following:

            int * function(int);

            Input type is   _____

     b. Choose whether following statement is true or false

         i. Remainder of float variable is obtained using %   (True/False)

         ii. Different function can have same names   (True/False)

         iii. function can return more than one value   (True/False)

         iv. Strings are character arrays   (True/False)

         v. Accessing RAM is faster as compared to Access of Cache   (True/False)

UNIVERSITY OF THE PUNJAB

PAPER: C++ Computer Programming Language   **TIME ALLOWED: 2 hrs. & 30 mins.**
Course Code: STAT-405   **MAX. MARKS: 50**

Attempt this Paper on Separate Answer Sheet provided.

Question2.   Write short note of following   (20) Marks

    a.   Use of NULL character in strings

    b.   Advantages of macros over functions

    c.   arrays

    d.   if statement

    e.   **for** loop vs **do-while** loop

Question3.   Write code to implement following programs   (7+9+7+7) Marks

    a.   User will enter 3 digit number through keyboard, write a program to check whether number is prime or not

    b.   Write a function that inputs array of 50 integers and return their mean and sum using structure

    c.   Write code to find even numbers less than 100.

    d.   Write a code concatenate two arrays.

UNIVERSITY OF THE PUNJAB

Attempt this Paper on Separate Answer Sheet provided.

Subjective

Q1.   Write short answer on the following questions.   **(20)**

    a)   Sampling frame
    b)   Census
    c)   Pretesting of questionnaire
    d)   Literature review

Q.2.   Explain open ended and close ended questions. Their advantage   **(15)**
    and disadvantage and the situation in which they are used.

Q.3.   What do you understand by the term determination of sample   **(15)**
    size which factors are involve in it? Explain in detail.

UNIVERSITY OF THE PUNJAB

Roll No. ....

Seventh Semester 2017
Examination: B.S. 4 Years Programme

PAPER: Research Methodology
Course Code: STAT-406

TIME ALLOWED: 30 mins.
MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

OBJECTIVE

Circle the correct answer.

1. Any numerical value calculated from sample data is called

   a) Error.
   b) Statistics.
   c) Bias.
   d) Standard deviation.

2. Conducting surveys is the most common method of generating

   a) Primary data
   b) Secondary data
   c) Qualitative data
   d) None of the above

3. Non-Sampling error can be reduced by

   a) Increasing sample size.
   b) Decreasing sample size
   c) Grouping the data
   d) None of the above

4. Which of the following is not a type of nonrandom sampling?

   a) Cluster sampling
   b) Convenience sampling
   c) Quota sampling
   d) Purposive sampling

5. What is a good research? The following are correct except

   a) Purpose clearly defined
   b) Research process detailed
   c) Research design thoroughly planned
   d) Findings presented ambiguously

P.T.O.

6. A simple random sample is one in which:

   a) From a random starting point, every nth unit from the sampling frame is selected.
   b) A non-probability strategy is used, making the results difficult to generalize.
   c) The researcher has a certain quota of respondents to fill for various social groups.
   d) Every unit of the population has an equal chance of being selected.

7. What is a cross-sectional design?

   a) The collection of data from more than one case at one moment in time
   b) A study of one particular section of society, e.g. the middle classes
   c) One that is devised when the researcher is in a bad mood
   d) A comparison of two or more variables over a long period of time

8. The person who leads a focus group discussion is called a _____.

   a) Anchor.
   b) Facilitator.
   c) Moderator.
   d) Recorder.

9. A literature Review requires

   a) Planning.
   b) Clear writing.
   c) Good writing.
   d) All of the above.

10. If a study is "reliable", this means that:

   a) The measures devised for concepts are stable on different occasions
   b) It was conducted by a reputable researcher who can be trusted
   c) The findings can be generalized to other social settings
   d) The methods are stated clearly enough for the research to be replicated

UNIVERSITY OF THE PUNJAB

Attempt this Paper on Separate Answer Sheet provided.

