		Semester – VI
		L T P C
16MMU603A	**PROBABILITY AND STATISTICS**	**4 2 0 4**

Scope: On successful completion of this course the learner gains a clear knowledge about the fundamentals of probability and Statistics, various measures and their applications.

Objectives: To enable the students to understand the meaning, definition and functions of probability and statistics, finding various measures such as mean, median, mode etc and various distributions functions and their characteristics, and its applications.

UNIT I

Meaning and definition of statistics, Measures of central tendency: Arithmetic Mean, Median, Mode. Measures of dispersion – Range, Coefficient of range, Quartile deviation, Coefficient of Quartile deviation, Standard deviation and Coefficient of variation.

UNIT II

Probability Concepts – trial, event, Sample space, mutually exclusive event, exclusive and exhaustive events, dependent and independent events, simple and compound events, Mathematical properties, permutation and combination, probability axioms, addition and multiplication theorem, real random variables (discrete and continuous), cumulative distribution function, probability density functions, mathematical expectation, moments, moment generating function, characteristic function.

UNIT III

Discrete distributions: uniform distribution, binomial distribution, Poisson distribution, and its properties, joint probability density functions, marginal and conditional distributions, expectation of function of two random variables, conditional expectations, independent random variables.

UNIT IV

Continuous distributions: uniform distribution, normal distribution, standard normal distribution, exponential distribution. Joint cumulative distribution function and its properties, joint probability density functions (No derivations) and simple problems. Bivariate distribution, correlation coefficient, joint moment generating function (jmgf) and calculation of covariance (from jmgf), linear regression for two variables.

UNIT V

Chebyshev's inequality, statement and interpretation of (weak) law of large numbers and strong law of large numbers, Central Limit theorem for independent and identically distributed random variables with finite variance, Markov Chains, Chapman-Kolmogorov equations, classification of states.

SUGGESTED READINGS

TEXT BOOK

1. Gupta S.P., (2001). Statistical Methods, Sultan Chand & Sons, New Delhi.

REFERENCES

1. Robert V. Hogg, Joseph W. McKean and Allen T. Craig., (2007). Introduction to Mathematical Statistics, Pearson Education, Asia.

2. Irwin Miller and Marylees Miller, John E. Freund, (2006). Mathematical Statistics with Application, Seventh Edition, Pearson Education, Asia.

3. Sheldon Ross., (2007). Introduction to Probability Model, Ninth Edition, Academic Press, Indian Reprint.

4. Pillai R.S.N., and Bagavathi V., (2002). Statistics , S. Chand & Company Ltd, New Delhi.

5. Srivastava T.N., and Shailaja Rego., (2012). 2e, Statistics for Management, Mc Graw Hill Education, New Delhi.

6. Dr.P.N.Arora, (1997). A foundation course statistics, S.Chand & Company Ltd, New Delhi.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)

(Established under Section 3 of UGC Act, 1956)

Pollachi Main Road, Coimbatore – 641 021, Tamil Nadu

Department of Mathematics

Subject : Mathematical Statistics	**Semester VI**	**L T P C**
Subject Code : 16MMU603A	**Class : III B.Sc Mathematics**	**4 2 0 4**

UNIT I

Meaning and definition of statistics, Measures of central tendency: Arithmetic Mean, Median, Mode. Measures of dispersion – Range, Coefficient of range, Quartile deviation, Coefficient of Quartile deviation, Standard deviation and Coefficient of variation.

SUGGESTED READINGS

TEXT BOOK

1. Gupta S.P., (2001). Statistical Methods, Sultan Chand & Sons, New Delhi.

REFERENCES

1. Robert V. Hogg, Joseph W. McKean and Allen T. Craig., (2007). Introduction to Mathematical Statistics, Pearson Education, Asia.
2. Irwin Miller and Marylees Miller, John E. Freund, (2006). Mathematical Statistics with Application, Seventh Edition, Pearson Education, Asia.
3. Sheldon Ross., (2007). Introduction to Probability Model, Ninth Edition, Academic Press, Indian Reprint.
4. Pillai R.S.N., and Bagavathi V., (2002). Statistics , S. Chand & Company Ltd, New Delhi.
5. Srivastava T.N., and Shailaja Rego., (2012). 2e, Statistics for Management, Mc Graw Hill Education, New Delhi.
6. Dr.P.N.Arora, (1997). A foundation course statistics, S.Chand & Company Ltd, New Delhi.

Unit – I

Introduction

Statistical tools are found useful in progressively increasing of disciplines. In ancient times the statistics or the data regarding the human force and wealth available in their land had been collected by the rulers. Now-a-days the fundamental concepts of statistics are considered by many to be essential part of their knowledge.

Origin and Growth

The origin of the word 'statistics' has been traced to the Latin word 'status', the Italian word 'statista' , the French word 'statistique' and the German word 'statistik'. All these words mean political state.

Meaning

The word 'statistics' is used in two different meanings. As a plural word it means data or numerical statements. As a singular word it means the science of statistics and statistical methods. The word 'statistics' is also used currently as singular to mean data.

Definitions

Statistics is "the science of collection, organization, presentation, analysis and interpretation of numerical data". – Dr S.P.Gupta.

Statistics are numerical statement of facts in any department of enquiry, placed in relation to each other". – Dr.A.L.Bowley.

Functions

The following are the important functions of statistics.

- σ Collection
- σ Numerical Presentation
- σ Diagrammatic Presentation
- σ Condensation

- σ Comparison
- σ Forecasting
- σ Policy Making
- σ Effect Measuring
- σ Estimation
- σ Tests of significance.

Characteristics

* Statistics is a Quantitative Science.

* It never considers a single item.

* The values should be different.

* Inductive logic is applied.

* Statistical results are true on the average.

* Statistics is liable to be misused.

Samples vs. Populations

Population: A complete set of observations or measurements about which conclusions are to be drawn.

Sample: A subset or part of a population.

Not necessarily random

Statistics vs. Parameters

Parameter: A summary characteristic of a population.

Summary of Central tendency, variability, shape, correlation

E.g., Population mean, Population Standard Deviation, Population Median, Proportion of population of registered voters voting for Bush, Population correlation between Systolic & Diastolic BP

Statistic: A summary characteristic of a sample. Any of the above computed from a sample taken from the population.

E.g., Sample mean, Sample Standard Deviation, median, correlation coefficient

MEASURES OF CENTRAL TENDENCY

Introduction

In this chapter we are going to deal with Measures of central tendency and about the measures of dispersion. The measures of central tendency concentrate about the values in the central part of the distribution. Plainly speaking an average of a statistical series is the value of the variable which is the representative of the entire distribution. If we know the average alone we cannot form a complete idea about the distribution so for the completeness of the idea we use Measures of dispersion.

Measures of Central Tendency

According to Professor Bowley the measures of central tendency are “statistical constants which enable us to comprehend in a single effort the significance of the whole “

The following three are the basic measures of central tendency in this chapter we deal with

- Arithmetic Mean or simply Mean
- Median
- Mode

Arithmetic Mean or Mean

Arithmetic Mean or simply Mean is the total values of the item divided by their number of the items. It is usually denoted by $\overline{X}$

Individual series

$\overline{X} = \Sigma X / N$

Example 1

The expenditure of ten families are given below .Calculate arithmetic mean.

30, 70, 10, 75, 500, 8, 42, 250, 40, 36

Solution

Here N=10

$\Sigma X = 30 + 70 + 10 + 75 + 500 + 8 + 42 + 250 + 40 + 36 = 1061$

$\overline{X} = 1061 / 10 = 106.1$

Discrete series

$\overline{X} = \Sigma f X / \Sigma f$

Example 2

Calculate the mean number of person per house.

No. of person : 2 3 4 5 6

No. of house : 10 25 30 25 10

Solution

X	f	f X
2	10	20
3	25	75

4	30	120
5	25	125
6	10	60
	$\Sigma f = 100$	$\Sigma fX = 400$

$\bar{X} = 400 / 100 = 4$.

Continuous series

$\bar{X} = \Sigma f m / \Sigma f$ where m represents the mid value

Mid-value = (upper boundary + lower boundary) / 2.

Example 3

Calculate the mean for the following.

Marks :	20-30	30-40	40-50	50-60	60-70	70-80
No. of student :	5	8	12	15	6	4

Solution

C.I	**f**	**m**	**f m**
20-30	5	25	125
30-40	8	35	280
40-50	12	45	540
50-60	15	55	825
60-70	6	65	390
70-80	4	75	300
	$\Sigma f = 50$		$\Sigma fm = 2460$

X = 2460 / 50 = 49.2.

Median

The median is the value for the middle most items when all the items are in the order of magnitude. It is denoted by M or Me.

Individual series

For odd number of items Median Position = (N+1) / 2

For even number of item

Position of the Median = [(N / 2)+((N/2)+1)] / 2

Example 1

Calculate median for the following.

22 10 6 7 12 8 5

Solution

Here N =7

Arrange in ascending order or descending order.

5 6 7 8 10 12 22

(N+1) / 2= (7+1) /2 = 4 th item = 8

Discrete series

Position of the median = (N+1) / 2 th item.

Example 2

Find the median for the following.

X : 10 15 17 18 21

F: 4 16 12 5 3

Solution

X f c.f

10	4	4
15	16	20
17	12	32
18	5	37
21	3	40

N= 40

(N+1) /2 = (40+1) / 2 = 20.5^{th} item

= (20^{th} item $+21^{st}$ item) /2 =(15+17) /2

= 16.

Continuous series

$$M = L + \frac{((N/2) - c.f) \times i}{f}$$

Where L - lower boundary, f - frequency, I - size of class interval and c.f - cumulative frequency.

Example 3

Calculate the median height (Ht) for the No .of Students (NoS) given below.

Ht: 145-150 150-155 155-160 160-165 165-170 170-175

NoS: 2 5 10 8 4 1

Solution

Height	No. of student	c.f
145-150	2	2
150-155	5	7
155-160	10	17
160-165	8	25

165-170	4	29
170-175	1	30

$\Sigma f = 30$

Position of the median = N/2 th item = 30 / 2 =15.

$$M = L + \frac{\left((N/2) - c.f\right) \times i}{f}$$

$$= 155 + \frac{(15-7)\times 5}{10} = 155 + (40/10) = 159.$$

Mode

Mode is the value which has the greatest frequency density. Mode is usually denoted by Z.

Individual series

In a set of observations the value which occur more number of time is known as Mode. In other way the most frequented value in a set of value is Mode.

Example 1

Determine the mode for the set of Individual observations given as follows 32, 35, 42, 32, 42, 32.

Solution

Mode = 32 Uni-model

Discrete series

Determine the Mode

Size of dress	No. of set
18	55
20	120
22	108

24 45

Here the mode represents highest frequency ie. I 120.

So, Mode = 20

Continuous series

$Z = L + [\, i(f_1 - f_0) / (2f_1 - f_0 - f_2)]$

Where L- lower boundary , f_1-frequency of the modal class, f_0 – frequency of the preceding modal class, f_2- frequency of the succeeding modal class, i-size of class interval , c.f- cumulative frequency.

Example

Determine the mode

Marks :	0-10	10-20	20-30	30-40	40-50
No.of student :	5	20	35	20	12

Solution

Marks	No. of student
0-10	5
10-20	20
20-30	35
30-40	20
40-50	12

$Z = L + [\, i(f_1 - f_0) / (2f_1 - f_0 - f_2)]$

$= 20+[10(35-20)/(2(35)-20-20)] = 20+5 = 25.$

Empirical relation

- Mode = 3 median -2 mean.

MEASURES OF DISPERSION

Measure of dispersion deals mainly with the following three measures

- Range
- Standard deviation
- Coefficient of variation

Range

Range is the difference between the greatest and the smallest value.

- Range = L – S , where L-largest value & S-Smallest value
- Coefficient of range = (L-S) /(L+S)

Individual series

Example

Find the value of range and its coefficient of range for the following data.

8 ,10, 5, 9,12,11

Solution

Range = L – S

= 12- 5 =7

Coefficient of range = (L-S) /(L+S)

= (12-5) / (12+5)

= 7 /17

= 0.4118

Continuous series

Range = L – S, where L-Mid-value of largest boundary & S-Mid-value of smallest boundary

Calculate the range for the following continuous frequency distribution

Marks :	20-30	30-40	40-50	50-60	60-70	70-80
No.of student :	5	8	12	15	6	4

Solution

C.I	**f**	**m**
20-30	5	25
30-40	8	35
40-50	12	45
50-60	15	55
60-70	6	65
70-80	4	75

Here $L=75$ & $S=25$

Range = $L - S = 75-25 = 50$

Quartile Deviation

Quartile Deviation is half of the difference between the first and the third quartiles. Hence it is called Semi Inter Quartile Range.

Coefficient of Quartile Deviation

A relative measure of dispersion based on the quartile deviation is called the coefficient of quartile deviation. It is defined as

It is pure number free of any units of measurement. It can be used for comparing the dispersion in two or more than two sets of data.

Example

The wheat production (in Kg) of 20 acres is given as: 1120, 1240, 1320, 1040, 1080, 1200, 1440, 1360, 1680, 1730, 1785, 1342, 1960, 1880, 1755, 1720, 1600, 1470, 1750, and 1885. Find the quartile deviation and coefficient of quartile deviation.

Solution

After arranging the observations in ascending order, we get

1040, 1080, 1120, 1200, 1240, 1320, 1342, 1360, 1440, 1470, 1600, 1680, 1720, 1730, 1750, 1755, 1785, 1880, 1885, 1960.

$$Q_1 = \text{Value of } \left(\frac{n+1}{4}\right) th \text{ item}$$

$$= \text{Value of } \left(\frac{20+1}{4}\right) th \text{ item}$$

$$= \text{Value of } (5.25) th \text{ item}$$

$$= 5th \text{ item} + 0.25(6th \text{ item} - 5th \text{ item}) = 1240 + 0.25(1320 - 1240)$$

$$Q_1 = 1240 + 20 = 1260$$

$$Q_3 = \text{Value of } \frac{3(n+1)}{4} th \text{ item}$$

$$= \text{Value of } \frac{3(20+1)}{4} th \text{ item}$$

$$= \text{Value of } (15.75) th \text{ item}$$

$$= 15th \text{ item} + 0.75(16th \text{ item} - 15th \text{ item}) = 1750 + 0.75(1755 - 1750)$$

$$Q_3 = 1750 + 3.75 = 1753.75$$

Quartile Deviation (Q.D)

$$= \frac{Q_3 - Q_1}{2} = \frac{1753.75 - 1260}{2} = \frac{492.75}{2} = 246.875$$

Coefficient of Quartile Deviation

$$= \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{1753.75 - 1260}{1753.75 + 1260} = 0.164$$

Standard deviation

The standard deviation is the root mean square deviation of the values from the arithmetic mean .It is a positive square root of variants. It is also called root mean square deviation. This is usually denoted by σ .

Individual series

$$\sigma = \sqrt{(\Sigma x^2 / N) - (\Sigma x / N)^2}$$

Example 1

Calculate standard deviation for the following data.

40,41,45,49,50,51,55,59,60,60.

Solution

X	X^2
40	1600
41	1681
45	2025
49	2401
50	2500
51	2601
55	3025
59	3481
60	3600
60	3600
510	$\Sigma x^2 = 26504$

$$\sigma = \sqrt{(\Sigma x^2 / N) - (\Sigma x / N)^2}$$

$= \sqrt{(26514/10) - (510/10)^2}$

$= 7.09$

Discrete series

$\sigma = \sqrt{(\Sigma fx^2 / \Sigma f) - (\Sigma fx / \Sigma f)^2}$

Example 2

Calculate standard deviation for the following data.

X : 0 1 2 3 4 5

F : 1 2 4 3 0 2

Solution

X	f	fx	x^2	fx^2
0	1	0	0	0
1	2	2	1	2
2	4	8	4	16
3	3	9	9	27
4	0	0	16	0
5	2	10	25	50
	$\Sigma f = 12$	$\Sigma fx = 29$		$\Sigma fx^2 = 95$

$\sigma = \sqrt{(\Sigma fx^2 / \Sigma f) - (\Sigma fx / \Sigma f)^2}$

$= \sqrt{(95/12) - (29/12)^2} = 1.44$

Continuous series

$$\sigma = \sqrt{(\Sigma fm^2 / \Sigma f) - (\Sigma fm / \Sigma f)^2}$$

Example 3

C.I	0-10	10-20	20-30	30-40	40-50
F	2	5	9	3	1

Solution

C.I	**f**	**m**	**fm**	**m^2**	**fm^2**
0-10	2	5	10	25	50
10-20	5	15	75	225	1125
20-30	9	25	225	625	5625
30-40	3	35	105	1225	3675
40-50	1	45	5	2025	2025
	20		460		12500

$$\sigma = \sqrt{(\Sigma fm^2 / \Sigma f) - (\Sigma fm / \Sigma f)^2}$$

$$= \sqrt{(12500/20) - (460/20)^2}$$

$= 9.79$

Coefficient of variation

Coefficient of variation = [standard deviation / arithmetic mean] x100

Example 1

Calculate the coefficient of variation.

Mean= 51, standard deviation = 7.09

Solution

Coefficient of variation = [standard deviation / arithmetic mean] x100

= (7.09 /51) x100

= 13.9

Questions

1) Calculate the Mean for the following.

X	20	30	35	15	10
f	2	3	4	3	2

2) Define Median and give Example.

3) Calculate the Range and its Coefficient for the following data.

X	:	12	14	16	18	20
f	:	1	3	5	3	1

4) What do you mean by Bimodal?

5) Calculate the Median for the following data.

80 100 50 90 120 110

6) Write the relation between Standard Deviation and Variance.

7) Calculate the Average number of students per class for the following data.

26 46 33 25 36 27 34 29

8) Find Median and Mode for the following data.

13 16 17 15 18 14 19 15 12

9) Define the term Range.

10) Find the Arithmetic Mean for the following data.

70 60 75 50 42 95 46

11) Calculate the Range and its Coefficient for the following data.

17 10 56 19 12 11 18 14

12) Define the term Quartile Deviation.

13) Find the median for 57, 58, 61, 42, 38, 65, 72, and 66

14) Write the empirical relation for Mode.

Exercise

1. Calculate the Mode for the following Continuous Frequency Distribution.

Salary (in Rs. 1000s)	0 -19	20-39	40 - 59	60-79	80-99
No. of Employees	5	20	35	20	12

2. Find the Mean and the Standard Deviation for the given below data set.

10	14	20	12	21	16	19	17	14	25

3. Calculate the Standard Deviation and Coefficient of Variance (CV) for the following data.

X	0 – 10	10 - 20	20 - 30	30 – 40	40 - 50
f	2	5	9	3	1

4. Calculate the Median for the following Continuous Frequency Distribution.

Wages (in Rs.)	0 - 19	20 - 39	40 - 59	60 – 79	80 - 99
No. of Workers	5	20	35	20	12

5. Calculate the Coefficient of Variation for the following data.

X	6	9	12	15	18
f	7	12	13	10	8

6. Calculate the Median for the following.

Hourly Wages (in Rs.)	40 - 50	50 – 60	60 - 70	70 - 80	80 - 90	90 - 100
Number of Employees	10	20	15	30	15	10

7. The following data give the details about salaries (in thousands of rupees) of seven employees randomly selected from a Pharmaceutical Company.

Serial No.	1	2	3	4	5	6	7
Salary per Annum (‘000)	89	57	104	73	26	121	81

Calculate the Standard Deviation and Coefficient of variance of the given data.