SUBJECTIVE TYPE

Q2.   (4 x 5=20)

a) Define time series. Differentiate between discrete and continuous time series.
b) Define strict and weak stationarity.
c) Show that random walk is a non-stationary process.
d) Show that $Y_t = Y_{t-1} + \lambda Y_{t-2} + \varepsilon_t$ is stationary when $-1 < \lambda < 0$.

Q3. The following table gives sample time series: (3 + 7=10)

t	1	2	3	4	5	6	7	8	9	10
Y_t	5	7	7	6	8	9	9	12	13	14

a) Plot the time series. Do you think the plot exhibits stationarity? Provide reasoning.

b) Take first order differences i.e. $W_t = Y_t - Y_{t-1}$. Calculate autocorrelation at lag 1.

Q4. Show that for the ARMA(1,1) process $Y_t = \phi Y_{t-1} + Z_t + \theta Z_{t-1}$ (10)

$$\rho_k = \begin{cases} (1 + \phi\theta)(\phi + \theta) \Big/ (1 + \theta^2 + \phi\theta) & k = 1 \\ \phi\rho_{k-1} & k = 2,3, \cdots \end{cases}$$

Q5. Show that the autocorrelation function of $Y_t = \frac{1}{3}Y_{t-1} + \frac{2}{9}Y_{t-2} + Z_t$ is given by   (10)

$$\rho_k = \frac{16}{21}(\frac{2}{3})^k + \frac{5}{21}(-\frac{1}{3})^k, \quad k = 0,1,2,3, \ldots$$

where $\{Z_t\}$ is a purely random process having zero mean and finite variance.

UNIVERSITY OF THE PUNJAB

Roll No. ....

Seventh Semester   2017
Examination: B.S. 4 Years Programme

PAPER: Time Series Analysis-I   TIME ALLOWED: 30 mins.
Course Code: STAT-411   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

OBJECTIVE TYPE

Q1.   **Tick the correct answer.**   (10 x 1=10)

1) The differencing of time series is used to:
   a) Stabilize the mean of time series
   b) Stabilize the variance of time series
   c) Make the data normally distributed
   d) All of the above

2) A time series is said to be completely random if all the autocorrelations at non-zero lag are
   a) positive
   b) negative
   c) zero
   d) one

3) The autocorrelation at lag zero
   a) is zero.
   b) is 1.
   c) is -1.
   d) Cannot be calculated.

4) Random walk is
   a) a stationary process.
   b) a non-stationary process.
   c) a purely random process.
   d) Both (a) and (c).

5) The autocorrelation function of a moving average process
   a) shows exponential decay.
   b) makes damped sine-cosine waves.
   c) is zero at all non-zero lags.
   d) cuts off to zero for lag greater than order of process.

6) The AR(1) process $Y_t = \phi Y_{t-1} + \varepsilon_t$ is
   a) always stationary.
   b) always non-stationary.
   c) stationary when $|\phi| > 1$.
   d) stationary when $|\phi| < 1$.

7) The autocorrelation at lag 2 for an AR(1) process $Y_t = 0.5 Y_{t-1} + \varepsilon_t$ is
   a) 0
   b) 1
   c) 0.5
   d) 0.25

8) An AR(2) process $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$ is stationary if
   a) $\phi_1 + \phi_2 < 1$
   b) $\phi_2 - \phi_1 < 1$
   c) $|\phi_2| < 1$
   d) All of the above

9) There is a duality between
   a) purely random process and random walk.
   b) random walk and moving average process.
   c) moving average process and autoregressive process.
   d) autoregressive process and random walk.

10) Under certain conditions, a finite order ARMA(p,q) process can be represented in the form of
    a) moving average process.
    b) autoregressive process.
    c) random walk.
    d) Both (a) and (b).

UNIVERSITY OF THE PUNJAB

Attempt this Paper on this Question Sheet only.

Objective Type

1. Operational research approach is

   a) multidisciplinary   b) Scientific   c) initiative   d) all of the above

2. A minimization problem can be converted into a maximization problem by changing the sign of the coefficients in the _____

   a) constraints   b) Objective function   c) Both A and B   d) None of the above

3. If the feasible region of the linear programming problem is empty , the solution is _____

   a) infeasible   b) unbounded   c) alternative   d) None of the above

4. If there are "m" original variables and "n" introduced (slack) variables , then there will be _____ columns in the simplex table

   a) m-n   b) m+n+3   c) m+n-1   d) m+n

5. All the parameters in linear programing model are assumed to be

   a) variables   b) constraints   c) functions   d) None of the above

6. In simplex method , we add _____ variables in the case of "="

   a) slack variable   b) Surplus variable   c) artificial variables   d) None of these

7. As for maximization in assignment problem, the objective is to maximize the _____

   a) profit   b) optimization   c) cost   d) None of these

8. Graphic method can be applied to solve a linear programming problem when there are only _____ variables

   a) One   b) More than one   c) Two   d) Three

9. Which of the following is a valid objective function for a linear programming problem?

   a) Max $5xy$ b) Min $4x + 3y + (2/3)z$ c) Max $5x^2 + 6y^2$ d) Min $(x_1 + x_2)/x_3$

10. Linear programing is a major innovation since _____ in the field of business, decision making , particularly under condition of certainty

    a) Industrial revolution
    c) World war II

    b) World war I
    d) French Revolution

UNIVERSITY OF THE PUNJAB

PAPER: Operation Research   **TIME ALLOWED: 2 hrs. & 30 mins.**
Course Code: STAT-412   **MAX. MARKS: 50**

Attempt this Paper on Separate Answer Sheet provided.

Subjective Type

Q.2.   Write short and concise answer to the following questions:

a)   Define OR and characteristics of OR
b)   Phases of solving a problem using OR
c)   Summarize the steps for solving a LP model
d)   Define the following with reference to LPP
     i) Unbounded solution   ii) Alternative optima
e)   How the problem of minimization can be solved by linear programming

(4 +4+4+4+4)

Q.3.   A jewelry store makes necklace and bracelets from gold and platinum. The store has 18 ounces of gold and 20 ounces of platinum. Each necklace requires 3 ounces of gold and 2 ounces of platinum, while each bracelet requires 2 ounces of gold and 4 ounces of platinum. The demand for bracelets is no more then four. A necklace earns \$300 in profit and a bracelet \$400. The store wants to determine the number of necklaces and bracelets to make in order to maximize profit.

a) formulate a linear programming model for this problem

b) solve it

(10)

Q.4.   Provide the graphical solution to following problem

a)   Min $Z = 4X_1 + X_2$
         Subject to
         $3X_1 + X_2 = 3$
         $4X_1 + 3X_2 \geq 6$
         $X_1 + 2X_2 \leq 4$
         $X_1, X_2 \geq 0$

b)   Maximize $Z = 80X_1 + 100X_2$ subject to

     $X_1 + 2x_2 \leq 720;$ $5X_1 + 4X_2 \leq 1800;$ $3X_1 + X_2 \leq 900;$ $X_1, X_2 \geq 0$

(7+7)

Q.5.   Provide simplex solution to the following problems

     Minimize cost$= 4X_1 + 5X_2$
     Subject to $X_1 + 2X_2 \geq 2.7$
               $3X_1 + X_2 \geq 75$
               $x_1 \geq 0, x_2 \geq 0$

(6)

UNIVERSITY OF THE PUNJAB

Attempt this Paper on this Question Sheet only.