8. Calculate the Arithmetic Mean for the following data.

Height (cms) : 160 161 162 163 164 165 166

No. of Persons : 27 36 43 78 65 48 28

9. Calculate the Coefficient of Variance for the following data. 77 73 75 70 72 76 75 72 74 7

Question	Opt 1	Opt 2	Opt 3	Opt 4	Answer
The word statistics is used as--------------------	singular word	a plural word	both singular and plural words	neither singular nor plural word	both singular and plural words
Classification is a process of arranging data in------------------	grouping of related facts in different classes	different rows	different columns and rows	different columns grouping of related facts in different classes	grouping of related facts in different classes
To represent two or more sets of interrelated data , we use -----------	bar diagram	pie diagram	histogram	multiple bar diagram	multiple bar diagram
Histogram is a graph of ------------------------	Time series	frequency distribution	cumulative frequency distribution	normal distribution	frequency distribution
Univariate data consists of -----------------------	one variable	two variables	three variable	four	one variable
Data are generally obtained from--------------------	Primary sources	Secondary sources	Both primary and secondary sources	neither primary nor secondary sources	Both primary and secondary sources
In geographical classification data are classified on the basis of------------------	area	attributes	time	location	area
In qualitative classification data are classified on the basis of------------------	area	attributes	time	location	attributes
In quantitative classification data are classified on the basis of-----------------	area	attributes	time	magnitude	magnitude
Number of source of data is--------------------	2	3	4	1	2
Squares and rectangles are---------------------------	Two dimensional diagram	One dimensional diagram	Three dimensional diagram	Multi dimensional diagram	Two dimensional diagram
Data originally collected for an investigation is known as------------------------	Tabulation	Primary data	Secondary data	Published data	Primary data
The heading of a row in a statistical table is known as -----------------------	stub	caption	title	heading	stub

Statistics can ----------	prove anything	disprove anything	neither prove nor disprove anything but it is just a tool	none of these	neither prove nor disprove anything but it is just a tool
Statistics is also a science of-----------------	estimates	both a and b	probabilities	neither a nor b	both a and b
Statistics is---------	quantitative science	a qualitative science	both quantitative and qualitative science	neither quantitative nor qualitative	both quantitative and qualitative science
Statistics considers-------------	a single item	a set of item	either a single item or a set of item	neither a single item or a set of item	a set of item
Statistics can be considered as----------------	an art	a science	both an art and science	neither an art nor a science	both an art and science
The other name of cumulative frequency curve is ----	Ogive	Bars	Histogram	Pie diagram	Ogive
Number of methods of collection of primary data is--------------------	2	3	4	5	5
Number of questions in a questionnaire should be------------------------	5	10	maximum	minimum	minimum
Sources of secondary data are-----------	Published sources	Unpublished sources	Either Published sources or Unpublished sources	primary source	Either Published sources or Unpublished sources
Compared with primary data , secondary data are------------	more reliable	less reliable	equally reliable	uniformly reliable	less reliable
----------- are column headings	stub	heading	bar	captions	captions
Mid value= ------	lower boundary/2	upper boundary/2	lower boundary+ upper boundary)/2	lower boundary+ upper boundary	lower boundary+ upper boundary)/2
The origin of the word statistics has been traced to the Latin word -------	statista	status	statistik	statistique	status
Graphs of frequency distribution are-----------------	histogram	pie diagram	bar chart	circle	histogram

cubes are-----------------------	Two dimensional diagram	One dimensional diagram	Three dimensional diagram	Multi dimensional diagram	Three dimensional diagram
------------------ is the difference between the value of the smallest item and the valueof the largest item.	class interval	frequency	number of items	range	range
------------- is one which is used by the individual or agency which collect it.	primary data	secondary data	both		primary data
Exclusive class intervals suit -------------------------	discrete variables	continuous variables	both	neither	continuous variables
A table is a systematic arrangement of statistical data in-----------------	columns	rows	both columns and rows	stubs	both columns and rows
The collected data in any statistical investigation are known as -------------	raw data	arranged data	classified data	tabulated data	raw data
The emitting form of a frequency polygon is called -----------------------	histogram	ogive	bar diagram	frequency curve	frequency curve
In chronological classification data are classified on the basis of---------------	time	attributes	class intervals	location	time
Bar diagrams are ------------- dimensional diagrams	two	three	one	multi	one
Diagrams and graphs are tools of -------------	collection of data	presentation	analysis	summarization	presentation
In a two dimensional diagram -------------------	only height is considered	only width is considered	height,width and thickness are considered	Both height and width are considered	only height is considered
Which one of the following is a measure of central	Median	range	variation	correlation	Median
The total of the values of the items divided by their number of items is known as	Median	Arithmetic mean	mode	range	Arithmetic mean
In the short-cut method of arithmetic mean, the deviation is taken as	x – A	A – x	(x – A) / c	(A – x) / c	x – A
The sum of the deviations of the values from their arithmetic mean is	– 1	one	two	zero	zero
The formula for the weighted arithmetic mean is	∑wx / ∑w	∑w / ∑wx	∑x / n	∑x / ∑f	∑wx / ∑w
Find the Mean of the following values. 5, 15, 20, 10,	5	18	41	20	18
Which of the followings represents median?	First quartile	Third quartile	Second quartile	Q.D	Second quartile

Which of the measure of central tendency is not affected by extreme values?	Mode	Median	sixth deciles	Mean	Median
Sum of square of the deviations about mean is	Maximum	one	zero	Minimum	Minimum
Median is the value of -------------- item when all the items are in order of magnitude.	First	second	Middle most	last	Middle most
Find the Median of the following data 160, 180, 175, 179, 164, 178, 171, 164, 176.	160	175	176	180	175
The position of the median for an individual series is	(N + 1) / 2	(N + 2) / 2	N/2	N/4	(N + 1) / 2
Mode is the value, which has	Average frequency density	less frequency density	greatest frequency density	graetest frequency	greatest frequency density
A frequency distribution having two modes is said to	unimodal	bimodal	trimodal	modal	bimodal
Mode has --------------- stable than mean.	less	more	same	most	less
Which of the following is not a measure of dispersion?	Range	quartile deviation	standard deviation	median	median
Range of the given values is given by	L- S	L+S	S+L	LS	L- S
Which one of the following is relative measure of dispersion?	Range	Q.D	S.D	coefficient of variation	coefficient of variation
Coefficient of variation is defined as	(AM * 100)/S.D	(S.D* 100)/A.M	S.D/A.M	(1/S.D)*100	(S.D* 100)/A.M
If the values of median and mean are 72 and 78 respectively, then find the mode.	16	60	70	76	60
Find Mean for the following 3, 4, 5.	4	2.25	3	2.28	4
The coefficient of range	L-S /L+S	L+S /L-S	L-S	L+S	L-S /L+S
Second quartile is also called as	Mode	mean	median	G.M	median
If A.M = 8, N=12, then find ∑X.	76	80	86	96	96
If the value of mode and mean is 60 and 66 then, find the value of median.	64	46	54	44	64
The formula for median for continuous series is	M = (N+1) / 2	M = L + [(N/2 + cf) / f] * i	M =L - (N/2+cf)/f* i	M = L + [(N/2 - cf) / f] * i	M = L + [(N/2 - cf) / f] * i

Median is	Average point	Midpoint	Most likely point	Most remote point	Midpoint
Mode is the value which	Is a mid point	Occur the most	Average of all	Most remote Likely	Occur the most
. ……………. Is known as positional average	Median	Mean	Mode	Range	Median
The median of marks 55, 60, 50, 40, 57, 45, 58, 65, 57, 48 of 10 students is	55	57	52.5	56	56
The middle most value of a frequency distribution table is known as	Mean	Median	Mode	Range.	Median
The middle most value of a frequency distribution table is known as	Mean	Median	Mode	Range	Median
Measures of central tendency is also known as	Dispersion	averages	correlation	tendency	correlation
From the given data 35,40,43,32,27 the coefficient	23	0.23	13	0.13	13
If S.D = 6, then find variance.	6	36	42	12	36
Which one of the following shows the relation between variance and standard deviation?	var = square root of S.D	S.D = square root of variance	variance = S.D	variance / S.D = 1	S.D = square root of variance
If variance is 64, then find S.D.	8	13	14	11	8
Which of the following measures of averages divide the observation into two parts	Mean	Median	Mode	Range	Median
Which of the following measures of averages divide the observation into four equal parts	Mean	Median	Mode	Quartile	Quartile
Arithmetic mean of the series 1, 3, 5, 7, 9 is	5	6	5.5	6.5	5
Arithmetic mean of the series 3, 4, 5, 6, 7 is ………	5.5	6	5	6.5	5
The Arithmetic mean for the series 3, 5, 5, 2, 6, 2, 9, 5, 8, 6, is…………….	5	6	5.5	6.5	5
The median value for the series 3, 5, 5, 2, 6, 2, 9, 5,	6	5	5.5	6.5	5
The mode for the series 3, 5, 6, 2, 6, 2, 9, 5, 8, 6 is	5	6	5.5	6.5	6
The Arithmetic mean for the series 51.6, 50.3, 48.9, 48.7, 48.5 is……….	49.8	50	48.9	49.6	49.8
The Median for the series 51.6, 50.3, 48.9, 48.7, 49.5, is…………	49.8	50	48.9	49.6	49.6

The Mode for the series 51.6, 50.3, 48.9, 48.7, 49.5 is………….	49.8	50	48.9	49.6	48.9
If standard deviation is 5, then the variance is	5	625	25	2.23068	25
Standard deviation is also called as	Root mean square deviation	mean square deviation	Root deviation	Root median square deviation	Root mean square deviation

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)

(Established under Section 3 of UGC Act, 1956)

Pollachi Main Road, Coimbatore – 641 021, Tamil Nadu

Department of Mathematics

Subject : Mathematical Statistics **Semester VI** **L T P C**

Subject Code : 16MMU603A **Class : III B.Sc Mathematics** **4 2 0 4**

UNIT II

Probability Concepts – trial, event, Sample space, mutually exclusive event, exclusive and exhaustive events, dependent and independent events, simple and compound events, Mathematical properties, permutation and combination, probability axioms, addition and multiplication theorem, real random variables (discrete and continuous), cumulative distribution function, probability density functions, mathematical expectation, moments, moment generating function, characteristic function.

SUGGESTED READINGS

TEXT BOOK

1. Gupta S.P., (2001). Statistical Methods, Sultan Chand & Sons, New Delhi.

REFERENCES

1. Robert V. Hogg, Joseph W. McKean and Allen T. Craig., (2007). Introduction to Mathematical Statistics, Pearson Education, Asia.
2. Irwin Miller and Marylees Miller, John E. Freund, (2006). Mathematical Statistics with Application, Seventh Edition, Pearson Education, Asia.
3. Sheldon Ross., (2007). Introduction to Probability Model, Ninth Edition, Academic Press, Indian Reprint.
4. Pillai R.S.N., and Bagavathi V., (2002). Statistics , S. Chand & Company Ltd, New Delhi.
5. Srivastava T.N., and Shailaja Rego., (2012). 2e, Statistics for Management, Mc Graw Hill Education, New Delhi.
6. Dr.P.N.Arora, (1997). A foundation course statistics, S.Chand & Company Ltd, New Delhi.

Probability

DEFINITION: A probability experiment is an action, or trial, through which specific results (counts, measurements or responses) are obtained. The result of a single trial in a probability experiment is an outcome. The set of all possible outcomes of a probability experiment is the sample space. An event consists of one or more outcomes and is a subset of the sample space.

Example 1 The experiment consists of tossing a coin then rolling a die. the sample space consists of

H						T					
1	2	3	4	5	6	1	2	3	4	5	6
H1	H2	H3	H4	H5	H6	T1	T2	T3	T4	T5	T6

How many outcomes are there? Do you agree, disagree, or have no opinion, and what is your gender?
An event that consists of a single outcome is called a simple event

DEFINITION: Classical (or theoretical) is used when each outcome in a sample space is equally likely to occur. The Classical probability of an event E is given by:

$$P(E)=\frac{\text{Number of outcomes in } E}{\text{Total number of outcomes in sample space}}$$

Example 3 Roll a die: What is the sample space? {1,2,3,4,5,6}
Event A: rolling a 3, p = 1/6 = 0.157. Note this is a simple event.
Event C: rolling < 5, p =4/6 = 0.667. Note this is not a simple event.

DEFINITION Empirical (or statistical) probability is based on observations obtained from probability experiments. The empirical probability of an event E is the relative frequency of event E:

$$P(E)=\frac{\text{Frequency of event } E}{\text{Total frequency}}=\frac{f}{n}$$

Example: Finding Empirical Probabilities . Each fish (Bluegill, Redgill, and Crappy) is equally likely to get caught. You catch and release the following.

Fish Type	Number of times caught, *f*
Bluegill	13
Redgill	17
Crappy	10
	$\Sigma f = 40$

Probability of catching a bluegill = 13/40 = 0.325

Law of Large Numbers (p 114): As an experiment is repeated over and over, the empirical probability of the event approaches the theoretical (actual) probability of the event.

Basic Concepts of Probability

In statistics, an experiment is a process leading to at least two possible outcomes with uncertainty as to which will occur.

The set of all possible outcomes of an experiment is called the sample space (S). Each outcome in S is called a sample point.

Example 1

Three items are selected at random from a manufacturing process. Each item is inspected and classified defective (D) or non-defective (N).

An event is a subset of a sample space, it consists of one or more outcomes with a common characteristic.

Example 2

The event that the number of defectives in above example is greater than 1.

The null space or empty space is a subset of the sample space that contains no outcomes (∅).

The intersection of two events A and B denoted by (A∩B) is the event containing all outcomes that are common to A and B.

Events are mutually exclusive if they have no elements in common.

The union of two events A and B denoted by (A∪B) is the event containing all the elements that belong to A or to B or to both.

Events are collectively exhaustive if no other outcome is possible for a given experiment.

The complement of an event A with respect to S is the set of outcomes of S that are not in A denoted by $(A' \text{ or } A^c \text{ or } \{\bar{A}\})$.

Probability of an Event

Notation : P(A) The probability of an event A

Probability postulates :

1. P(S) = 1
2. P(∅) = 0
3. 0 ≤ P(A) ≤ 1

Methods of Assigning Probabilities

1. The classical approach

   If an experiment can result in any one of N different equally likely outcomes, and if exactly n of these outcomes correspond to event A, then

$$P(A)=\frac{n}{N}$$

2. The relative frequency approach

   If some number N of experiments are conducted and the event A occurs in N_A of them, then

$$P(A)=\lim_{N\to ¥}\frac{N_A}{N}$$

3. The subjective approach

   Subjective probability is a personal assessment of the likelihood of an event.

Principle of Counting - Permutation and Combination

Counting Sample Points

Fundamental principle of counting:

If an operation can be performed in N_1 ways, a second operation can be performed in N_2 ways, and so forth, then the sequence of k operations can be performed in

$$N_1\ N_2\ N_3\ \ldots\ N_k \quad \text{ways}$$

Example 3

Suppose a licence plates containing two letters following by three digits with the first digit not zero. How many different licence plates can be printed?

A <u>permutation</u> is an arrangement of all or part of a set of objects.

Example 4

The possible permutations from 3 letters A, B, C

The number of permutations of n distinct objects is n!.

The number of permutations of n distinct objects taken r at a time is

$$nPr = \frac{n!}{(n-r)!}$$

Example 5

In how many ways can 10 people be seated on a bench if only 4 seats are available?

The <u>combination</u> is a collection of n objects taken r at a time in any selections of r objects where order does not count. The number of combinations of n objects taken r at a time is

$$nCr = \frac{n!}{r!(n-r)!}$$

Example 6

A box contains 8 eggs, 3 of which are rotten. Three eggs are picked at random. Find the probabilities of the following events.

a) Exactly two eggs are rotten.

b) All eggs are rotten.

c) No egg is rotten.

Addition Rule and Complimentary Rule

Addition Rule

1. For events that are not mutually exclusive

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Example 7

A card is drawn from a complete deck of playing cards. What is the probability that the card is a heart or an ace?

For mutually exclusive events

$$P(A \cup B) = P(A) + P(B)$$

Complementary Rule

If A and A′ are complementary events then

$$P(A') = 1 - P(A)$$

Conditional Probability, Statistically Independence and Multiplication Rule

Conditional Probability

Let A and B be two events. The conditional probability of event A, given event B, denoted by P(A|B) is defined as

$$P(A|B)=\frac{P(A\cap B)}{P(B)}$$ provided that P(B) > 0.

Similarly, the conditional probability of B given A is defined as

$$P(B|A)=\frac{P(A\cap B)}{P(A)}$$ provided that P(A) > 0.

Statistically Independence

Two events are independent when the occurrence or non-occurrence of one event has no effect on the probability of occurrence of the other event.

Definition : Two events A and B are independent if and only if

$P(A\cap B) = P(A)P(B)$

Multiplication Rule

1. For dependent events

   $P(A\cap B) = P(A)P(B|A)$ or

   $=P(B)P(A|B)$

2. For independent events

   $P(A\cap B) = P(A)P(B)$

Theorem of Total Probability

If the events $B_1, B_2, \ldots, B_k$ constitute a partition of the sample space S such that $P(B_i) \neq 0$ $(i = 1, \ldots, k)$ then for any event A of S

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \ldots + P(B_k)P(A|B_k)$$

Baye's Theorem

If $B_1, B_2, \ldots, B_k$ are mutually exclusive events such that $B_1 \cup B_2 \cup \ldots \cup B_k$ contains all sample points of S, then for any event A of S with $P(A) \neq 0$,

$$P(B_i|A)=\frac{P(B_i)P(A|B_i)}{P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+\ldots+P(B_k)P(A|B_k)}$$

for i = 1, 2, 3, …, k

Possible Questions

Unit II

PART-B

1. Explain the functions of Random variable by an example.
2. Write a short note about the following terms:
   i. Conditional Probability
   ii. Independent and Dependent Event
   iii. Mutually Exclusive Event
3. Explain the Characteristics of Random variable with an example.
4. Write a short note about the following terms:
   a) Event and Mutually Exclusive Event
   b) Exclusive and Exhaustive Events
   c) Dependent and Independent Events
   d) Simple and Compound Events
5. (i) Explain the of axioms of the theory probability.
   (ii) State and prove Bayes theorem
6. State and prove the Addition and Multiplication theorems of Probability
7. Write a short note about the following terms:
   i) Random Event and Independent Event
   ii) Differentiate between Permutation and Combination.

PART-C

8. Write a short note about the following terms:
   a) Conditional Probability
   b) Bayes' Theorem
   c) Event and Mutually Exclusive Event
   d) Exclusive and Exhaustive Events
   e) Dependent and Independent Events

Question	Opt 1	Opt 2	Opt 3	Opt 4	Answer
The regression line cut each other at the point of-------	Average of X only	Average of Y only	Average of X and Y	the median of X on Y	Average of X and Y
Given the coefficient of correlation being 0.8, the coefficient of determination will be	0.98	0.64	0.66	0.54	0.64
Given the coefficient of correlation being 0.9, the coefficient of determination will be	0.98	0.81	0.66	0.54	0.81
If the coefficient of determination being 0.49, what is the coefficient of correlation	0.7	0.8	0.9	0.6	0.7
Given the coefficient of determination being 0.36, the coefficient of correlation will be	0.3	0.4	0.6	0.5	0.6
Which one of the following refers the term Correlation?	Relationship between two values	Relationship between two variables	Average relationship between two variables	Relationship between two things	Relationship between two variables
If r = +1, then the relationship between the given two variables is	perfectly positive	perfectly negative	no correlation	high positive	perfectly positive
If r = - 1, then the relationship between the given two variables is	perfectly positive	perfectly negative	no correlation	low Positive	perfectly negative
If r = 0, then the relationship between the given two variables is	Perfectly positive	perfectly negative	no correlation	both positive and negative	no correlation
Coefficient of correlation value lies between	1 and –1	0 and 1	0 and ∞	0 and –1.	1 and –1
While drawing a scatter diagram if all points appear to form a straight line getting Downward from left to right, then it is inferred that there is	Perfect positive correlation	simple positive correlation	Perfect negative correlation	no correlation	Perfect negative correlation
The range of the rank correlation coefficient is	0 to 1	–1 to 1	0 to ∞	– ∞ to ∞	–1 to 1
If r =1, then the angle between two lines of regression is	Zero degree	sixty degree	ninety degree	thirty degree	ninety degree

Regression coefficient is independent of	Origin	scale	both origin and scale	neither origin nor scale.	Origin
If the correlation coefficient between two variables X and Y is negative, then the Regression coefficient of Y on X is	Positive	negative	not certain	zero	negative
If the correlation coefficient between two variables X and Y is positive, then the Regression coefficient of X on Y is	Positive	negative	not certain	zero	Positive
There will be only one regression line in case of two variables if	r =0	r = +1	r = –1	r is either +1 or –1	r =0
The regression line cut each other at the point of	Average of X only	Average of Y only	Average of X and Y	the median of X on Y	Average of X and Y
If b_{xy} and b_{yx} represent regression coefficients and if $b_{yx} > 1$ then b_{xy} is	Less than one	greater than one	equal to one	equal to zero	Less than one
Rank correlation was discovered by	R.A.Fisher	Sir Francis Galton	Karl Pearson	Spearman	Spearman
Formula for Rank correlation is	1- ($6\Sigma d^2$ /(n(n2-1)))	1- ($6\Sigma d^2$ /(n(n2+1)))	1+ ($6\Sigma d^2$ /(n(n2+1)))	1 /(n(n2-1))	1- ($6\Sigma d^2$ /(n(n2-1)))
With b_{xy}=0.5, r = 0.8 and the variance of Y=16, the standard deviation of X=	6.4	2.5	10	25.6	2.5
The coefficient of correlation r =	$(b_{xy.}\, b_{yx})^{1/4}$	$(b_{xy}\, .b_{yx})^{-1/2}$	$(b_{xy}\, .b_{yx})^{1/3}$	$(b_{xy.}\, b_{yx})^{1/2}$	$(b_{xy.}\, b_{yx})^{1/2}$
If two regression coefficients are positive then the coefficient of correlation must be	Zero	negative	positive	one	positive
If two-regression coefficients are negative then the coefficient of correlation must be	Positive	negative	zero	one	Positive
The regression equation of X on Y is	X = a + bY	X = a + bX	X= a - bY	Y= a + bX	X = a + bY
The regression equation of Y on X is	X = a + bY	X = a + bX	X= a - bY	Y= a + bX	Y= a + bX
The given two variables are perfectly positive, if	r = +1	r = -1	r = 0	$r \neq +1$	r = +1
The relationship between two variables by plotting the values on a chart, known as-	coefficient of correlation	Scatter diagram	Correlogram	rank correlation	Scatter diagram

If x and y are independent variables then,	$cov(x,y) \neq 0$	$cov(x,y)= 1$	$cov(x,y)= 0$	$cov(x,y) >1$	$cov(x,y)= 0$
Correlation coefficient is the ----------- of the two regression coefficients.	Mode	Geometric mean	Arithmetic mean	median	Geometric mean
$b_{xy} = 0.4$, $b_{yx} = 0.9$ then r =	0.6	0.3	0.1	-0.6	0.6
$b_{xy}=1/5$, $r=8/15$, $s_x = 5$ then $s_y=$	40/13	13/40	40/3	3	40/3
The geometric mean of the two regression coefficients.	Correlation coefficient	regression coefficients	coefficient of range	coefficient of variation	Correlation coefficient
If two variables are uncorrelated, then the lines of regression	Do not exist	coincide	Parallel to each other	perpendicular to each other	perpendicular to each other
If the given two variables are correlated perfectly negative, then	$r = +1$	$r = -1$	$r = 0$	$r \neq +1$	$r = -1$
If the given two variables have no correlation, then	$r = +1$	$r = -1$	$r = 0$	$r \neq +1$	$r = 0$
If the correlation coefficient between two variables X and Y is ---------, the Regression coefficient of Y on X is positive	Negative	positive	not certain	zero	positive
If the correlation coefficient between two variables X and Y is ---------, the Regression coefficient of Y on X is negative	Negative	positive	not certain	zero	Negative
------------ is independent of origin and scale.	Correlation coefficient	regression coefficients	coefficient of range	coefficient of variation	Correlation coefficient
The angle between two lines of regression is ninety degree, if -------------	$r = 2$	$r = 0$	$r = 1$	$r = -1$	$r = 1$
--------- is used to measure closeness of relationship between variables.	Regression	mean	Rank correlation	correlation	correlation
If r is either +1 or –1, then there will be only one -------- line in case of two variables	Correlation	regression	rank correlation	mean	regression
When $b_{xy}=0.85$ and $b_{yx}= 0.89$, then correlation coefficient r =	0.98	0.5	0.68	0.87	0.87

If b_{xy} and b_{yx} represent regression coefficients and if $b_{xy} < 1$, then b_{yx} is	less than 1	greater than one	equal to one	equal to zero	greater than one
While drawing a scatter diagram if all points appear to form a straight line getting Downward from left to right, then it is inferred that there is--------	Perfect positive correlation	simple positive correlation	Perfect negative correlation	no correlation	Perfect negative correlation
If r =1, the angle between two lines of regression is----------	Zero degree	sixty degree	ninety degree	thirty degree	ninety degree
Regression coefficient is independent of--------	Origin	scale	both origin and scale	neither origin nor scale.	Origin
There will be only one regression line in case of two variables if------------	r =0	r = +1	r = –1	r is either +1 or –1	r =0

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established under Section 3 of UGC Act, 1956)
Pollachi Main Road, Coimbatore – 641 021, Tamil Nadu

Department of Mathematics

Subject : Probability and Statistics **Semester VI** **L T P C**
Subject Code : 16MMU603A **Class : III B.Sc Mathematics** **4 0 0 4**

UNIT III

Discrete distributions: uniform distribution, binomial distribution, Poisson distribution, and its properties, joint probability density functions, marginal and conditional distributions, expectation of function of two random variables, conditional expectations, independent random variables.

SUGGESTED READINGS

TEXT BOOK

1. Gupta S.P., (2001). Statistical Methods, Sultan Chand & Sons, New Delhi.

REFERENCES

1. Robert V. Hogg, Joseph W. McKean and Allen T. Craig., (2007). Introduction to Mathematical
 Statistics, Pearson Education, Asia.
2. Irwin Miller and Marylees Miller, John E. Freund, (2006). Mathematical Statistics with Application, Seventh Edition, Pearson Education, Asia.
3. Sheldon Ross., (2007). Introduction to Probability Model, Ninth Edition, Academic Press, Indian Reprint.
4. Pillai R.S.N., and Bagavathi V., (2002). Statistics , S. Chand & Company Ltd, New Delhi.
5. Srivastava T.N., and Shailaja Rego., (2012). 2e, Statistics for Management, Mc Graw Hill Education, New Delhi.
6. Dr.P.N.Arora, (1997). A foundation course statistics, S.Chand & Company Ltd, New Delhi.

Sampling, Sampling Distributions

Samples vs. Populations

Population: A complete set of observations or measurements about which conclusions are to be drawn.

Sample: A subset or part of a population.

Not necessarily random

Statistics vs. Parameters

Parameter: A summary characteristic of a population.

Summary of Central tendency, variability, shape, correlation

E.g., Population mean, Population Standard Deviation, Population Median, Proportion of population of registered voters voting for Bush, Population correlation between Systolic & Diastolic BP

Statistic: A summary characteristic of a sample. Any of the above computed from a sample taken from the population.

E.g., Sample mean, Sample Standard Deviation, median, correlation coefficient

Inferential Statistics

We take a sample and compute a description of a characteristic of the sample – central tendency (usually), variability or shape. That is, we compute the value of a sample statistic.

We use the sample statistic to make an educated guess about the corresponding population parameter.

The basic concept is easy. The devil is in the details.

DEFINITIONS

- A random variable *X* represents a numerical value associated with each outcome of a probability experiment.
- A random variable is discrete if it has a finite or countable number of possible outcomes that can be listed.
- A random variable is continuous if it has an uncountable number of possible outcomes, represented by an interval on the number line.

The number of calls a salesperson makes in one day is an example of a discrete random variable, while the time in hours he spends making calls in one day is an example of a continuous random variable.

A discrete probability distribution lists each possible value the random variable can assume, together with its probability. A probability distribution must satisfy the following conditions:

The probability of each value of the discrete random variable is between 0 and 1: $0 \le P(x) \le 1$

The sum of all the probabilities is 1: $\sum P(x) = 1$

Guidelines for constructing a discrete probability distribution: (p164)

1. Make a frequency distribution for the possible outcomes
2. Find the sum of the frequencies
3. Find the probability of each possible outcome by dividing its frequency by the sum of the frequencies.
4. Check that each probability is between 0 and 1 and that the sum is 1.

Example (p 164) Individuals are rated on a score of 1 to 5 for passive-aggressive traits, where 1 is extremely passive and 5 is extremely aggressive.

Score, *X*	**Frequency, *f***	***P(X)***
1	24	0.16
2	33	0.22
3	42	0.28
4	30	0.2
5	21	0.14
Total	150	1.00

The mean (also called the expected value) of a discrete random variable is given by:

$$Expected\ Value = E(x) = \mu = \sum xP(x)$$

Note that each value of x is multiplied by its corresponding probability and the products are added.

Example: Find the mean for passive-aggressive traits above:

X	**P(X)**	**XP(X)**
1	0.16	1*0.16 = 0.16

2	0.22	2*0.22 = 0.44
3	0.28	3*0.28 = 0.84
4	0.2	4*0.20 = 0.80
5	0.14	5*0.14 = 0.70
	$\Sigma P(X) = 1$	$\Sigma XP(X) = 2.94$

The variance of a discrete random variable is the expected value of $(x-\mu)^2$:

$$\sigma^2 = E(x-\mu)^2 = \sum (x-\mu)^2 P(x)$$

The standard deviation is

$$\sigma = \sqrt{\sigma^2}$$

Example (p 167) Find the Variance and Standard Deviation of the passive-aggressive measure in the above example

X	**P(X)**	$x-\mu$	$(x-\mu)^2$	$P(x)(x-\mu)^2$
1	0.16	-1.94	3.764	0.602
2	0.22	-0.94	0.884	0.194
3	0.28	0.06	0.004	0.001
4	0.20	1.06	1.124	0.225
5	0.14	2.06	4.244	0.594
X	$\Sigma P(X) = 1$			$\Sigma P(x)(x-\mu)^2 = 1.616$

So,

$$Var(x) = \sigma^2 = 1.616$$

$$\sigma = \sqrt{1.616} \approx 1.27.$$

Binomial Distributions

A binomial experiment is a probability experiment that satisfies the following conditions:

1. The experiment is repeated for a fixed number of trials where each trial is independent of the other trials.
2. There are only two possible outcomes of interest for each trial. The outcomes can be classified as a success (S) or as a failure (F).
3. The probability of a success P(S) is the same for each trial.
4. The random variable x counts the number of successful trials.

Notation for Binomial Experiments

Symbol	Description
n	The number of times the trial is repeated
$p = P(S)$	The probability of success in a single trial
$q = P(F)$	The probability of failure in a single trial $(q = 1 - p)$
x	The random variable represents a count of the number of successes in n trials: $x = 0,1,2,3,\ldots,n$

Suppose we have 9 trials. If we let 0 mean failure and 1 mean success, the probability of getting the results: 0 0 1 0 1 1 0 0 0 is p^3q^6. (See Mood and Graybill p 66.) This is a specific way of getting 3 successes: on the third fifth and sixth tries. Each try can be viewed as a box, and the number of ways we can place 3 1's in 9 boxes is the same as the number of ways we can choose the first 3 players from 9 on a baseball team. This is $\binom{9}{3}$. In general the probability of a specific arrangement of *x* 1's and *n*-*x* 0's is $p^x q^{n-x}$ and there are $\binom{n}{x}$ arrangements. This leads to the following formula for the binomial distribution.

In a binomial experiment, the probability of exactly *x* successes in *n* trials is:

$$P(x) = {}_nC_x p^x q^{n-x} = \frac{n!}{(n-x)!x!} p^x q^{n-x} = \binom{n}{x} p^x q^{n-x}, x = 0,1,2,...,n$$

This is often referred to as $b(x; n, p)$.

We can also see how this formula is derived from a simple example: Suppose we perform have 3 trials. The possible results are:

Sample Points	*Probability of sample point*	*Value of x*
SSS	p^3	*3*
SSF	p^2q	*2*
SFS	p^2q	*2*
SFF	pq^2	*1*
FSS	p^2q	*2*
FSF	pq^2	*1*
FFS	pq^2	*1*
FFF	q^3	*0*

So, $P(0) = \binom{3}{0} q^3, P(1) = \binom{3}{1} pq^2, P(2) = \binom{3}{2} p^2 q, P(3) = \binom{3}{3} p^3$

Appendix B, Table 2 gives, for the binomial distribution, the probabilities of x successes in n trials, for values of n = 2-16,20 for x = 0 to n, for various probabilities of success.

Population Parameters of a Binomial Distribution
$\mu = np$
$\sigma^2 = npq$
$\sigma = \sqrt{npq}$

The following are derivations of the mean for n = 1 and 2. (Mendenhall p 123)

$n=1$	$E(x)=\sum_{x=0}^{1} xp(x)=0q+1p=p$
$n=2$	$E(x)=\sum_{x=0}^{2} xp(x)=0q^2+1*2pq+2p^2=2p(q+p)=2p$

The following is a derivation of the variance for n = 1. (Mendenhall p 123)

$$\sigma^2=E(x-\mu)^2=\sum_{x=0}^{1}(x-\mu)^2 p(x)=(0-p)^2 q+(1-p)^2 p=p^2 q+q^2 p=pq(q+p)=pq.$$

Performing the Test

To perform a hypothesis test, you must find a z-score based on the value of the parameter specified in the null hypothesis.

$$z=\left|\frac{\bar{x}-\mu_{H_o}}{\sigma_{\bar{x}}}\right|$$

Note that in forming this z-score, we are using the standard error of the mean in the denominator. That's because your sample mean is distributed normally with that standard deviation, not the standard deviation of the population as a whole. We can rewrite the above like so:

$$z=\left|\frac{x-\mu_{H_o}}{\sigma/\sqrt{n}}\right|$$

We will then compare this to a critical value of z from the standard normal table. If it's greater than the z-critical, we reject the null and accept the alternative hypothesis. Otherwise, we do not reject the null, nor do we accept the alternative.

Example: Let's do the two-tail test on CSUN's graduation time. Let's say we know the standard deviation of the population is 2 years, and we sampled 49 CSUN graduates and found a sample mean of 6.9. Then we calculate:

$$z=\frac{\bar{x}-\mu_{H_o}}{\sigma/\sqrt{n}}=\frac{6.9-6.5}{2/\sqrt{49}}=1.4$$

We need a z-critical value for a significance level of 0.10. Since this is a two-tail test, we want 0.05 in each tail, so find the value of z in Table 3 that gives you an area as close to 0.95 as possible. This turns out to be 1.64. Since $1.4 < 1.64$, we do not reject the null. The administration's claim cannot be rejected.

Example: Now let's do the one-tail test on CSUN's graduation time. All the calculations are the same, except now we want the whole 10% in the right tail. That gives us a z-critical of 1.28. Since $1.4 > 1.28$, we reject the null and accept the alternative. We think the administration has underestimated the true mean graduation time.

NOTE: The test we just did is a right-tail test, because the null hypothesis is rejected only for a sufficiently <u>high</u> sample mean. But what if the null hypothesis had been that CSUN's average graduation time was <u>greater</u> than or equal to 6.5? In that case we would have done a left-tail test. In addition to the z-value calculated above being greater than z-critical, you also need to make sure the sample mean is less than the hypothesized mean (6.5 in this case). Alternatively, just calculate the z-value above without absolute value signs, and then put a negative sign on your z-critical.

Why did we reject in the two-tail case and accept in the one-tail case? Because in the two-tail case, some of the weight of α had to go in the left tail, which turned out to be irrelevant in this case. That meant there was less weight to go in the right tail, and thus less chance of rejecting the null as a result of a high sample mean.

V. Getting Rid of the Bogus Assumptions

We assumed above that true standard deviation was known. Just as with CI's, this is a weird assumption. Why would we know the true standard deviation but not the true mean?
When we have a large sample, we can get away with substituting the sample standard deviation for the true one and continuing to use the z-distribution. This gives us the following z-score formula:

$$z = \left| \frac{\bar{x} - \mu_{H_o}}{s / \sqrt{n}} \right|$$

But what if you don't know the true standard deviation and the sample size is small? Then we have to use the t-distribution. We calculate a t-score instead of a z-score:

$$t = \left| \frac{\bar{x} - \mu_{H_o}}{s / \sqrt{n}} \right|$$

And then we find a t-critical value instead of a z-critical value.

Example: Same example as above, doing a one-tail test. But this time, we don't know the standard deviation is 2, and our sample size was only 17. Our sample standard deviation turns out to be 1.9, and we use this to find our t-score:

$$t = \frac{x - \mu_{H_o}}{\sigma / \sqrt{n}} = \frac{6.9 - 6.5}{1.9 / \sqrt{16}} = 0.84$$

In the t-table, with df = 16 -1 = 15 and 90% confidence level, t-critical is 1.75. We do not reject the null.

If we had wanted a one-tail test, we'd have looked in the column of the table headed by 0.1000 (ignore the 0.8000 confidence level below, because that assumes a two-tail test). We get 1.341. Since 0.84 < 1.341, we do not reject the null.

Example: In the above example, for the two-tail test, the p-value is 2(0.0808) = 0.1616. That's the lowest alpha that will lead to rejection of the null in the two-tail test.

It is possible to find p-values when we're using the t-statistics as well. But the t-table in a book doesn't give us enough information to find the p-value with much precision. A statistical software program can do it for us, though.

Three theoretical facts and one practical fact about the distribution of sample means . . .

The theoretical facts are about 1) central tendency, 2) variability, and 3) shape . . .

1. The mean of the population of sample means will be the same as the mean of the population from which the samples were taken. The mean of the means is the mean. $\mu_M = \mu$.

Implication: The sample mean is an unbiased estimate of the population mean. If you take a random sample from a population, it is just as likely to be smaller than the population mean as it is to be larger than the population mean.

2. The standard deviation of the population of sample means – called the standard error of the mean - will be equal to d original population's standard deviation divided by the square root of N, the size of each sample. (Corty, Eq. 5.1, p 142)

In Corty's notation,

$$\sigma_M = \frac{\sigma}{\sqrt{N}}$$

The standard deviation (σ_M) is called the standard error of the mean.

Implication: Means are less variable than individual scores. Means are likely to be closer to the population mean than individual scores. You can make a sample mean as close as you want to the population mean if you can afford a large sample.