Q.1 Tick the correct answer for each of the followings.　　　　(1 × 10)

i) Maximum Likelihood Estimator (MLE) possesses the following properties of good estimator
　　a) Unbiasedness, Consistency, efficiency & sufficiency.
　　b) Unbiasedness, Consistency, efficiency & Invariance.
　　c) Consistency, efficiency, invariance & sufficiency.
　　d) All of above.

ii) Likelihood function is _____
　　a) Mandatory for MLE.
　　b) Not mandatory for MLE.
　　c) Partially mandatory for MLE.
　　d) None of the above

iii) Moment estimators are
　　a) Sufficient.
　　b) Consistent.
　　c) Efficient.
　　d) All of the above.

iv) Minimum chi-square _____
　　a) Is an estimator.
　　b) Measures minimum association between attributes.
　　c) Measures minimum variance between attributes.
　　d) Both a &c.

v) Likelihood Ratio test uses
　　a) Relational hypotheses.
　　b) Non-relational hypotheses.
　　c) Both a & b.
　　d) None of the above.

vi) Average Sample Number (ASN) is
　　a) a statistic.
　　b) an estimator.
　　c) an averaging tool.
　　d) All of the above.

vii) When sample size is large enough
　　a) OLS is the best choice
　　b) MLE is the best choice.
　　c) Moment estimator is the best choice.
　　d) Minimum chi-square is the best choice.

viii) Neyman-Pearson Lemma provides
　　a) Non-Randomized test.
　　b) Randomized test.
　　c) Most powerful test.
　　d) Admissible test.

ix) Modified minimum chi-square differs from minimum chi-square in respect of
　　a) Numerator.
　　b) Denominator.
　　c) Basic approach.
　　d) None of the above.

x) A test based on test-statistic is classified as
　　a) Randomized test.
　　b) Non-randomized test.
　　c) Admissible test.
　　d) Most powerful test.

UNIVERSITY OF THE PUNJAB

Attempt this Paper on Separate Answer Sheet provided.

Q.2   Write short and brief answer to the followings.

    i) Why moment estimator is considered important even in this modern era of computerization?   **(05)**

    ii) Differentiate between confidence interval and large sample confidence interval.   **(04)**

    iii) What is the utility of likelihood function in MLE?   **(02)**

    iv) What do you mean by 95% confidence interval for a parameter?   **(03)**

    v) Compare the properties of MLE and OLS estimate.   **(04)**

    vi) Why we have to use sequential sampling?   **(02)**

Q.3.   Let $P(X = x) = \binom{x + \theta_2 - 1}{x} \theta_1^{\theta_2} (1 - \theta_1)^x$ $x = 0, 1, \ldots\ldots\ldots, \infty$ **(10)**

Find the moment estimators of θ_1 and θ_2

Q.4.   Let $x_1, x_2, \ldots\ldots, x_n$ denote a random sample from a distribution which has p.d.f. $f(x_i)$ that is positive on only non negative integers. It is desired to test the simple hypothesis $H_o : f(x) = e^{-1}/x!$, $x = 0, 1, 2, \ldots\ldots$ against alternative simple hypothesis $H_1 : f(x) = (1/2)^{x+1}$ $x = 0, 1, 2, \ldots\ldots$ Derive the expression for BCR (Best critical region). Consider the case of n=1 and k=1, k being any positive integer in the expression $\dfrac{L(\theta', x_1, x_2, \ldots\ldots, x_n)}{L(\theta'', x_1, x_2, \ldots\ldots, x_n)} \leq k$ where   **(12)**

$H_o : \theta = \theta'$, $H_1 = \theta = \theta''$.

Find the power of the test for this combination of n and k   when

    (i) H_o is true   (ii) H_1 is true

Q.5.   Obtain the MLE for λ *and* λ^2 from the Rayleigh distribution with p.d.f   **(08)**

$f(x) = (2x/\lambda^2) \exp[-x^2/\lambda^2]$, $0 < x < \infty$ so prove that MLE for this problem is Invariant.

UNIVERSITY OF THE PUNJAB

Eighth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Applied Econometrics (Theory)   **TIME ALLOWED: 30 mins.**

Course Code: STAT-409   **MAX. MARKS: 10**

Attempt this Paper on this Question Sheet only.

SECTION (I)

Q.1.   State whether each of the following statements is true or false OR uncertain.   *(10)*

(i)   When error terms of G.L.R Model are Non-speherical, OLS estimators are BLUE.