3. The shape of the distribution of the population of sample means will be the normal distribution if the original distribution is normal or approach the normal as N gets larger in all other cases. This fact is called the Central Limit Theorem. It is the foundation upon which most of modern day inferential statistics rests. See Corty, p. 141.

Why do we care about #3: Because we'll need to compute probabilities associated with sample means when doing inferential statistics. To compute those probabilities, we need a probability distribution.

Practical fact

4. The distribution of Z's computed from each sample, using the formula

$$Z = \frac{\text{X-bar} - \mu_M}{\sigma}$$

$\sqrt{N}$

will be or approach (as sample size gets large) the Standard Normal Distribution with mean = 0 and SD = 1.

Another test question: What are three facts about the distribution of sample means – a fact about central, a fact about variability, and a fact about shape of the distribution of sample means?

Continuous (Normal) Probability Distributions

[NOTE: The following notes were compiled from previous notes used when I taught other statistics courses.]

CHAPTER OBJECTIVES

- Understand the difference between discrete and continuous distributions.
- Compute the mean and the standard deviation for a uniform distribution.
- Compute probabilities using the uniform distribution.
- List the characteristics of the normal probability distribution.
- Define and calculate *z* values.
- Determine the probability an observation will lie between two points using the standard normal distribution.
- Determine the probability an observation will be above or below a given value using the standard normal distribution.
- Use the normal distribution to approximate the binomial probability distribution.

CHARACTERISTICS OF A NORMAL PROBABILITY DISTRIBUTION

- The normal curve is bell-shaped and has a single peak at the exact center of the distribution.
- The arithmetic mean, median, and mode of the distribution are equal and located at the peak.
- Half the area under the curve is above and half is below this center point (peak).
- The normal probability distribution is symmetrical about its mean.

- It is asymptotic - the curve gets closer and closer to the x-axis but never actually touches it.

NOTE

You can also have normal distributions with the same means but different standard deviations.

You can also have normal distributions with the same standard deviation but with different means.

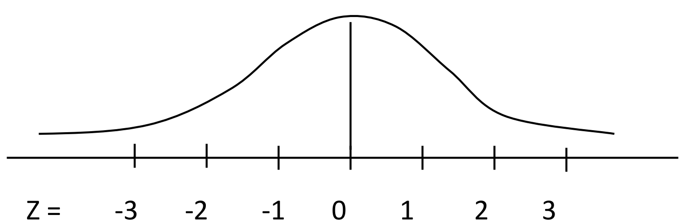
You can also have normal distributions with different means and different standard deviations.

THE STANDARD NORMAL PROBABILITY DISTRIBUTION

A normal distribution with a mean of 0 and a standard deviation of 1 is called the standard normal distribution.

z value: The distance between a selected value, designated X, and the population mean μ, divided by the population standard deviation, σ.

$$\mathbf{z} = \frac{(\mathbf{X} - \mu)}{\sigma} \qquad (5\text{-}2)$$



EXAMPLE 1: Monthly incomes of MBA Graduates.

The monthly incomes of recent MBA graduates in a large corporation are normally distributed with a mean of $2,000 and a standard deviation of $200. What is the z value for an income X of $2,200? $1,700?

For X = $2,200 and since $z = (X - \mu)/\sigma$, then $z = (2,200 - 2,000)/200 = 1$.

For X = $1,700 and since $z = (X - \mu)/\sigma$, then $z = (1,700 - 2,000)/200 = -1.5$.

A z value of 1 indicates that the value of $2,200 is 1 standard deviation above the mean of $2,000.

A z value of -1.5 indicates that the value of $1,700 is 1.5 standard deviations below the mean of $2,000.

AREAS UNDER THE NORMAL CURVE (See Section 5-3 “Applications of Normal Distributions.”)

From the Empirical Rule, we should remember the following:

About 68 percent of the area under the normal curve is within plus one and minus one standard deviation of the mean, written as $\mu \pm 1\sigma$.

About 95 percent of the area under the normal curve is within plus and minus two standard deviations of the mean, written $\mu \pm 2\sigma$.

Practically all (99.74 percent) of the area under the normal curve is within three standard deviations of the mean, written $\mu \pm 3\sigma$.

THE NORMAL APPROXIMATION TO THE BINOMIAL

Using the normal distribution (a continuous distribution) as a substitute for a binomial distribution (a discrete distribution) for large values of n seems reasonable because as n increases, a binomial distribution gets closer and closer to a normal distribution.

When do we use the normal approximation to the binomial distribution?

The normal probability distribution is generally deemed a good approximation to the binomial probability distribution when $np \geq 5$ and $nq \geq 5$. [Remember $q = 1 – p$.]

Recall for the binomial experiment:

1. There are only two mutually exclusive outcomes (success or failure) on each trial.
2. A binomial distribution results from counting the number of successes.
3. Each trial is independent.
4. The probability p is fixed from trial to trial, and the number of trials n is also fixed.

In order to use the normal distribution to approximate the binomial, we must use the mean and the standard deviation. Of the 200 homes sampled how many would you expect to have video recorders?

$\mu = np = (200)(0.15) = 30$. [Also $n(1 - p) = (200)(0.85) = 170$].

What is the variance?

$\sigma 2$ = np(1 - p) = (200)(0.15)(1- 0.15) = 25.5.

What is the standard deviation?

σ = $\sqrt{25.5}$ = 5.0498.

What is the probability that less than 40 homes in the sample have video recorders?

Form of H_0 and H_1 for one-sample mean:

H_0: $\mu = 115$

H_1: $\mu \neq 115$

- Hypotheses are always about population parameters, not sample statistics

H_0: μ = population value

H_1: $\mu \neq$ population value

- This hypothesis is a *non-directional* (two-tailed) hypothesis
- Null hypothesis: No effect
- Alternative hypothesis: Some effect (doesn't specify an increase or decrease)

Criterion for rejecting H_0: Creating a Decision Rule:

Will compute a test statistic (types vary based on data, design & question)

Then decide if the value of the test statistic is "improbable" under H_0

Traditionally, a test statistic is considered "unlikely" if it is expected to occur:

$\leq$ 5 in a 100: has a probability of .05 or less ($p \leq 0.05$)

- Look in *tails* of *sampling distribution for the unlikely outcomes*
- Divide distribution into two parts:

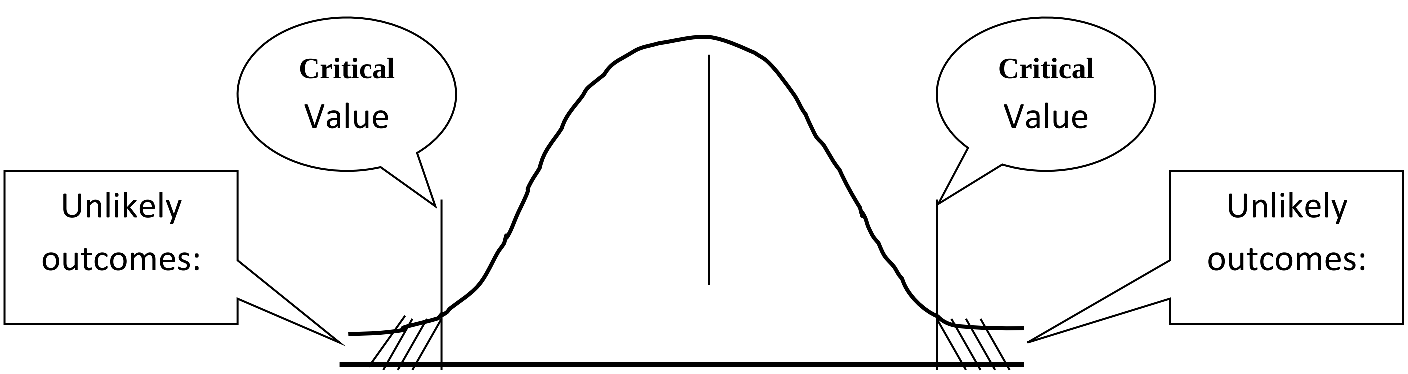
Values that are likely if H_0 is true	Values that are very unlikely if H_0 is true
Values close to H_0	Values far from H_0
Values in the *middle*	Values in the *tails*

Selecting a "significance level": α

Probability chosen as criteria for "unlikely"

Common convention: $\alpha = .05$ (5%)

May set a smaller α to be more conservative ($p \leq 0.01, 0.001$)



Critical value(s) = boundary(ies) b/n likely & unlikely outcomes

Rejection region = area(s) beyond critical value(s); outcomes that lead to a rejection of H_0

Decision rule:

Reject H_0 when *observed* test-statistic equals or exceeds *Critical value*

…when statistic falls in the *rejection region*

Otherwise, Fail to Reject (Retain) H_0

Collect data and Calculate "observed" test statistic:

z-test for one sample mean:

$$z = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}}$$

z = sample mean – hypothesized population μ

standard error

z = observed difference

difference due to chance

Don't forget to compute standard error first!

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

Compare Test Statistic to Critical Values:

- Does observed z equal or exceed CV?

  (Does it fall in the *rejection region*?)

- If YES,

  Reject H_0 = "statistically significant" finding

- If NO,

  Fail to Reject H_0 = "non-significant" finding

Interpret results:

- Return to research question
- *statistical significance* = not likely to be due to chance
- Never "prove" H_0 or H_1

Summary of Statistical Hypothesis Testing:

1. Formulate a research question
2. Formulate a research/alternative hypothesis
3. Formulate the null hypothesis
4. Collect data
5. Reference a sampling distribution of the particular statistic assuming that H_0 is true (in the cases so far, a sampling distribution of the mean)
6. Decide on a significance level (α), typically .05
7. Identify the critical value(s)
8. Compute the appropriate test statistic
9. Compare the test statistic to the critical value(s)
10. Reject H_0 if the test statistic is equal to or exceeds the critical value, retain otherwise

Possible Questions

PART-B

1. State properties and applications of *t*- distribution?
2. A group of 5 patients treated with medicine A weigh 39, 48, 60 and 41 kgs; second group of 7 patients treated with medicine B weigh 38, 42, 56, 64, 68, 69, 62 kgs. Do you agree with claim that medicine B increases weight significantly?
   (Use $\alpha = 5\%$ and $t_{0.05} = 1.812$)
3. Write the properties normal distribution?
4. Certain pesticide is packed into bags by a machine. Random samples of 10 bags are drawn and their contents are found to weigh (in kg) as follows.
   50 49 52 44 45 48 46 45 49 45
   Test if the average packing can be taken to be 50 kg.
5. Certain brand of rice is packed into bags by a machine. Random samples of 15 bags are drawn and their contents are found to weigh (in kg) as follows.
   50 49 52 44 45 48 46 45 49 45 50 52 54 53 51
   Test if the average packing can be taken to be 50 kg.
6. Mention Snedecor's F-distribution, its properties and applications?

PART-C

7. i) Define Normal Distribution and write its important characteristics.
   ii) Describe the characteristics of χ^2 - distribution.

Question	Opt 1	Opt 2	Opt 3	Opt 4	Answer
Completely randomized design is similar to ----------	three way	one way	two way	t test	one way
Randomized block design is similar to ----------	two way	three way	one way	many	two way
ANOVA is the technique of analysis of ------	standard deviation	variance	mean	range	variance
Under one way classification , the influence of only ----- attribute or factor is considered	two	three	one	many	one
Under two way classification , the influence of only ----- attribute or factor is considered	four	two	three	one	two
The word ---------------- is used to indicate various statistical measures like mean, standard deviation, correlation etc, in the universe.	Statistic	parameter	hypothesis	none of these	parameter
The term STATISTIC refers to the statistical measures relating to the -----------.	Population	hypothesis	sample	universe	sample
A hypothesis may be classified as -----------------.	Simple	Composite	null	all the above	all the above
Level of significance is the probability of---------------.	Type I error	Type II error	Not committing error	any of the above	any of the above
Degrees of freedom are related to -------------------.	No. of observations in a set	hypothesis under test	No. of independent observations in a set	No. of dependent observations in a set	No. of independent observations in a set
A critical function provides the basis for ----------------.	Accepting H_0	rejecting H_0	no decision about H_0	all the above	all the above

Student's t-test is applicable in case of -------------------.	Small samples	for sample of size between 5 and 30	Large samples	none of the above	Small samples
Student's t-test is applicable only when ---------------------.	The variate values are independent	the variable is distributed normally	The sample is not large	all the above	all the above
If the calculated value is less than the table value then we accept the --------------- hypothesis.	Alternative	null	both	sample	null
Small sample test is also known as -------------------.	Exact test	t – test	normal test	F-test	t – test
The formula for c^2 is -------------	å(O–E) ^{2}/E	å(E+O) ^{2}/E	å(O-E) /E	å(O-E) ^{2}/O	å(O–E) ^{2}/E
If a statistic 't' follows student's t distribution with n degrees of freedom then t^2 follows ------	c^2distribution with (n-1) degrees of freedom	c^2distribution with n degrees of freedom	c^2distribution with n^2 degrees of freedom	c^2distribution with (n+1) degrees of freedom	c^2distribution with (n-1) degrees of freedom
The distribution used to test goodness of fit is------------	F distribution	c^2distribution	t distribution	Z distribution	c^2distribution
Degree of freedom for statistic chi-square incase of contingency table of order 2x2 is----------	3	4	2	1	1
Larger group from which the sample is drawn is called ------------.	Sample	sampling	universe	parameter	universe
Any hypothesis concerning a population is called a ----------------.	Sample	population	statistical measure	statistical hypothesis	statistical hypothesis
Rejecting Ho when it is true leads ----------.	Type I error	Type II error	correct decision	either (a) or (b)	Type I error
Accept Ho when it is true leads -----------.	Type I error	Type II error	correct decision	either (a) or (b)	correct decision
Type II error occurs only if ----------.	Reject Ho when it is true	Accept Ho when it is false	Accept Ho when it is true	reject Ho when it is false	Accept Ho when it is false

The correct decision is ---------------.	Reject Ho when it is true	Accept Ho when it is false	Reject Ho when it is false	none of these	Reject Ho when it is false
The maximum probability of committing type I error, which we specified in a test is known as -----------------.	Null hypothesis	alternative hypothesis	DOF	level of significance	level of significance
If the computed value is less than the critical value, then ---------------.	Null hypothesis is accepted	Null hypothesis is rejected	Alternative hypothesis is accepted	population	Null hypothesis is accepted
If the computed value is greater than the critical value, then ---------------.	Null hypothesis is accepted	Null hypothesis is rejected	Alternative hypothesis is accepted	small sample	Null hypothesis is rejected
In sampling distribution the standard error is ----------------.	np	pq	npq	sqrt(npq)	sqrt(npq)
If the sample size is greater than 30, then the sample is called --------------.	Large sample	small sample	population	Null hypothesis	Large sample
If the sample size is less than 30, then the sample is called --------------.	Large sample	small sample	population	alternative hypothesis	small sample
Z – test is applicable only when the sample size is -----------.	zero	one	small	large	large
The degrees of freedom for two samples in t – test is ---------.	$n_1 + n_2 + 1$	$n_1 + n_2 – 2$	$n_1 + n_2 + 2$	$n_1 + n_2 – 1$	$n_1 + n_2 – 2$
An assumption of t – test is population of the sample is -----------.	Binomial	Poisson	normal	exponential	normal
The degrees of freedom of chi – square test is ------------.	(r – 1)(c – 1)	(r + 1)(c + 1)	(r + 1)(c – 1)	(r – 1)(c + 1)	(r – 1)(c – 1)
In chi – square test, if the values of expected frequency are less than 5, then they are combined together with the neighbouring frequencies. This is known as ----------.	Goodness of fit	DOF	LOS	pooling	pooling

The expected frequency of chi – square test can be calculated as ----------.	(RT + CT) / GT	(RT – CT) / GT	(RT * CT) / GT	(RT*CT)	(RT * CT) / GT
In F – test, the variance of population from which samples are drawn are ----------.	equal	not equal	small	large	equal
If the data is given in the form of a series of variables, then the DOF is ---------.	n	n-1	n+1	(r – 1)(c – 1)	n-1
The characteristic of the chi–square test is -------------.	DOF	LOS	ANOVA	independence of attributes	independence of attributes
If $S_1^2 > S_2^2$, then the F – statistic is ----------.	S_1 / S_2	S_2 / S_1	S_1^2 / S_2^2	S_1^3 / S_2^3	S_1^2 / S_2^2
The value of Z test at 5% level of significance is ------------.	3.96	2.96	1.96	0.96	1.96
In -------, the variance of population from which samples are drawn are equal	t-test	Chi-Square test	Z-test	F-test	F-test
F – statistics is ---------------------.	Variance between the samples / variance within the samples	Variance within the samples / variance between the samples	Variance between the rows / variance between the columns	Variance within the rows / variance within the columns	Variance between the samples / variance within the samples
Analysis of variance utilizes:	t-test	Chi-Square test	Z-test	F-test	F-test
F – test whish is also known as ------	Chi-Square test	Z-test	varience ratio test	t-test	varience ratio test
The technique of analysis of variance refered to as ------	ANOVA	F – test	Z – test	Chi- square test	ANOVA
The two variations, variation within the samples and variations between the samples are tested for their significance by -------	Chi- square test	F – test	t-test	Z – test	F – test

Under ------- classification , the influence of only one attribute or factor is considered.	two way	three way	one way	many	one way
Under --------- classification , the influence of two attribute or factors is considered	two way	three way	one way	many	two way

Chapter 8

Law of Large Numbers

8.1 Law of Large Numbers for Discrete Random Variables

We are now in a position to prove our first fundamental theorem of probability. We have seen that an intuitive way to view the probability of a certain outcome is as the frequency with which that outcome occurs in the long run, when the experiment is repeated a large number of times. We have also defined probability mathematically as a value of a distribution function for the random variable representing the experiment. The Law of Large Numbers, which is a theorem proved about the mathematical model of probability, shows that this model is consistent with the frequency interpretation of probability. This theorem is sometimes called the *law of averages.* To find out what would happen if this law were not true, see the article by Robert M. Coates.[1]

Chebyshev Inequality

To discuss the Law of Large Numbers, we first need an important inequality called the *Chebyshev Inequality.*

Theorem 8.1 (Chebyshev Inequality) Let X be a discrete random variable with expected value $\mu = E(X)$, and let $\epsilon > 0$ be any positive real number. Then

$$P(|X - \mu| \geq \epsilon) \leq \frac{V(X)}{\epsilon^2}\ .$$

Proof. Let $m(x)$ denote the distribution function of X. Then the probability that X differs from μ by at least ϵ is given by

$$P(|X - \mu| \geq \epsilon) = \sum_{|x-\mu|\geq\epsilon} m(x)\ .$$

[1]R. M. Coates, "The Law," *The World of Mathematics,* ed. James R. Newman (New York: Simon and Schuster, 1956.

We know that

$$V(X) = \sum_x (x-\mu)^2 m(x) \ ,$$

and this is clearly at least as large as

$$\sum_{|x-\mu|\geq\epsilon} (x-\mu)^2 m(x) \ ,$$

since all the summands are positive and we have restricted the range of summation in the second sum. But this last sum is at least

$$\begin{aligned}\sum_{|x-\mu|\geq\epsilon} \epsilon^2 m(x) &= \epsilon^2 \sum_{|x-\mu|\geq\epsilon} m(x) \\ &= \epsilon^2 P(|X-\mu| \geq \epsilon) \ .\end{aligned}$$

So,

$$P(|X-\mu| \geq \epsilon) \leq \frac{V(X)}{\epsilon^2} \ .$$

□

Note that X in the above theorem can be any discrete random variable, and ϵ any positive number.