(ii)   When high multicollinearity is present then, determinant of correlation matrix approaches to one.

(iii)   Geary OR Run test is used for detection of Heteroskedasticity.

(iv)   In the presence of autocorrelation, OLS regression gives efficient prediction.

(v)   For over specified model, the OLS estimators are unbiased.

(vi)   Pareto curve cannot be used for income distribution analysis.

(vii)   When orthogonal polynomial functions are used as regressors in regression model, then OLS estimator are uncorrelated.

(viii)   Instrumental variable and the regressor should be highly correlated.

(ix)   For unidentified equation, the OLS method gives unique estimates of structural parameters.

(x)   It is not absolutely essential that dummy variables take the values of 0 and 1.

UNIVERSITY OF THE PUNJAB

Eighth Semester - 2017

Examination: B.S. 4 Years Programme

PAPER: Applied Econometrics (Theory)
Course Code: STAT-409

TIME ALLOWED: 2 hrs. & 30 mins.
MAX. MARKS: 50

Attempt this Paper on Separate Answer Sheet provided.

SECTION (II)

Q.2. Define the following: **(20)**

(i) Assumptions of M.L. estimation method for S.L.R. model

(ii) Over identified equation

(iii) Pareto Curve

(iv) Auto correlation

(v) Distributed Lag model

(vi) Multicollinearity

(vii) Statement of Aitken Theorem

(viii) Stochastic Regression

(ix) Rank Condition of identification

(x) Orthogonal Polynomials

Q.3. Discuss the consequences of using OLS estimators in the presence of first order auto-correlated error terms. **(06)**

Q.4. For the model **(12)**

$$y_{1t} = \beta_{12} y_{2t} + \gamma_{11} x_{1t} + u_{1t}$$
$$y_{2t} = \beta_{21} y_{1t} + \gamma_{22} x_{2t} + \gamma_{23} x_{3t} + u_t$$

You are given the following information:

1. The least square estimates of the reduced-form coefficients are

$$\begin{bmatrix} 5 & 10 & 2 \\ 10 & 10 & 5 \end{bmatrix}$$

2. The estimates of variances of the regression coefficients in the first reduced-form equation are 1, 0.5, 0.1.
3. The corresponding covariance's are estimated to be all zero.
4. The estimated variance of the error on the first reduced-form equation is 2.0.

Compute the consistent estimates of structural parameters of the first structural equation.

Q.5. Define Heteroskedasticity and discuss two measures to remove / minimize it. **(06)**

Q.6. Let $\underline{Y} = X\underline{\beta} + \underline{\in}$ such that $\underline{\in} \sim N(0, \sigma^2 V)$ obtain BLUE of σ^2. **(06)**

UNIVERSITY OF THE PUNJAB

Roll No. ....

Eighth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Time Series Analysis-II   TIME ALLOWED: 30 mins.
Course Code: STAT-413   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

Question No.1. Four options are given against each statement, encircle the correct option.   (10×1)

1) The residual autocorrelation coefficients obtained from a sample time series of 250 data points are 0.2, -0.15 and -0.1 at lags 1,2 and 3 respectively. Assuming that the coefficients are approximately normally distributed, what is the value of Box-Pierce statistic?

    a) 18.09   **b)** 37.50   **c)** 18.125   **d)** 53.06

2) When a forecaster uses the _____ method, he or she assumes that the time series components are changing quickly over time.

    a) Maximum Likelihood   **b)** Simple exponential smoothing

    c) Minimum Mean Square   **c)** Box-Jenkins

3) An AR(p) (p is an integer greater than zero) model will have

    a)   An acf and a pacf that both decline geometrically
    b)   An acf that declines geometrically and a pacf that is zero after p lags
    c)   A pacf that declines geometrically and an acf that is zero after p lags
    d)   An acf and a pacf that both are zero after p lags

4) Under Simple exponential smoothing method for forecasting if we want our forecast to be very responsive to the most recent observation, the value of alpha should be;

    a) Small   **b)** Large

    c) Moderate   **d)** The value of alpha does not matter

5) Updating forecasts helps us to minimize the _____ forecast error.

    a) Long-term   **b)** Short-term

    c) Both (a) & (b)   **d)** Neither (a) nor (b)

P.T.O.

6) In the acronymn "SARIMA," what does the "I" stand for?

a) Independence   b) Integrated

c) Intraclass   d) Irreversible

7) What is the name of the process defined by,

$$(1 + 0.6B)(1 - B)Y_t = (1 - 0.9B^3)Z_t$$

a) ARIMA(1,0,3)   b) ARIMA(3,1,1)

c) ARIMA(1,1,1)   d) ARIMA(1,1,3)

8) Yule-walker estimates is another name of _____ estimates.

a) Maximum Likelihood   b) Least Squares

c) Minimum Mean Square   c) Moment

9) Iterative Least square is a procedure to estimate parameters of _____ Process(s).

a) Moving average
b) Autoregressive
c) Mixed Autoregressive and Moving average
d) Both a) and c)

10) Which one of the following would not be classified as a time series technique?

a) Extrapolation of trend curves   b) Simple moving average
c) Box-Jenkins methodology   d) Regression models.

UNIVERSITY OF THE PUNJAB

Eighth Semester - 2017

Roll No. ....

PAPER: Time Series Analysis-II	**TIME ALLOWED: 2 hrs. & 30 mins.**
Course Code: STAT-413	**MAX. MARKS: 50**

Attempt this Paper on Separate Answer Sheet provided.

Note: Attempt all questions.

Question No.2. Answer the following.   (4 × 5 = 20)

a) Discuss the role of ACF & PACF in identification of a time series model.

b) Consider the $SARIMA(1, 0, 0) \times (0, 1, 1)_{12}$ Model and show that

$$X_n(2) = X_{n-10} + \varphi^2(X_n - X_{n-12}) + \theta\varphi Z_{n-11} + \theta Z_{n-10}$$

c) Describe Portmanteau Lack of fit tests..

d) Find the MMSE forecast of X_{n+l} for an AR(1) process, with zero mean, using MA representation approach

Question No. 3. Given an AR(2) process, $X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + Z_t$,   (10)

Find MLE of σ_z^2, also find approximate MLEs of φ_1 and φ_2.

Question No. 4. Explain the meaning and importance of updating forecasts. Derive the   (10)

rule to update the forecast previously made with origin at "n" using the information X_{n+1}.

Question No. 5. For the model $(1 - B)(1 - 0.2B)X_t = (1 - 0.5B)Z_t$ (10)

Find forecast for one and two steps ahead and show that

$$X_n(l) = 1.2X_n(l - 1) - 0.2X_n(l - 2), for\ l \geq 3$$

If $Z_n = 1$, $X_n = 4$, $X_{n-1} = 3$ & $\sigma_Z^2 = 0.10$.

Also show that

$X_n(2) = 3.64$ and $Var[e_n(1)] = 0.10$.

UNIVERSITY OF THE PUNJAB

Roll No. ....

Eighth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Multivariate Analysis
Course Code: STAT-414

TIME ALLOWED: 30 mins.
MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

OBJECTIVE

Q1. Read the following items carefully and encircle the correct option listed below at each item.

(One mark for each)

i) A multivariate generalization of student's t-distribution is
a) Hotelling T^2 b) MANOVA   c) PCA   d) None of Theses

ii) The linear combinations of independent variables developed by discriminant analysis that will best discriminate between the categories of the dependent variable are:
a) discriminant functions   c) discriminant scores
b) characteristic profiles   d) classification matrix

iii) Discriminant analysis can be used to answer questions such as:
a) How much of the variation in sales can be explained by advertising expenditures, prices, and level of distribution?
b) In terms of demographic characteristics, how do customers who exhibit store loyalty differ from those who do not?
c) What are the distinguishing characteristics of consumers who respond to direct mail solicitations?
d) Both b and c are correct.