Example 8.1 Let X by any random variable with $E(X) = \mu$ and $V(X) = \sigma^2$. Then, if $\epsilon = k\sigma$, Chebyshev's Inequality states that

$$P(|X-\mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2} \ .$$

Thus, for any random variable, the probability of a deviation from the mean of more than k standard deviations is $\leq 1/k^2$. If, for example, $k = 5$, $1/k^2 = .04$. □

Chebyshev's Inequality is the best possible inequality in the sense that, for any $\epsilon > 0$, it is possible to give an example of a random variable for which Chebyshev's Inequality is in fact an equality. To see this, given $\epsilon > 0$, choose X with distribution

$$p_X = \begin{pmatrix} -\epsilon & +\epsilon \\ 1/2 & 1/2 \end{pmatrix} \ .$$

Then $E(X) = 0$, $V(X) = \epsilon^2$, and

$$P(|X-\mu| \geq \epsilon) = \frac{V(X)}{\epsilon^2} = 1 \ .$$

We are now prepared to state and prove the Law of Large Numbers.

Law of Large Numbers

Theorem 8.2 (Law of Large Numbers) Let X_1, X_2, ..., X_n be an independent trials process, with finite expected value $\mu = E(X_j)$ and finite variance $\sigma^2 = V(X_j)$. Let $S_n = X_1 + X_2 + \cdots + X_n$. Then for any $\epsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \to 0$$

as $n \to \infty$. Equivalently,

$$P\left(\left|\frac{S_n}{n} - \mu\right| < \epsilon\right) \to 1$$

as $n \to \infty$.

Proof. Since X_1, X_2, ..., X_n are independent and have the same distributions, we can apply Theorem 6.9. We obtain

$$V(S_n) = n\sigma^2 ,$$

and

$$V(\frac{S_n}{n}) = \frac{\sigma^2}{n} .$$

Also we know that

$$E(\frac{S_n}{n}) = \mu .$$

By Chebyshev's Inequality, for any $\epsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2} .$$

Thus, for fixed ϵ,

$$P\left(\left|\frac{S_n}{n} - \mu\right| \geq \epsilon\right) \to 0$$

as $n \to \infty$, or equivalently,

$$P\left(\left|\frac{S_n}{n} - \mu\right| < \epsilon\right) \to 1$$

as $n \to \infty$. $\square$

Law of Averages

Note that S_n/n is an average of the individual outcomes, and one often calls the Law of Large Numbers the "law of averages." It is a striking fact that we can start with a random experiment about which little can be predicted and, by taking averages, obtain an experiment in which the outcome can be predicted with a high degree of certainty. The Law of Large Numbers, as we have stated it, is often called the "Weak Law of Large Numbers" to distinguish it from the "Strong Law of Large Numbers" described in Exercise 15.

Consider the important special case of Bernoulli trials with probability p for success. Let $X_j = 1$ if the jth outcome is a success and 0 if it is a failure. Then $S_n = X_1 + X_2 + \cdots + X_n$ is the number of successes in n trials and $\mu = E(X_1) = p$. The Law of Large Numbers states that for any $\epsilon > 0$

$$P\left(\left|\frac{S_n}{n} - p\right| < \epsilon\right) \to 1$$

as $n \to \infty$. The above statement says that, in a large number of repetitions of a Bernoulli experiment, we can expect the proportion of times the event will occur to be near p. This shows that our mathematical model of probability agrees with our frequency interpretation of probability.

Coin Tossing

Let us consider the special case of tossing a coin n times with S_n the number of heads that turn up. Then the random variable S_n/n represents the fraction of times heads turns up and will have values between 0 and 1. The Law of Large Numbers predicts that the outcomes for this random variable will, for large n, be near 1/2.

In Figure 8.1, we have plotted the distribution for this example for increasing values of n. We have marked the outcomes between .45 and .55 by dots at the top of the spikes. We see that as n increases the distribution gets more and more concentrated around .5 and a larger and larger percentage of the total area is contained within the interval $(.45, .55)$, as predicted by the Law of Large Numbers.

Die Rolling

Example 8.2 Consider n rolls of a die. Let X_j be the outcome of the jth roll. Then $S_n = X_1 + X_2 + \cdots + X_n$ is the sum of the first n rolls. This is an independent trials process with $E(X_j) = 7/2$. Thus, by the Law of Large Numbers, for any $\epsilon > 0$

$$P\left(\left|\frac{S_n}{n} - \frac{7}{2}\right| \geq \epsilon\right) \to 0$$

as $n \to \infty$. An equivalent way to state this is that, for any $\epsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - \frac{7}{2}\right| < \epsilon\right) \to 1$$

as $n \to \infty$. □

Numerical Comparisons

It should be emphasized that, although Chebyshev's Inequality proves the Law of Large Numbers, it is actually a very crude inequality for the probabilities involved. However, its strength lies in the fact that it is true for any random variable at all, and it allows us to prove a very powerful theorem.

In the following example, we compare the estimates given by Chebyshev's Inequality with the actual values.

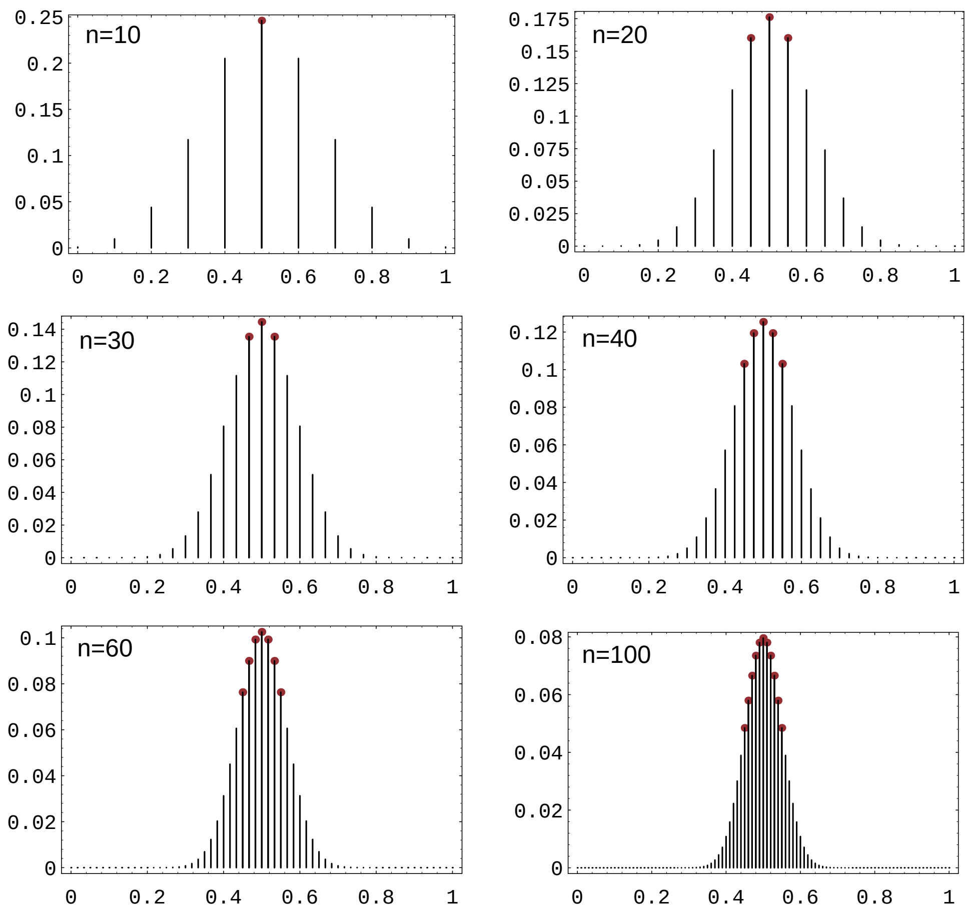


Figure 8.1: Bernoulli trials distributions.

Example 8.3 Let X_1, X_2, ..., X_n be a Bernoulli trials process with probability .3 for success and .7 for failure. Let $X_j = 1$ if the jth outcome is a success and 0 otherwise. Then, $E(X_j) = .3$ and $V(X_j) = (.3)(.7) = .21$. If

$$A_n = \frac{S_n}{n} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

is the *average* of the X_i, then $E(A_n) = .3$ and $V(A_n) = V(S_n)/n^2 = .21/n$. Chebyshev's Inequality states that if, for example, $\epsilon = .1$,

$$P(|A_n - .3| \geq .1) \leq \frac{.21}{n(.1)^2} = \frac{21}{n} \ .$$

Thus, if $n = 100$,

$$P(|A_{100} - .3| \geq .1) \leq .21 \ ,$$

or if $n = 1000$,

$$P(|A_{1000} - .3| \geq .1) \leq .021 \ .$$

These can be rewritten as

$$\begin{aligned} P(.2 < A_{100} < .4) &\geq .79 \ , \\ P(.2 < A_{1000} < .4) &\geq .979 \ . \end{aligned}$$

These values should be compared with the actual values, which are (to six decimal places)

$$\begin{aligned} P(.2 < A_{100} < .4) &\approx .962549 \\ P(.2 < A_{1000} < .4) &\approx 1 \ . \end{aligned}$$

The program **Law** can be used to carry out the above calculations in a systematic way. □

Historical Remarks

The Law of Large Numbers was first proved by the Swiss mathematician James Bernoulli in the fourth part of his work *Ars Conjectandi* published posthumously in 1713.[2] As often happens with a first proof, Bernoulli's proof was much more difficult than the proof we have presented using Chebyshev's inequality. Chebyshev developed his inequality to prove a general form of the Law of Large Numbers (see Exercise 12). The inequality itself appeared much earlier in a work by Bienaymé, and in discussing its history Maistrov remarks that it was referred to as the Bienaymé-Chebyshev Inequality for a long time.[3]

In *Ars Conjectandi* Bernoulli provides his reader with a long discussion of the meaning of his theorem with lots of examples. In modern notation he has an event

[2] J. Bernoulli, *The Art of Conjecturing IV,* trans. Bing Sung, Technical Report No. 2, Dept. of Statistics, Harvard Univ., 1966

[3] L. E. Maistrov, *Probability Theory: A Historical Approach,* trans. and ed. Samual Kotz, (New York: Academic Press, 1974), p. 202

that occurs with probability p but he does not know p. He wants to estimate p by the fraction $\bar{p}$ of the times the event occurs when the experiment is repeated a number of times. He discusses in detail the problem of estimating, by this method, the proportion of white balls in an urn that contains an unknown number of white and black balls. He would do this by drawing a sequence of balls from the urn, replacing the ball drawn after each draw, and estimating the unknown proportion of white balls in the urn by the proportion of the balls drawn that are white. He shows that, by choosing n large enough he can obtain any desired accuracy and reliability for the estimate. He also provides a lively discussion of the applicability of his theorem to estimating the probability of dying of a particular disease, of different kinds of weather occurring, and so forth.

In speaking of the number of trials necessary for making a judgement, Bernoulli observes that the "man on the street" believes the "law of averages."

> Further, it cannot escape anyone that for judging in this way about any event at all, it is not enough to use one or two trials, but rather a great number of trials is required. And sometimes the stupidest man—by some instinct of nature *per se* and by no previous instruction (this is truly amazing)— knows for sure that the more observations of this sort that are taken, the less the danger will be of straying from the mark.[4]

But he goes on to say that he must contemplate another possibility.

> Something futher must be contemplated here which perhaps no one has thought about till now. It certainly remains to be inquired whether after the number of observations has been increased, the probability is increased of attaining the true ratio between the number of cases in which some event can happen and in which it cannot happen, so that this probability finally exceeds any given degree of certainty; or whether the problem has, so to speak, its own asymptote—that is, whether some degree of certainty is given which one can never exceed.[5]

Bernoulli recognized the importance of this theorem, writing:

> Therefore, this is the problem which I now set forth and make known after I have already pondered over it for twenty years. Both its novelty and its very great usefullness, coupled with its just as great difficulty, can exceed in weight and value all the remaining chapters of this thesis.[6]

Bernoulli concludes his long proof with the remark:

> Whence, finally, this one thing seems to follow: that if observations of all events were to be continued throughout all eternity, (and hence the ultimate probability would tend toward perfect certainty), everything in

[4] Bernoulli, op. cit., p. 38.
[5] ibid., p. 39.
[6] ibid., p. 42.

the world would be perceived to happen in fixed ratios and according to a constant law of alternation, so that even in the most accidental and fortuitous occurrences we would be bound to recognize, as it were, a certain necessity and, so to speak, a certain fate.

I do now know whether Plato wished to aim at this in his doctrine of the universal return of things, according to which he predicted that all things will return to their original state after countless ages have past.[7]

Exercises

1 A fair coin is tossed 100 times. The expected number of heads is 50, and the standard deviation for the number of heads is $(100 \cdot 1/2 \cdot 1/2)^{1/2} = 5$. What does Chebyshev's Inequality tell you about the probability that the number of heads that turn up deviates from the expected number 50 by three or more standard deviations (i.e., by at least 15)?

2 Write a program that uses the function binomial(n, p, x) to compute the exact probability that you estimated in Exercise 1. Compare the two results.

3 Write a program to toss a coin 10,000 times. Let S_n be the number of heads in the first n tosses. Have your program print out, after every 1000 tosses, $S_n - n/2$. On the basis of this simulation, is it correct to say that you can expect heads about half of the time when you toss a coin a large number of times?

4 A 1-dollar bet on craps has an expected winning of $-.0141$. What does the Law of Large Numbers say about your winnings if you make a large number of 1-dollar bets at the craps table? Does it assure you that your losses will be small? Does it assure you that if n is very large you will lose?

5 Let X be a random variable with $E(X) = 0$ and $V(X) = 1$. What integer value k will assure us that $P(|X| \geq k) \leq .01$?

6 Let S_n be the number of successes in n Bernoulli trials with probability p for success on each trial. Show, using Chebyshev's Inequality, that for any $\epsilon > 0$

$$P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) \leq \frac{p(1-p)}{n\epsilon^2}\ .$$

7 Find the maximum possible value for $p(1-p)$ if $0 < p < 1$. Using this result and Exercise 6, show that the estimate

$$P\left(\left|\frac{S_n}{n} - p\right| \geq \epsilon\right) \leq \frac{1}{4n\epsilon^2}$$

is valid for any p.

[7] ibid., pp. 65–66.

8 A fair coin is tossed a large number of times. Does the Law of Large Numbers assure us that, if n is large enough, with probability $> .99$ the number of heads that turn up will not deviate from $n/2$ by more than 100?

9 In Exercise 6.2.15, you showed that, for the hat check problem, the number S_n of people who get their own hats back has $E(S_n) = V(S_n) = 1$. Using Chebyshev's Inequality, show that $P(S_n \geq 11) \leq .01$ for any $n \geq 11$.

10 Let X by any random variable which takes on values 0, 1, 2, …, n and has $E(X) = V(X) = 1$. Show that, for any integer k,

$$P(X \geq k+1) \leq \frac{1}{k^2} \ .$$

11 We have two coins: one is a fair coin and the other is a coin that produces heads with probability 3/4. One of the two coins is picked at random, and this coin is tossed n times. Let S_n be the number of heads that turns up in these n tosses. Does the Law of Large Numbers allow us to predict the proportion of heads that will turn up in the long run? After we have observed a large number of tosses, can we tell which coin was chosen? How many tosses suffice to make us 95 percent sure?

12 (Chebyshev[8]) Assume that $X_1, X_2, \ldots, X_n$ are independent random variables with possibly different distributions and let S_n be their sum. Let $m_k = E(X_k)$, $\sigma_k^2 = V(X_k)$, and $M_n = m_1 + m_2 + \cdots + m_n$. Assume that $\sigma_k^2 < R$ for all k. Prove that, for any $\epsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - \frac{M_n}{n}\right| < \epsilon\right) \to 1$$

as $n \to \infty$.

13 A fair coin is tossed repeatedly. Before each toss, you are allowed to decide whether to bet on the outcome. Can you describe a betting system with infinitely many bets which will enable you, in the long run, to win more than half of your bets? (Note that we are disallowing a betting system that says to bet until you are ahead, then quit.) Write a computer program that implements this betting system. As stated above, your program must decide whether to bet on a particular outcome before that outcome is determined. For example, you might select only outcomes that come after there have been three tails in a row. See if you can get more than 50% heads by your "system."

***14** Prove the following analogue of Chebyshev's Inequality:

$$P(|X - E(X)| \geq \epsilon) \leq \frac{1}{\epsilon} E(|X - E(X)|) \ .$$

[8] P. L. Chebyshev, "On Mean Values," *J. Math. Pure. Appl.,* vol. 12 (1867), pp. 177–184.

***15** We have proved a theorem often called the "Weak Law of Large Numbers." Most people's intuition and our computer simulations suggest that, if we toss a coin a sequence of times, the proportion of heads will really approach 1/2; that is, if S_n is the number of heads in n times, then we will have

$$A_n = \frac{S_n}{n} \to \frac{1}{2}$$

as $n \to \infty$. Of course, we cannot be sure of this since we are not able to toss the coin an infinite number of times, and, if we could, the coin could come up heads every time. However, the "Strong Law of Large Numbers," proved in more advanced courses, states that

$$P\left(\frac{S_n}{n} \to \frac{1}{2}\right) = 1 .$$

Describe a sample space Ω that would make it possible for us to talk about the event

$$E = \left\{ \omega : \frac{S_n}{n} \to \frac{1}{2} \right\} .$$

Could we assign the equiprobable measure to this space? (See Example 2.18.)

***16** In this problem, you will construct a sequence of random variables which satisfies the Weak Law of Large Numbers, but not the Strong Law of Large Numbers (see Exercise 15). For each positive integer n, let the random variable X_n be defined by

$$\begin{aligned} P(X_n = \pm n2^n) &= f(n) , \\ P(X_n = 0) &= 1 - 2f(n) , \end{aligned}$$

where $f(n)$ is a function that will be chosen later (and which satisfies $0 \leq f(n) \leq 1/2$ for all positive integers n). Let $S_n = X_1 + X_2 + \cdots + X_n$.

(a) Show that $\mu(S_n) = 0$ for all n.

(b) Show that if $X_n > 0$, then $S_n \geq 2^n$.

(c) Use part (b) to show that $S_n/n \to 0$ as $n \to \infty$ if and only if there exists an n_0 such that $X_k = 0$ for all $k \geq n_0$. Show that this happens with probability 0 if we require that $f(n) < 1/2$ for all n. This shows that the sequence $\{X_n\}$ does not satisfy the Strong Law of Large Numbers.

(d) We now turn our attention to the Weak Law of Large Numbers. Given a positive ϵ, we wish to estimate

$$P\left(\left|\frac{S_n}{n}\right| \geq \epsilon\right) .$$

Suppose that $X_k = 0$ for $m < k \leq n$. Show that

$$|S_n| \leq 2^{2m} .$$

(e) Show that if we define $g(n) = (1/2)\log_2(\epsilon n)$, then

$$2^{2m} < \epsilon n \ .$$

This shows that if $X_k = 0$ for $g(n) < k \leq n$, then

$$|S_n| < \epsilon n \ ,$$

or

$$\left| \frac{S_n}{n} \right| < \epsilon \ .$$

We wish to show that the probability of this event tends to 1 as $n \to \infty$, or equivalently, that the probability of the complementary event tends to 0 as $n \to \infty$. The complementary event is the event that $X_k \neq 0$ for some k with $g(n) < k \leq n$. Show that the probability of this event equals

$$1 - \prod_{k=\lceil g(n) \rceil}^{n} \bigl(1 - 2f(n)\bigr) \ ,$$

and show that this expression is less than

$$1 - \prod_{k=\lceil g(n) \rceil}^{\infty} \bigl(1 - 2f(n)\bigr) \ .$$

(f) Show that by making $f(n) \to 0$ rapidly enough, the expression in part (e) can be made to approach 1 as $n \to \infty$. This shows that the sequence $\{X_n\}$ satisfies the Weak Law of Large Numbers.