iv) What does it mean if the discriminant function is estimated and the square of the canonical correlation is .64?
a) 64% of the variance in the dependent variable is explained by the model.
b) The null hypothesis is not rejected. Therefore, there is no significant discrimination between groups.
c) 64% of the explained variance is accounted for.
d) b and c are correct

v) A technique for the study of interrelationships among variables, usually for the purposes of data reduction and the discovery of underlying constructs or latent dimensions is known as:
a) Multiple Regression   c) Discriminant Analysis
b) Factor Analysis   d) Canonical Analysis

vi) To determine which variables relate to which factors, a researcher would use:

a) Factor Loadings   b) Communalities   c) Eigen Values   d) None of these

vii) Which of the following can be used to determine how many factors to take from a factor analysis
a) Eigen Values   b) Scree Plot   c) % of variance   d) All of these

viii) The factors extracted in a factor analysis are made more clear and interpretable by:
a) converting raw scores to z scores before analysis.
b) eliminating variables that have low correlations with other variables.
c) applying a square root transformation to the raw data before analysis.
d) statistically rotating factors.

ix) Correlating a set of four predictor variables with five criterion variables uses which analysis?
a) Multiple Regression   c) Discriminant Analysis
b) Factor Analysis   d) Canonical Analysis

x) Generally speaking, multivariate analysis requires:
a) Fairly Large Samples   b) Small Samples
c) less concern over meeting assumptions than do univariate tests.
d) sampling from a population that is not normally distributed.

UNIVERSITY OF THE PUNJAB

Eighth Semester - 2017

Examination: B.S. 4 Years Programme

Roll No. ....

PAPER: Multivariate Analysis

Course Code: STAT-414

TIME ALLOWED: 2 hrs. & 30 mins.

MAX. MARKS: 50

Attempt this Paper on Separate Answer Sheet provided.

SHORT QUESTIONS

Q2. Differentiate between the following terms: *(4,4,4,4,4)*

i) Principal Component Analysis and Factor Analysis

ii) Communalities and Factor Loadings

iii) Common Factors and Specific Factors

iv) Discriminant Analysis and Canonical Correlations

v) Mahalnobis Distances and Cook's Distances

SUBJECTIVE

Q3. Derive the distribution of Hotelling T^2-statistic for testing $H_0 : \mu = \mu_0$. *(08)*

Q4. Consider the sample correlation matrix based on 150 observations: *(07)*

$$R = \begin{bmatrix} 1.00 & 0.75 & 0.65 & 0.66 \\ & 1.00 & 0.69 & 0.73 \\ & & 1.00 & 0.67 \\ & & & 1.00 \end{bmatrix} \text{Test the hypothesis that: } H_0 : \rho = \begin{bmatrix} 1 & \rho & \rho & \rho \\ & 1 & \rho & \rho \\ & & 1 & \rho \\ & & & 1 \end{bmatrix}$$

Q5. The following mean vector and covariance matrices are based on $n_1 = n_2 = 100$ observations:

$$\bar{X}_1 = \begin{bmatrix} 6.213 \\ 3.133 \end{bmatrix}, \bar{X}_2 = \begin{bmatrix} 7.412 \\ 5.321 \end{bmatrix}, S_1 = \begin{bmatrix} 1.813 & 0.321 \\ & 0.937 \end{bmatrix}, S_2 = \begin{bmatrix} 2.193 & 1.654 \\ & 3.789 \end{bmatrix}$$

Find the Fisher's Linear Discriminant function and Discriminant rule. Also allocate the new observations $\bar{X}' = \begin{bmatrix} 7.2 & 3.1 \end{bmatrix}$ to any of these populations. *(08)*

Q6. A class of 140 tenth grade children received four tests, two open books $X_1' = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$ and two closed books $X_2' = \begin{bmatrix} X_3 & X_4 \end{bmatrix}$. The mean vector and joint covariance matrix is given below:

$$\mu = \begin{bmatrix} -3 \\ 2 \\ 0 \\ 1 \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} 8 & 2 & 3 & 1 \\ & 5 & -1 & 3 \\ & & 6 & -2 \\ & & & 7 \end{bmatrix}$$

Find the canonical correlations between X_1 & X_2 and the first pair of canonical variables.