***17** Let us toss a biased coin that comes up heads with probability p and assume the validity of the Strong Law of Large Numbers as described in Exercise 15. Then, with probability 1,

$$\frac{S_n}{n} \to p$$

as $n \to \infty$. If $f(x)$ is a continuous function on the unit interval, then we also have

$$f\left(\frac{S_n}{n}\right) \to f(p) \ .$$

Finally, we could hope that

$$E\left(f\left(\frac{S_n}{n}\right)\right) \to E(f(p)) = f(p) \ .$$

Show that, if all this is correct, as in fact it is, we would have proven that any continuous function on the unit interval is a limit of polynomial functions. This is a sketch of a probabilistic proof of an important theorem in mathematics called the *Weierstrass approximation theorem.*

8.2 Law of Large Numbers for Continuous Random Variables

In the previous section we discussed in some detail the Law of Large Numbers for discrete probability distributions. This law has a natural analogue for continuous probability distributions, which we consider somewhat more briefly here.

Chebyshev Inequality

Just as in the discrete case, we begin our discussion with the Chebyshev Inequality.

Theorem 8.3 (Chebyshev Inequality) Let X be a continuous random variable with density function $f(x)$. Suppose X has a finite expected value $\mu = E(X)$ and finite variance $\sigma^2 = V(X)$. Then for any positive number $\epsilon > 0$ we have

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2} \ .$$

□

The proof is completely analogous to the proof in the discrete case, and we omit it.

Note that this theorem says nothing if $\sigma^2 = V(X)$ is infinite.

Example 8.4 Let X be any continuous random variable with $E(X) = \mu$ and $V(X) = \sigma^2$. Then, if $\epsilon = k\sigma = k$ standard deviations for some integer k, then

$$P(|X - \mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2} \ ,$$

just as in the discrete case. □

Law of Large Numbers

With the Chebyshev Inequality we can now state and prove the Law of Large Numbers for the continuous case.

Theorem 8.4 (Law of Large Numbers) Let $X_1, X_2, \ldots, X_n$ be an independent trials process with a continuous density function f, finite expected value μ, and finite variance σ^2. Let $S_n = X_1 + X_2 + \cdots + X_n$ be the sum of the X_i. Then for any real number $\epsilon > 0$ we have

$$\lim_{n \to \infty} P\left(\left| \frac{S_n}{n} - \mu \right| \geq \epsilon \right) = 0 \ ,$$

or equivalently,

$$\lim_{n \to \infty} P\left(\left| \frac{S_n}{n} - \mu \right| < \epsilon \right) = 1 \ .$$

□

Note that this theorem is not necessarily true if σ^2 is infinite (see Example 8.8).

As in the discrete case, the Law of Large Numbers says that the average value of n independent trials tends to the expected value as $n \to \infty$, in the precise sense that, given $\epsilon > 0$, the probability that the average value and the expected value differ by more than ϵ tends to 0 as $n \to \infty$.

Once again, we suppress the proof, as it is identical to the proof in the discrete case.

Uniform Case

Example 8.5 Suppose we choose at random n numbers from the interval $[0, 1]$ with uniform distribution. Then if X_i describes the ith choice, we have

$$\begin{aligned} \mu &= E(X_i) = \int_0^1 x\,dx = \frac{1}{2}\ , \\ \sigma^2 &= V(X_i) = \int_0^1 x^2\,dx - \mu^2 \\ &= \frac{1}{3} - \frac{1}{4} = \frac{1}{12}\ . \end{aligned}$$

Hence,

$$\begin{aligned} E\left(\frac{S_n}{n}\right) &= \frac{1}{2}\ , \\ V\left(\frac{S_n}{n}\right) &= \frac{1}{12n}\ , \end{aligned}$$

and for any $\epsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \epsilon\right) \leq \frac{1}{12n\epsilon^2}\ .$$

This says that if we choose n numbers at random from $[0, 1]$, then the chances are better than $1 - 1/(12n\epsilon^2)$ that the difference $|S_n/n - 1/2|$ is less than ϵ. Note that ϵ plays the role of the amount of error we are willing to tolerate: If we choose $\epsilon = 0.1$, say, then the chances that $|S_n/n - 1/2|$ is less than 0.1 are better than $1 - 100/(12n)$. For $n = 100$, this is about .92, but if $n = 1000$, this is better than .99 and if $n = 10{,}000$, this is better than .999.

We can illustrate what the Law of Large Numbers says for this example graphically. The density for $A_n = S_n/n$ is determined by

$$f_{A_n}(x) = nf_{S_n}(nx)\ .$$

We have seen in Section 7.2, that we can compute the density $f_{S_n}(x)$ for the sum of n uniform random variables. In Figure 8.2 we have used this to plot the density for A_n for various values of n. We have shaded in the area for which A_n would lie between .45 and .55. We see that as we increase n, we obtain more and more of the total area inside the shaded region. The Law of Large Numbers tells us that we can obtain as much of the total area as we please inside the shaded region by choosing n large enough (see also Figure 8.1). □

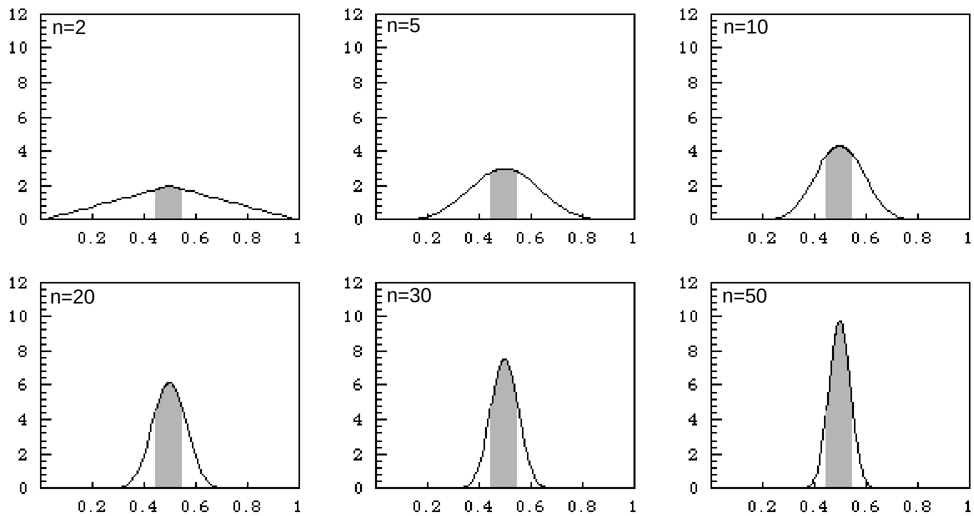


Figure 8.2: Illustration of Law of Large Numbers — uniform case.

Normal Case

Example 8.6 Suppose we choose n real numbers at random, using a normal distribution with mean 0 and variance 1. Then

$$\begin{aligned}\mu &= E(X_i) = 0 \ , \\ \sigma^2 &= V(X_i) = 1 \ .\end{aligned}$$

Hence,

$$\begin{aligned}E\left(\frac{S_n}{n}\right) &= 0 \ , \\ V\left(\frac{S_n}{n}\right) &= \frac{1}{n} \ ,\end{aligned}$$

and, for any $\epsilon > 0$,

$$P\left(\left|\frac{S_n}{n} - 0\right| \geq \epsilon\right) \leq \frac{1}{n\epsilon^2} \ .$$

In this case it is possible to compare the Chebyshev estimate for $P(|S_n/n - \mu| \geq \epsilon)$ in the Law of Large Numbers with exact values, since we know the density function for S_n/n exactly (see Example 7.9). The comparison is shown in Table 8.1, for $\epsilon = .1$. The data in this table was produced by the program **LawContinuous**. We see here that the Chebyshev estimates are in general *not* very accurate. $\square$

n	$P(\|S_n/n\| \geq .1)$	Chebyshev
100	.31731	1.00000
200	.15730	.50000
300	.08326	.33333
400	.04550	.25000
500	.02535	.20000
600	.01431	.16667
700	.00815	.14286
800	.00468	.12500
900	.00270	.11111
1000	.00157	.10000

Table 8.1: Chebyshev estimates.

Monte Carlo Method

Here is a somewhat more interesting example.

Example 8.7 Let $g(x)$ be a continuous function defined for $x \in [0,1]$ with values in $[0,1]$. In Section 2.1, we showed how to estimate the area of the region under the graph of $g(x)$ by the Monte Carlo method, that is, by choosing a large number of random values for x and y with uniform distribution and seeing what fraction of the points $P(x,y)$ fell inside the region under the graph (see Example 2.2).

Here is a better way to estimate the same area (see Figure 8.3). Let us choose a large number of independent values X_n at random from $[0,1]$ with uniform density, set $Y_n = g(X_n)$, and find the average value of the Y_n. Then this average is our estimate for the area. To see this, note that if the density function for X_n is uniform,

$$\begin{aligned}\mu &= E(Y_n) = \int_0^1 g(x)f(x)\,dx \\ &= \int_0^1 g(x)\,dx \\ &= \text{average value of } g(x)\ ,\end{aligned}$$

while the variance is

$$\sigma^2 = E((Y_n-\mu)^2) = \int_0^1 (g(x)-\mu)^2\,dx < 1\ ,$$

since for all x in $[0,1]$, $g(x)$ is in $[0,1]$, hence μ is in $[0,1]$, and so $|g(x)-\mu| \leq 1$. Now let $A_n = (1/n)(Y_1 + Y_2 + \cdots + Y_n)$. Then by Chebyshev's Inequality, we have

$$P(|A_n - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2} < \frac{1}{n\epsilon^2}\ .$$

This says that to get within ϵ of the true value for $\mu = \int_0^1 g(x)\,dx$ with probability at least p, we should choose n so that $1/n\epsilon^2 \leq 1-p$ (i.e., so that $n \geq 1/\epsilon^2(1-p)$). Note that this method tells us how large to take n to get a desired accuracy. $\square$

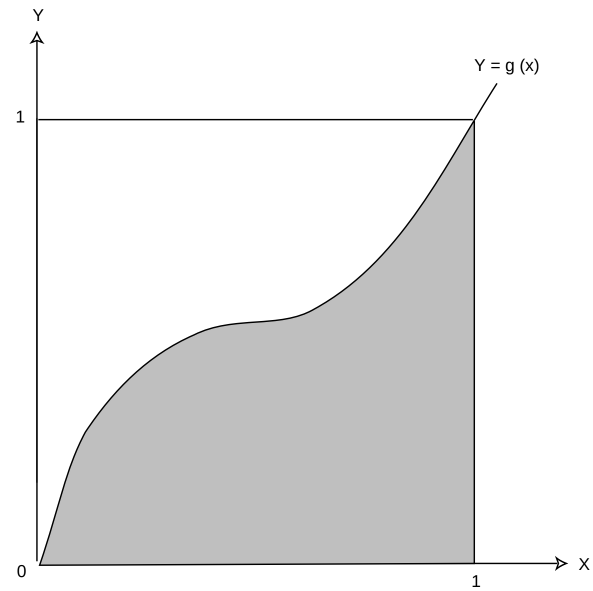


Figure 8.3: Area problem.

The Law of Large Numbers requires that the variance σ^2 of the original underlying density be finite: $\sigma^2 < \infty$. In cases where this fails to hold, the Law of Large Numbers may fail, too. An example follows.

Cauchy Case

Example 8.8 Suppose we choose n numbers from $(-\infty, +\infty)$ with a Cauchy density with parameter $a = 1$. We know that for the Cauchy density the expected value and variance are undefined (see Example 6.28). In this case, the density function for

$$A_n = \frac{S_n}{n}$$

is given by (see Example 7.6)

$$f_{A_n}(x) = \frac{1}{\pi(1 + x^2)} \ ,$$

that is, *the density function for A_n is the same for all n.* In this case, as n increases, the density function does not change at all, and the Law of Large Numbers does not hold. □

Exercises

1 Let X be a continuous random variable with mean $\mu = 10$ and variance $\sigma^2 = 100/3$. Using Chebyshev's Inequality, find an upper bound for the following probabilities.

(a) $P(|X-10| \geq 2)$.

(b) $P(|X-10| \geq 5)$.

(c) $P(|X-10| \geq 9)$.

(d) $P(|X-10| \geq 20)$.

2 Let X be a continuous random variable with values unformly distributed over the interval $[0, 20]$.

(a) Find the mean and variance of X.

(b) Calculate $P(|X-10| \geq 2)$, $P(|X-10| \geq 5)$, $P(|X-10| \geq 9)$, and $P(|X-10| \geq 20)$ exactly. How do your answers compare with those of Exercise 1? How good is Chebyshev's Inequality in this case?

3 Let X be the random variable of Exercise 2.

(a) Calculate the function $f(x) = P(|X-10| \geq x)$.

(b) Now graph the function $f(x)$, and on the same axes, graph the Chebyshev function $g(x) = 100/(3x^2)$. Show that $f(x) \leq g(x)$ for all $x > 0$, but that $g(x)$ is not a very good approximation for $f(x)$.

4 Let X be a continuous random variable with values exponentially distributed over $[0, \infty)$ with parameter $\lambda = 0.1$.

(a) Find the mean and variance of X.

(b) Using Chebyshev's Inequality, find an upper bound for the following probabilities: $P(|X-10| \geq 2)$, $P(|X-10| \geq 5)$, $P(|X-10| \geq 9)$, and $P(|X-10| \geq 20)$.

(c) Calculate these probabilities exactly, and compare with the bounds in (b).

5 Let X be a continuous random variable with values normally distributed over $(-\infty, +\infty)$ with mean $\mu = 0$ and variance $\sigma^2 = 1$.

(a) Using Chebyshev's Inequality, find upper bounds for the following probabilities: $P(|X| \geq 1)$, $P(|X| \geq 2)$, and $P(|X| \geq 3)$.

(b) The area under the normal curve between -1 and 1 is .6827, between -2 and 2 is .9545, and between -3 and 3 it is .9973 (see the table in Appendix A). Compare your bounds in (a) with these exact values. How good is Chebyshev's Inequality in this case?

6 If X is normally distributed, with mean μ and variance σ^2, find an upper bound for the following probabilities, using Chebyshev's Inequality.

(a) $P(|X-\mu| \geq \sigma)$.

(b) $P(|X-\mu| \geq 2\sigma)$.

(c) $P(|X-\mu| \geq 3\sigma)$.

(d) $P(|X - \mu| \geq 4\sigma)$.

Now find the exact value using the program **NormalArea** or the normal table in Appendix A, and compare.

7 If X is a random variable with mean $\mu \neq 0$ and variance σ^2, define the *relative deviation* D of X from its mean by

$$D = \left| \frac{X - \mu}{\mu} \right| .$$

(a) Show that $P(D \geq a) \leq \sigma^2/(\mu^2 a^2)$.

(b) If X is the random variable of Exercise 1, find an upper bound for $P(D \geq .2)$, $P(D \geq .5)$, $P(D \geq .9)$, and $P(D \geq 2)$.

8 Let X be a continuous random variable and define the *standardized version* X^* of X by:

$$X^* = \frac{X - \mu}{\sigma} .$$

(a) Show that $P(|X^*| \geq a) \leq 1/a^2$.

(b) If X is the random variable of Exercise 1, find bounds for $P(|X^*| \geq 2)$, $P(|X^*| \geq 5)$, and $P(|X^*| \geq 9)$.

9 (a) Suppose a number X is chosen at random from $[0, 20]$ with uniform probability. Find a lower bound for the probability that X lies between 8 and 12, using Chebyshev's Inequality.

(b) Now suppose 20 real numbers are chosen independently from $[0, 20]$ with uniform probability. Find a lower bound for the probability that their average lies between 8 and 12.

(c) Now suppose 100 real numbers are chosen independently from $[0, 20]$. Find a lower bound for the probability that their average lies between 8 and 12.

10 A student's score on a particular calculus final is a random variable with values of $[0, 100]$, mean 70, and variance 25.

(a) Find a lower bound for the probability that the student's score will fall between 65 and 75.

(b) If 100 students take the final, find a lower bound for the probability that the class average will fall between 65 and 75.

11 The Pilsdorff beer company runs a fleet of trucks along the 100 mile road from Hangtown to Dry Gulch, and maintains a garage halfway in between. Each of the trucks is apt to break down at a point X miles from Hangtown, where X is a random variable uniformly distributed over $[0, 100]$.

(a) Find a lower bound for the probability $P(|X - 50| \leq 10)$.

(b) Suppose that in one bad week, 20 trucks break down. Find a lower bound for the probability $P(|A_{20} - 50| \leq 10)$, where A_{20} is the average of the distances from Hangtown at the time of breakdown.

12 A share of common stock in the Pilsdorff beer company has a price Y_n on the nth business day of the year. Finn observes that the price change $X_n = Y_{n+1} - Y_n$ appears to be a random variable with mean $\mu = 0$ and variance $\sigma^2 = 1/4$. If $Y_1 = 30$, find a lower bound for the following probabilities, under the assumption that the X_n's are mutually independent.

(a) $P(25 \leq Y_2 \leq 35)$.

(b) $P(25 \leq Y_{11} \leq 35)$.

(c) $P(25 \leq Y_{101} \leq 35)$.

13 Suppose one hundred numbers X_1, X_2, ..., X_{100} are chosen independently at random from $[0, 20]$. Let $S = X_1 + X_2 + \cdots + X_{100}$ be the sum, $A = S/100$ the average, and $S^* = (S - 1000)/(10/\sqrt{3})$ the standardized sum. Find lower bounds for the probabilities

(a) $P(|S - 1000| \leq 100)$.

(b) $P(|A - 10| \leq 1)$.

(c) $P(|S^*| \leq \sqrt{3})$.

14 Let X be a continuous random variable normally distributed on $(-\infty, +\infty)$ with mean 0 and variance 1. Using the normal table provided in Appendix A, or the program **NormalArea**, find values for the function $f(x) = P(|X| \geq x)$ as x increases from 0 to 4.0 in steps of .25. Note that for $x \geq 0$ the table gives $NA(0, x) = P(0 \leq X \leq x)$ and thus $P(|X| \geq x) = 2(.5 - NA(0, x)$. Plot by hand the graph of $f(x)$ using these values, and the graph of the Chebyshev function $g(x) = 1/x^2$, and compare (see Exercise 3).

15 Repeat Exercise 14, but this time with mean 10 and variance 3. Note that the table in Appendix A presents values for a standard normal variable. Find the standardized version X^* for X, find values for $f^*(x) = P(|X^*| \geq x)$ as in Exercise 14, and then rescale these values for $f(x) = P(|X - 10| \geq x)$. Graph and compare this function with the Chebyshev function $g(x) = 3/x^2$.

16 Let $Z = X/Y$ where X and Y have normal densities with mean 0 and standard deviation 1. Then it can be shown that Z has a Cauchy density.

(a) Write a program to illustrate this result by plotting a bar graph of 1000 samples obtained by forming the ratio of two standard normal outcomes. Compare your bar graph with the graph of the Cauchy density. Depending upon which computer language you use, you may or may not need to tell the computer how to simulate a normal random variable. A method for doing this was described in Section 5.2.

(b) We have seen that the Law of Large Numbers does not apply to the Cauchy density (see Example 8.8). Simulate a large number of experiments with Cauchy density and compute the average of your results. Do these averages seem to be approaching a limit? If so can you explain why this might be?

17 Show that, if $X \geq 0$, then $P(X \geq a) \leq E(X)/a$.

18 (Lamperti[9]) Let X be a non-negative random variable. What is the best upper bound you can give for $P(X \geq a)$ if you know

(a) $E(X) = 20$.

(b) $E(X) = 20$ and $V(X) = 25$.

(c) $E(X) = 20$, $V(X) = 25$, and X is symmetric about its mean.

[9] Private communication.

KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)

(Established under Section 3 of UGC Act, 1956)

Pollachi Main Road, Coimbatore – 641 021, Tamil Nadu

Department of Mathematics

Subject : Probability and Statistics	**Semester VI**	**L T P C**
Subject Code : 16MMU603A	**Class : III B.Sc Mathematics**	**4 2 0 4**

UNIT IV

Continuous distributions: uniform distribution, normal distribution, standard normal distribution, exponential distribution. Joint cumulative distribution function and its properties, joint probability density functions (No derivations) and simple problems. Bivariate distribution, correlation coefficient, joint moment generating function (jmgf) and calculation of covariance (from jmgf), linear regression for two variables.

SUGGESTED READINGS

TEXT BOOK

1. Gupta S.P., (2001). Statistical Methods, Sultan Chand & Sons, New Delhi.

REFERENCES

1. Robert V. Hogg, Joseph W. McKean and Allen T. Craig., (2007). Introduction to Mathematical Statistics, Pearson Education, Asia.
2. Irwin Miller and Marylees Miller, John E. Freund, (2006). Mathematical Statistics with Application, Seventh Edition, Pearson Education, Asia.
3. Sheldon Ross., (2007). Introduction to Probability Model, Ninth Edition, Academic Press, Indian Reprint.
4. Pillai R.S.N., and Bagavathi V., (2002). Statistics , S. Chand & Company Ltd, New Delhi.
5. Srivastava T.N., and Shailaja Rego., (2012). 2e, Statistics for Management, Mc Graw Hill Education, New Delhi.
6. Dr.P.N.Arora, (1997). A foundation course statistics, S.Chand & Company Ltd, New Delhi.

Question	Opt 1	Opt 2	Opt 3	Opt 4	Answer
Study to portray accurately characteristics of a particular individual, situation or a group is called ------------ research	Exploratory	Diagnostic	Descriptive	Hypothesis testing	Descriptive
Critical evaluation made by the researcher with the facts and information already available is called ----------- research.	Analytical	Exploratory	Diagnostic	Hypothesis testing	Analytical
Research to find reason, why people think or do certain things is an example of -----------	Quantitative Research	Applied Research	Qualitative research	Fundamental research	Qualitative research
Which one is considered a major component of the research study	Interpretation	research report	finding	draft	research report
Research task remains incomplete till the ___________ has been presented.	Report	objective	finding	objective and finding	Report
What is the last step in a research study	Writing report	writing finding	writing limitations	research report	Writing report
Which is the final step in report writing………..	Writing report	writing finding	writing drafts	writing limitations	writing drafts
What is usually appended to the research work	Editing	bibliography	coding	research report	bibliography
The ___________ is one which gives emphasis on simplicity and attractiveness	popular report	research report	article report	writing limitations	popular report
___________ should slow originality and should necessarily be on attempt to solve some intellectual problem	Interpretation	research report	finding	draft	research report
The researcher must remain caution about the _________ that can possibly arise in the process of interpreting results	Analysis	conclusions	findings	error	error
Which one should be considered while interpreting a given data	Validity	reliability	both validity and reliability	technical jargon	reliability

_________ is asking questions face to face	Indirect method	mailed questionnaire	through post	personal interview	personal interview
Journals, books, magazines etc are useful sources of collecting----------	Primary data	secondary data	both primary and secondary data	objective	secondary data
The collected raw data to detect errors and are called,	Editing	coding	classification	all the above	Editing
The formal, systematic and intensive process of carrying on a scientific method of analysis is _______________	Research Design	research	interpretation	research analysis	research
------ Refers to the process of assigning numerals or symbols to answers of response	Coding	editing	classification	all the above	Coding
The research study, which is based on describing the characteristic of a particular individual or group	Experience survey	Descriptive	Diagnostic	Exploratory	Descriptive
Research is a______________	Finding	assumption	statement	all the above	all the above
The research, which has the purpose of improving a product or a process testing theoretical concepts in actual problem situations is-------research.	Statistical	Applied	Domestic	Biological	Applied
The chart of research process indicates that the process consists of a number of ----.	Closely related activities	unrelated activities	Closely unrelated activities	moderately related activities	Closely related activities

The objective of a good design is ------------- ---.	Maximize the bias andmaximize the reliability of data	Minimize the bias and minimize the reliability of data	Minimize the bias and maximize the reliability of data	Maximize the bias and maximize the reliability of data	Maximize the bias and maximize the reliability of data
A ---------- is used whenever a full written report of the study is required.	Popular report	Technical report	article	monograph	Technical report
The ----------- is one which gives emphasis on simplicity and attractiveness.	Popular report	Technical report	article	monograph	Popular report
Study to portray accurately characteristics of a particular individual, situation or a group is called ------------ research	Exploratory	Diagnostic	Descriptive	Hypothesis testing	Descriptive
Critical evaluation made by the researcher with the facts and information already available is called ----------- research.	Analytical	Exploratory	Diagnostic	Hypothesis testing	Analytical
Research to find reason, why people think or do certain things is an example of -----------	Quantitative Research	Applied Research	Qualitative research	Fundamental research	Qualitative research
Which one is considered a major component of the research study	Interpretation	research report	finding	draft	research report
Research task remains incomplete till the ___________ has been presented.	Report	objective	finding	objective and finding	Report
What is the last step in a research study	Writing report	writing finding	limitations	research report	Writing report
Which is the final step in report writing………..	Writing report	writing finding	writing drafts	writing limitations	writing drafts
What is usually appended to the research work	Editing	bibliography	coding	research report	bibliography
The ___________ is one which gives emphasis on simplicity and attractiveness	popular report	research report	article report	writing limitations	popular report

___________ should slow originality and should necessarily be on attempt to solve some intellectual problem	Interpretation	research report	finding	draft	research report
The researcher must remain caution about the _________ that can possibly arise in the process of interpreting results	Analysis	conclusions	findings	error	error
Which one should be considered while interpreting a given data	Validity	reliability	both validity and reliability	technical jargon	reliability
__________ is asking questions face to face	Indirect method	mailed questionnaire	through post	personal interview	personal interview
Journals, books, magazines etc are useful sources of collecting----------	Primary data	secondary data	both primary and secondary data	objective	secondary data
The collected raw data to detect errors and are called,	Editing	coding	classification	all the above	Editing
The formal, systematic and intensive process of carrying on a scientific method of analysis is _______________	Research Design	research	interpretation	research analysis	research
------ Refers to the process of assigning numerals or symbols to answers of response	Coding	editing	classification	all the above	Coding
The research study, which is based on describing the characteristic of a particular individual or group	Experience survey	Descriptive	Diagnostic	Exploratory	Descriptive
Research is a______________	Finding	assumption	statement	all the above	all the above
The research, which has the purpose of improving a product or a process testing theoretical concepts in actual problem situations is-------research.	Statistical	Applied	Domestic	Biological	Applied

The chart of research process indicates that the process consists of a number of ----.	Closely related activities	unrelated activities	Closely unrelated activities	moderately related activities	Closely related activities
The objective of a good design is -------------.	Maximize the bias andmaximize the reliability of data	Minimize the bias and minimize the reliability of data	Minimize the bias and maximize the reliability of data	Maximize the bias and maximize the reliability of data	Maximize the bias and maximize the reliability of data
A ---------- is used whenever a full written report of the study is required.	Popular report	Technical report	article	monograph	Technical report
The ----------- is one which gives emphasis on simplicity and attractiveness.	Popular report	Technical report	article	monograph	Popular report
The square of the S.D is ……………….	Variance	Coefficient of variation	Square of variance	Square of coefficient of variation	Variance
Analysis of variance is a statistical method of comparing the ________ of several populations.	Standard deviations	Means	Variances	Proportions	Means
The analysis of variance is a statistical test that is used to compare how many group means?	Three	More than three	Three or more	Two or more	Two or more
Analysis of variance utilizes:	F-test	Chi-Square test	Z-test	t-test	F-test
What is two-way ANOVA?	An ANOVA with two variables and one factor	An ANOVA with one variable and two factors	An ANOVA with one variable and three factors	An ANOVA with both categorical and scale variables	An ANOVA with one variable and two factors
Which of the following is the correct F ratio in the one-way ANOVA?	MSA/MSE	MSBL/MSE	MST/MSE	MSE/MST	MST/MSE
For validity of F-test in Anova, parent population should be……….	Binomial	Poisson	Normal	Exponential	Normal

______ sum of squares measures the variability of the observed values around their respective tabulated values	Treatment	Error	Interaction	Total	Error
The ________ sum of squares measures the variability of the sample treatment means around the overall mean.	Total	Treatment	Error	Interaction	Treatment
If the true means of the k populations are equal, then MST/MSE should be:	more than 1.00	Close to 1.00	Close to -1.00	A negative value between 0 and - 1	Close to 1.00
If MSE of ANOVA for six treatment groups is known, you can compute…………	Degree of freedom	The standard deviation of each treatment group	Variance	The pooled standard deviation	The pooled standard deviation
To determine whether the test statistic of ANOVA is statistically significant,to determine critical value we need	Sample size, number of groups	Mean, sample standard deviation	Expected frequency, obtained frequency	MSTR, MSE	Sample size, number of groups
Which of the following is an assumption of one-way ANOVA comparing samples from 3 or more experimental treatments?	Variables follow F-distribution	Variables follow normal distribution	Samples are dependent each other	Variables have different variances	Variables follow normal distribution
The error deviations within the SSE statistic measure distances:	Within groups	Between groups	Between each value and the grand mean	Betweeen samples	Within groups
In one-way ANOVA, which of the following is used within the F-ratio as a measurement of the variance of individual observations?	SSTR	MSTR	SSE	MSE	SSE
When conducting a one-way ANOVA, the _______ the between-treatment variability is when compared to the within-treatment variability	More random larger	Smalller	Larger	More random smaller	Smaller

When conducting a one-way ANOVA,the value of *F* DATA will be tend to be ……..	More random larger	Smalller	More random smaller	Larger	Smaller
When conducting an ANOVA, *F* DATA will always fall within what range?	Between negative infinity and infinity	Between 0 and 1	Between 0 and infinity	Between 1 and infinity	Between 0 and infinity
If F DATA = 5, the result is statistically significant	Always	Sometimes	Never	Is impossible	Sometimes
If F DATA= 0.9, the result is statistically significant	Always	Sometimes	Never	Is impossible	Never
When comparing three treatments in a one-way ANOVA ,the alternate hypothesis is…..	All three treatments have different effect on the mean response.	Exactly two of the three treatments have the same effect on the mean response.	At least two treatments are different from each other in terms of their effect on the mean response	All the treatments have same effect	At least two treatments are different from each other in terms of their effect on the mean response
If the sample means for each of *k* treatment groups were identical,the observed value of the ANOVA test statistic?	1	0	A value between 0.0 and 1.0	A negative value	0
If the null hypothesis is rejected, the probability of obtaining a *F* - ratio > the value in the *F* table as the 95th % is:	0.5	>0.5	<0.5	1	<0.5
ANOVA was used to test the outcomes of three drug treatments. Each drug was given to 20 individuals. If MSE =16, What is the standard deviation for all 60 individuals sampled for this study?	6.928	48	16	4	4
Analysis of variance technique originated in the field of………………	Agriculture	Industry	Biology	Genetics	Agriculture

With 90, 35, 25 as TSS, SSR and SSC , in case of two way classification, SSE is…..	50	40	30	20	30
Variation between classes or variation due to different basis of classification is commonly known as ……………...	Treatments	Total sum of squares	Sum of squares	Sum of squares due to error	Treatments
The total variation in observations in Anova is classified as:	Treatments and inherent variation	SSE and SST	MSE and MST	TSS and SSE	Treatments and inherent variation
In Anova, variance ratio is given by……………	MST/MSE	MSE/MST	SSE/SST	TSS/SSE	MST/MSE
Degree of freedom for TSS is……………..	N-1	k-1	h-1	(k-1)(h-1)	N-1
For Anova, MST stands for……………………...	Mean sum of squares of treatment	Mean sum of squares of varieties	Mean sum of squares of tables	Mean sum of sources of treatment	Mean sum of squares of treatment
An ANOVA procedure is applied to data of 4 samples,where each sample contains 10 observations.Then degree of freedom for critical value of F are	4 numerator and 9 denominator	3 numerator and 40 denominator	3 numerator and 36 denominator	4 numerator and 10 denominator	3 numerator and 36 denominator
The power function of a test is denoted by…………………….	M(w,Q)	M(Q,Qo)	P(w,Q)	P(w,Qo)	M(w,Q)
Sum of power function and operation characteristic is…………	Unity	Zero	two	Negative	Unity
Operation characteristic is denoted by…………………	L(w,Q)	M(w,Q)	L(w,Qo)	M(w,Qo)	L(w,Q)
Operation characteristic is also known as …………………	Test characteristic	Power function	best characteristic	unique characteristic	Test characteristic
The formula to find OC is L(w,Q)=	1-Power Function	2xPower Function	Power Funtion -1	2xConfidance Interval	1-Power Function
Operation Characteristic is of a test is related to…	Power Function	Best Test	Unique Test	Uniformally Best Test	Power Function

If the Hypothesis is correct the operation charectristics will be…………	1	0	-1	0.5	1
If the Hypothesis is wrong the operation charectristics will be……….	0	1	0.5	0.333333	0
In which test we verify a null hypothesis against any other definite alternate hypothesis?	Best Test	Unique Test	Uniformally Best Test	Unbiased Test	Best Test
A Best Test is a Test such that the critical region for which ______ attains least value for a given α.	Beta	1-Beta	Alpha	1-Alpha	1-Beta
A Test whose power funtion attains its mean at point Q = Qo is called ____ Test	Unique	Unbiased	Power	Operation Characteristic	Unique
A Best Unique Test exist ____________	Always	Never	Sometimes	When Q not = to Qo	Sometimes
Operation Characteristic is related to	Power Function	Unique Test	Best Test	Uniformally Best Test	Power Function
Power is the ability to detect:	A statistically significant effect where one exists	A psychologically important effect where one exists	Both (a) and (b) above	Design flaws	A statistically significant effect where one exists
Calculating how much of the total variance is due to error and the experimental manipulation is called:	Calculating the variance	Partitioning the variance	Producing the variance	Summarizing the variance	Partitioning the variance
ANOVA is useful for:	Teasing out the individual effects of factors on an Independent Variables	Analyzing data from research with more than one Independent Variable and one Dependent Variable	Analyzing correlational data	Individual effects of factors on an Dependent Variables	Analyzing data from research with more than one Independent Variable and one Dependent Variable

What is the definition of a simple effect?	The effect of one variable on another	The difference between two conditions of one Independent Variable at one level of another Independent Variable	The easiest way to get a significant result	Difference between two Dependent Variables	The difference between two conditions of one Independent Variable at one level of another Independent Variable
In a study with gender as the manipulated variable, the Independent Variable is:	Within participants	Correlational	Between participants	Regressional	Between participants
Which of the following statements are true of experiments?	The Independent Variable is manipulated by the experimenter	The Dependent Variable is assumed to be dependent upon the IV	They are difficult to conduct	both (a) and (b)	both (a) and (b)
All other things being equal, repeated-measures designs:	Have exactly the same power as independent designs	Are often less powerful than independent designs	Are often more powerful than independent designs	Are rarely less powerful when compare to than independent designs	Are often more powerful than independent designs
Professor P. Nutt is examining the differences between the scores of three groups of participants. If the groups show homogeneity of variance, this means that the variances for the groups:	Are similar	Are dissimilar	Are exactly the same	Are enormously different	Are similar
Differences between groups, which result from our experimental manipulation, are called:	Individual differences	Treatment effects	Experiment error	Within-participants effects	Treatment effects

Herr Hazelnuss is thinking about whether he should use a related or unrelated design for one of his studies. As usual, there are advantages and disadvantages to both. He has four conditions. If, in a related design, he uses 10 participants, how many would he need for an unrelated design?	40	20	10	100	40
Individual differences within each group of participants are called:	Treatment effects	Between-participants error	Within-participants error	Individual biases	Within-participants error
Calculating how much of the total variance is due to error and the experimental manipulation is called:	Calculating the variance	Partitioning the variance	Producing the variance	Summarizing the variance	Partitioning the variance
The decision on how many factors to keep is decided on:	Statistical criteria	Theoretical criteria	Both (a) and (b)	Neither (a) nor (b)	Both (a) and (b)
It is possible to extract:	As many factors as variables	More factors than variables	More variables than factors	Correlation between the actual and predicted variables	As many factors as variables
Four groups have the following means on the covariate: 35, 42, 28, 65. What is the grand mean?	43.5	42.5	56.7	58.9	42.5
You can perform ANCOVA on:	Two groups	Three groups	Four groups	All of the above	All of the above
When carrying out a pretest--posttest study, researchers often wish to:	Partial out the effect of the dependent variable	Partial out the effect of the pretest	Reduce the correlation between the pretest and posttest scores	Correlation between the two tests scores	Partial out the effect of the pretest

Using difference scores in a pretest--posttest design does not partial out the effect of the pretest for the following reason:	The pretest scores are not normally correlated with the posttest scores	The pretest scores are normally correlated with the different scores	The posttest scores are normally correlated with the different scores	Up normal relationship with the different scores	The pretest scores are normally correlated with the different scores
Experimental designs are characterized by:	Two conditions	No control condition	Random allocation of participants to conditions	More than two conditions	Random allocation of participants to conditions
Between-participants designs can be:	Either quasi-experimental or experimental	Only experimental	Only quasi-experimental	Only correlational	Either quasi-experimental or experimental
A continuous variable can be described as:	Able to take only certain discrete values within a range of scores	Able to take any value within a range of scores	Being made up of categories	Being made up of variables	Able to take any value within a range of scores
In a within-participants design with two conditions, if you do not use counterbalancing of the conditions then your study is likely to suffer from:	Order effects	Effects of time of day	Lack of participants	Effects of participants	Order effects
Demand effects are possible confounding variables where:	Participants behave in the way they think the experimenter wants them to behave	Participants perform poorly because they are tired or bored	Participants perform well because they have practiced the experimental task	Participants perform strongly	Participants behave in the way they think the experimenter wants them to behave

Power can be calculated by a knowledge of:	The statistical test, the type of design and the effect size	The statistical test, the criterion significance level and the effect size	The criterion significance level, the effect size and the type of design	The criterion significance level, the effect size and the sample size	The criterion significance level, the effect size and the sample size
Relative to large effect sizes, small effect sizes are:	Easier to detect	Harder to detect	As easy to detect	As difficult to detect	As difficult to detect
Differences between groups, which result from our experimental manipulation, are called:	Individual differences	Treatment effects	Experiment error	Within-participants effects	Treatment effects
Completely randomized design is similar to ----------	three way	one way	two way	t test	one way
Randomized block design is similar to ----------	two way	three way	one way	many	two way
ANOVA is the technique of analysis of ------	standard deviation	variance	mean	range	variance
Under one way classification , the influence of only ----- attribute or factor is considered	two	three	one	many	one
Under two way classification , the influence of only ----- attribute or factor is considered	four	two	three	one	two

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established under Section 3 of UGC Act, 1956)
Pollachi Main Road, Coimbatore – 641 021, Tamil Nadu

Department of Mathematics

Subject : Probability and Statistics	**Semester VI**	**L T P C**
Subject Code : 16MMU603A	**Class : III B.Sc Mathematics**	**4 2 0 4**

UNIT IV

Continuous distributions: uniform distribution, normal distribution, standard normal distribution, exponential distribution. Joint cumulative distribution function and its properties, joint probability density functions (No derivations) and simple problems. Bivariate distribution, correlation coefficient, joint moment generating function (jmgf) and calculation of covariance (from jmgf), linear regression for two variables.