(07)

UNIVERSITY OF THE PUNJAB

Roll No. ....

Eighth Semester - 2017
Examination: B.S. 4 Years Programme

PAPER: Total Quality Management   TIME ALLOWED: 30 mins.
Course Code: STAT-415   MAX. MARKS: 10

Attempt this Paper on this Question Sheet only.

Q. No. 1: Encircle the correct option.

1. The appropriate control chart for number of defective items is called --------------
   A. p-chart   B. np-chart
   C. c-chart   D. u-chart

2. The philosophy of TQM is ------------------------
   A. Cost saving   B. Accept the lot
   C. Continuous Improvement   D. Reject the lot

3. Under Six Sigma level, the processes could produce -------- defects per million opportunities.
   A. 3.2   B. 3.4
   C. 3.6   D. 3.8

4. In SWOT analysis, the -------------------- are characteristics that place the firm at an advantage relative to others.
   A. Strengths   B. Weaknesses
   C. Opportunities   D. Threats

5. _____is intended to assess a company's performance in terms of environmental performance.
   A. ISO 14000   B. ISO 9000
   C. Six sigma   D. All of the above

6. -------------------- costs are associated with the design, implementation and maintenance of the TQM system.
   A. Prevention   B. Appraisal
   C. Internal Failure   D. External Failure

7. Total quality management has evolved from the quality assurance methods that were first developed around the time of the ----------------------
   A. First World War   B. Second World War
   C. Third World War   D. None of the above

8. The benefits of TQM includes ------------------------
   A. Improved quality   B. Employee participation
   C. Customer satisfaction   D. All of the above

9. ---------------- is a reference or measurement standard used for comparison.
   A. Control chart   B. Six-sigma
   C. Benchmarking   D. Double Sampling Plan

10. Defective product or material that cannot be repaired, used or sold is called -------------.
    A. Waste   B. Rework
    C. Scrap   D. None of the above

UNIVERSITY OF THE PUNJAB

Eighth Semester - 2017

Examination: B.S. 4 Years Programme

PAPER: Total Quality Management	**TIME ALLOWED: 2 hrs. & 30 mins.**
Course Code: STAT-415	**MAX. MARKS: 50**

Attempt this Paper on Separate Answer Sheet provided.

Attempt all questions.

Q. No. 2. Write a short note on the following: (5 each)
  i. Total Quality Management
  ii. Competitive benchmarking
  iii. International Standard Organization (ISO)
  iv. Failure costs in TQM

Q. No. 3. Describe Deming's Fourteen points for quality improvement in detail.   **(07)**

Q. No. 4. Take a sampling plan with $n_1 = 50$, $c_1 = 1$, $n_1 + n_2 = 150$, $c_2 = 4$ **(08)**
If the incoming lots have fraction nonconforming $p = 0.05$ then what is the probability of final acceptance? Calculate the probability of rejection on the first sample?
Where n_1 & n_2 are the sample sizes and c_1 and c_2 are the acceptance numbers.

Q. No. 5. An item is made in lots of 200 each. The lots are given 100% inspection. The   **(07)**
record sheet for the first 25 lots inspected showed that a total of 75 items were defective.
  a. Determine the trial control limits for an np-chart showing number of defectives in each lot.
  b. Assume that all points fall within the control limits. What is your estimate of the process average fraction defective p' ?
If this p' remains unchanged. What is the probability that (i) the 26th lot will contain exactly 7 defectives? (ii) it will contain 7 or more defectives?

Q. No. 6. What is benchmarking? Also describe the advantages and disadvantages of   **(08)**
benchmarking?