SUGGESTED READINGS

TEXT BOOK

1. Gupta S.P., (2001). Statistical Methods, Sultan Chand & Sons, New Delhi.

REFERENCES

1. Robert V. Hogg, Joseph W. McKean and Allen T. Craig., (2007). Introduction to Mathematical
Statistics, Pearson Education, Asia.
2. Irwin Miller and Marylees Miller, John E. Freund, (2006). Mathematical Statistics with Application, Seventh Edition, Pearson Education, Asia.
3. Sheldon Ross., (2007). Introduction to Probability Model, Ninth Edition, Academic Press, Indian Reprint.
4. Pillai R.S.N., and Bagavathi V., (2002). Statistics , S. Chand & Company Ltd, New Delhi.
5. Srivastava T.N., and Shailaja Rego., (2012). 2e, Statistics for Management, Mc Graw Hill Education, New Delhi.
6. Dr.P.N.Arora, (1997). A foundation course statistics, S.Chand & Company Ltd, New Delhi.

Question	Opt 1	Opt 2	Opt 3	Opt 4	Answer
Psychometric Methods book is written by________________	J.P.Guilford	Likert	L.L.Thurstone	Louis Guttman	J.P.Guilford
Respondents are asked to rank their choices in________________	Comparative scaling	arbitrary scaling	rating scale	differential scale	Comparative scaling
________ is developed on ad-hoc basis	Differential scale	arbitrary scale	rating scale	ranking scale	arbitrary scale
_______________scale is developed by utilizing item analysis approach	Comparative scale	likert scale	differential scale	rating scale	likert scale
Scalogram analysis is developed by______________	J.P.Guilford	Likert	L.L.Thurstone	Louis Guttman	Louis Guttman
A complete enumeration of all items in the population is known as _____________	sampling unit	sample design	census inquiry	all the above	census inquiry
The selected respondents constitute______________	population	sample	sample size	population size	sample
The selection process of respondents is called_____________	survey	sampling technique	sample survey	census inquiry	sampling technique
The survey conducted to select the respondents is called______________	sampling technique	sample survey	census inquiry	population size	sample survey
A sample design is a definite plan for obtaining a sample from a given_____________	universe	sample design	population	sample survey	population
The number of items in universe can be_________________	finite	infinite	both	zero	both
The population of a city, number of workers in a company is_______________	infinite	finite	both	zero	finite
Source list is also known as___________	sampling size	sampling size	sampling frame	population size	sampling frame
The size of the sample should be _____________	large	optimum	small	all the above	optimum

Inappropriateness in sampling frame will result in ______________________	systematic bias	optimum	problems	sampling error	systematic bias
Sampling error ___________ with increase in size of sample	decrease	increase	both	optimum	decrease
Sampling error can be measured from______________	sample design	sample size	population	sample design and sample size	sample design and sample size
On the representation basis samples may be___________________	probability sampling	non-probability sampling	both	restricted	both
On element selection basis the samples may be______________	restricted	unrestricted	both	probability sampling	both
Non-probability sampling is also known as__________________	quota sampling	purposive sampling	deliberate sampling	all the three	all the three
Quota sampling is an example of___________________	probability sampling	non-probability sampling	both	purposive sampling	non-probability sampling
Probability sampling is also known as___________________	random sampling	choice sampling	random and choice sampling	multistage sampling	random and choice sampling
Lottery method of selecting data is an example of________________	random sampling	choice sampling	purposive sampling	quota sampling	random sampling
Systematic sampling is an improved version of_____________________	quota sampling	simple random sampling	choice sampling	purposive sampling	simple random sampling
If population is not drawn from homogeneous group________________ technique is applied	simple random sampling	quota sampling	choice sampling	stratified sampling	stratified sampling
In_____________________ total population is divided into number of relatively small sub divisions	cluster sampling	choice sampling	stratified sampling	quota sampling	cluster sampling
When a particular lot is to be accepted or rejected on the basis of single sampling it is known as____________________	double sampling	single sampling	area sampling	purposive sampling	single sampling

Survey designed to determine attitude of students toward new teaching plan is known as_____________________________-	cross stratification sampling	stratification sampling	cluster sampling	multi stage sampling	cross stratification sampling
Sample design is determined_______________ datas are collected	before	after	both	based on the survey	before
Indeterminary principle step comes in_________________	step in sample design	criteria to select sample procedure	both	step doesnot occur	criteria to select sample procedure
The measurement of sampling error is called as_____________________	precision of sampling plan	sampling survey	sampling plan	representation basis	precision of sampling plan
The different sub populations divided to constitute a sample is known as__________________	stratified sampling	survey	population	strata	strata
Every nth item is selected in____________________	stratified sampling	systematic sampling	judgement sampling	all the above	systematic sampling
_________________is conducted for determining a more appropriate and efficient stratification plan	survey	sample	pilot study	sample plan	pilot study
______________________is considered more appropriate when universe happens to be small	purposive sampling	area sampling	cluster sampling	simple random sampling	purposive sampling
When we use rating scales we judge an object in ________terms against some specified criteria.	real	absolute	imaginary	perfect	absolute
Rating scale is also known as_______________	Categorical scale	arbitrary scale	cumulative scales	all the above	Categorical scale
The graphical scale is ________and is commonly used in practice.	Problematic	critical	simple	real	simple
________________is also known as numerical scale	Itemized rating scale	graphical rating scale	cumulative scale	likert scale	Itemized rating scale

The chief merit of itemized rating scale is it provides___________ information	more	deep	critical	all the above	more
________________occurs when the respondents are either easy raters or hard raters	error of hallo effect	error of leniency	error of central tendency	cumulative scales	error of leniency
_______________ occurs when the rater carries a generalized impression of the subject from one rating to another.	error of hallo effect	error of leniency	error of central tendency	graphical rating scale	error of hallo effect
When the raters are reluctant to give extreme judgments, the result is____________	error of hallo effect	error of leniency	error of central tendency	cluster sampling	error of central tendency
Systematic bias is also known as___________________	Error of hallo effect	error of leniency	error of central tendency	cumulative scales	Error of hallo effect
_____________occurs when the rater is asked to rate more factors, which has no evidence for judgment.	error of hallo effect	error of leniency	error of central tendency	cluster sampling	error of hallo effect
_______________ is also known as ranking scale	rating scale	comparative scale	likert scale	graphical rating scale	comparative scale
We make relative judgments against similar objects in _______________	comparative scale	likert scale	differential scale	rating scale	comparative scale
Paired comparisions provide_______________ data.	nominal	ordinal	ratios	interval	ordinal
Ordinal data can be converted to ____________ data through Law of comparative judgment.	nominal	ordinal	ratio	interval	interval
Law of comparative judgment is developed by_______________	J.P.Guilford	Likert	L.L.Thurstone	all the three	L.L.Thurstone
--------- Scales have an absolute or true zero of measurement	Ordinal	Nominal	interval	ratio	ratio

The section of ---------- constitutes the main body of the report where in the results of the study are presented in clear.	Appendix	results	methods	Ordinal	results
Study to portray accurately characteristics of a particular individual, situation or a group is called ------------ research	Exploratory	Diagnostic	Descriptive	Hypothesis testing	Descriptive
Critical evaluation made by the researcher with the facts and information already available is called ----------- research.	Analytical	Exploratory	Diagnostic	Hypothesis testing	Analytical
Research to find reason, why people think or do certain things is an example of -----------	Quantitative Research	Applied Research	Qualitative research	Fundamental research	Qualitative research
Which one is considered a major component of the research study	Interpretation	research report	finding	draft	research report
Research task remains incomplete till the ___________ has been presented.	Report	objective	finding	objective and finding	Report
What is the last step in a research study	Writing report	writing finding	writing limitations	research report	Writing report
Which is the final step in report writing………..	Writing report	writing finding	writing drafts	writing limitations	writing drafts
What is usually appended to the research work	Editing	bibliography	coding	research report	bibliography
The ___________ is one which gives emphasis on simplicity and attractiveness	popular report	research report	article report	writing limitations	popular report
___________ should slow originality and should necessarily be on attempt to solve some intellectual problem	Interpretation	research report	finding	draft	research report

The researcher must remain caution about the ________ that can possibly arise in the process of interpreting results	Analysis	conclusions	findings	error	error
Which one should be considered while interpreting a given data	Validity	reliability	both validity and reliability	technical jargon	reliability
_________ is asking questions face to face	Indirect method	mailed questionnaire	through post	personal interview	personal interview
Journals, books, magazines etc are useful sources of collecting----------	Primary data	secondary data	both primary and secondary data	objective	secondary data
The collected raw data to detect errors and are called,	Editing	coding	classification	all the above	Editing
The formal, systematic and intensive process of carrying on a scientific method of analysis is _______________	Research Design	research	interpretation	research analysis	research
------ Refers to the process of assigning numerals or symbols to answers of response	Coding	editing	classification	all the above	Coding
The research study, which is based on describing the characteristic of a particular individual or group	Experience survey	Descriptive	Diagnostic	Exploratory	Descriptive
Research is a______________	Finding	assumption	statement	all the above	all the above
The research, which has the purpose of improving a product or a process testing theoretical concepts in actual problem situations is-------research.	Statistical	Applied	Domestic	Biological	Applied

The chart of research process indicates that the process consists of a number of ---.	Closely related activities	unrelated activities	Closely unrelated activities	moderately related activities	Closely related activities
The objective of a good design is -------------.	Maximize the bias andmaximize the reliability of data	Minimize the bias and minimize the reliability of data	Minimize the bias and maximize the reliability of data	Maximize the bias and maximize the reliability of data	Maximize the bias and maximize the reliability of data
A ---------- is used whenever a full written report of the study is required.	Popular report	Technical report	article	monograph	Technical report
The ----------- is one which gives emphasis on simplicity and attractiveness.	Popular report	Technical report	article	monograph	Popular report
Which of the following are measurements of scale?	Nominal	ordinal	interval	all the above	all the above
____Scale is a system of assigning numbers, symbols to events in order to label them.	Interval	ordinal	Nominal	ratio	Nominal
The qualitative phenomena are considered in the -------------- scale.	Ordinal	Nominal	interval	ratio	Ordinal
---------- Scales can have an arbitrary zero, but it is not possible to determine the absolute zero.	Ordinal	Nominal	interval	ratio	interval

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University)
(Established under Section 3 of UGC Act, 1956)
Pollachi Main Road, Coimbatore – 641 021, Tamil Nadu

Department of Mathematics

Subject : Mathematical Statistics **Semester VI** **L T P C**
Subject Code : 16MMU603A **Class : III B.Sc Mathematics** **4 2 0 4**

Glossary of Statistical Terms

2 X 5 factorial design	A factorial design with one variable having two levels and the other having five levels.
Alpha	The probability of a Type I error.
Abscissa	Horizontal axis.
Additive law of probability	The rule giving the probability of the occurrence of one or more mutually exclusive events.
Adjacent values	Actual data points that are no more extreme than the inner fences.
Alternative hypothesis (H_1)	The hypothesis that is adopted when H_0 is rejected. Usually the same as the research hypothesis.
ß (Beta)	The probability of a Type II error.
Categorical data	Data representing counts or number of observations in each category.
Cell	The combination of a particular row and column ‹the set of observations obtained under identical treatment conditions.
Central limit theorem	The theorem that specifies the nature of the sampling distribution of the mean.
Chi-square test	A statistical test often used for analyzing categorical data.
Conditional probability	The probability of one event *given* the occurrence of some other event.
Confidence interval	An interval, with limits at either end, with a specified probability of including the parameter being estimated.
Confidence limits	An interval, with limits at either end, with a specified probability of including the parameter being estimated.
Constant	A number that does not change in value in a given situation.
Continuous variables	Variables that take on *any* value.
Correlation	Relationship between variables.
Correlation coefficient	A measure of the relationship between variables.
Count data	Data representing counts or number of observations in each category.
Covariance	A statistic representing the degree to which two variables vary together.
Criterion variable	The variable to be predicted.
Critical value	The value of a test statistic at or beyond which we will reject H0 .
Decision making	A procedure for making logical decisions on the basis of sample data.
Degrees of	The number of independent pieces of information remaining after

freedom (*df*)	estimating one or more parameters.
Density	Height of the curve for a given value of X- closely related to the probability of an observation in an interval around X.
Dependent variables	The variable being measured. The data or score.
Discrete variables	Variables that take on a small set of possible values.
Dispersion	The degree to which individual data points are distributed around the mean.
Distribution free tests	Statistical tests that do not rely on parameter estimation or precise distributional assumptions.
Effect size	The difference between two population means divided by the standard deviation of either population.
Efficiency	The degree to which repeated values for a statistic cluster around the parameter.
Event	The outcome of a trial.
Exhaustive	A set of events that represents all possible outcomes.
Expected value	The average value calculated for a statistic over an infinite number of samples.
Expected frequencies	The expected value for the number of observations in a cell if H_0 is true.
Experimental hypothesis	Another name for the research hypothesis.
Exploratory data analysis (EDA)	A set of techniques developed by Tukey for presenting data in visually meaningful ways.
External Validity	The ability to generalize the results from this experiment to a larger population.
Frequency distribution	A distribution in which the values of the dependent variable are tabled or plotted against their frequency of occurrence.
Frequency data	Data representing counts or number of observations in each category.
Goodness of fit test	A test for comparing observed frequencies with theoretically predicted frequencies.
Grand total ($\sum X$)	The sum of all of the observations.
Hypothesis testing	A process by which decisions are made concerning the values of parameters.
Independent variables	Those variables controlled by the experimenter.
Independent events	Events are independent when the occurrence of one has no effect on the probability of the occurrence of the other.
Interaction	A situation in a factorial design in which the effects of one independent variable depend upon the level of another independent variable.
Intercept	The value of Y when X is 0.
Interval scale	Scale on which equal intervals between objects represent equal differences ‹differences are meaningful.
Interval estimate	A range of values estimated to include the parameter.
Joint probability	The probability of the co-occurrence of two or more events.
Kurtosis	A measure of the peakedness of a distribution.
Leptokurtic	A distribution that has relatively more scores in the center and in the tails.

Linear relationship	A situation in which the best-fitting regression line is a straight line.
Linear regression	Regression in which the relationship is linear.
Marginal totals	Totals for the levels of one variable summed across the levels of the other variable.
Matched samples	An experimental design in which the same subject is observed under more than one treatment.
Mean absolute deviation (m.a.d.)	Mean of the absolute deviations about the mean.
Mean	The sum of the scores divided by the number of scores.
Measurement	The assignment of numbers to objects.
Measurement data	Data obtained by measuring objects or events.
Measures of central tendency	Numerical values referring to the center of the distribution.
Median location	The location of the median in an ordered series.
Median (Med)	The score corresponding to the point having 50% of the observations below it when observations are arranged in numerical order.
Mesokurtic	A distribution with a neutral degree of kurtosis.
Midpoints	Center of interval -- average of upper and lower limits.
Mode (Mo)	The most commonly occurring score.
Monotonic relationship	A relationship represented by a regression line that is continually increasing (or decreasing), but perhaps not in a straight line.
Multiplicative law of probability	The rule giving the probability of the joint occurrence of independent events.
Mutually exclusive	Two events are mutually exclusive when the occurrence of one precludes the occurrence of the other.
Negative relationship	A relationship in which increases in one variable are associated with decreases in the other.
Negatively skewed	A distribution that trails off to the left.
Nominal scale	Numbers used only to distinguish among objects.
normal distribution	A specific distribution having a characteristic bell-shaped form.
Ordinal scale	Numbers used only to place objects in order.
Ordinate	Vertical axis.
Outlier	An extreme point that stands out from the rest of the distribution.
p level	The probability that a particular result would occur by chance if H_0 is true. The exact probability of a Type I error.
Parameters	Numerical values summarizing population data.
Parametric tests	Statistical tests that involve assumptions about, or estimation of, population parameters.
Pearson product-moment correlation coefficient (*r*)	The most common correlation coefficient.

Percentile	The point below which a specified percentage of the observations fall.
Phi	The correlation coefficient when both of the variables are measured as dichotomies.
Platykurtic	A distribution that is relatively thick in the "shoulders."
Pooled variance	A weighted average of the separate sample variances.
Population variance	Variance of the population ‹usually estimated, rarely computed.
Population	Complete set of events in which you are interested.
Positively skewed	A distribution that trails off to the right.
Power	The probability of correctly rejecting a false H_0 .
Predictor variable	The variable from which a prediction is made.
Protected *t*	A technique in which we run *t* tests between pairs of means only if the analysis of variance was significant.
Quantitative data	Data obtained by measuring objects or events.
Random sample	A sample in which each member of the population has an equal chance of inclusion.
Random Assignment	Assigning participants to groups or cells on a random basis.
Range	The distance from the lowest to the highest score.
Range restrictions	Refers to cases in which the range over which X or Y varies is artificially limited.
Ranked data	Data for which the observations have been replaced by their numerical ranks from lowest to highest.
Rank - randomization tests	A class of nonparametric tests based on the theoretical distribution of randomly assigned ranks.
Ratio scale	A scale with a true zero point -- ratios are meaningful.
Real lower limit	The points halfway between the top of one interval and the bottom of the next.
Real upper limit	The points halfway between the top of one interval and the bottom of the next.
Rectangular distribution	A distribution in which all outcomes are equally likely.
Regression	The prediction of one variable from knowledge of one or more other variables.
Regression equation	The equation that predicts Y from X.
Regression coefficients	The general name given to the slope and the intercept ‹most often refers just to the slope.
Rejection region	The set of outcomes of an experiment that will lead to rejection of H_0 .
Rejection level	The probability with which we are willing to reject H0 when it is in fact correct.
Related samples	An experimental design in which the same subject is observed under more than one treatment.
Relative frequency view	Definition of probability in terms of past performance.

Research hypothesis	The hypothesis that the experiment was designed to investigate.
Sample	Set of actual observations. Subset of the population.
Sample statistics	Statistics calculated from a sample and used primarily to describe the sample.
Sample variance (s^2)	Sum of the squared deviations about the mean divided by N - 1.
Sample with replacement	Sampling in which the item drawn on trial N is replaced before the drawing on trial N + 1.
Sampling distribution of differences between means	The distribution of the differences between means over repeated sampling from the same population(s).
Sampling distribution of the mean	The distribution of sample means over repeated sampling from one population.
Sampling distributions	The distribution of a statistic over repeated sampling from a specified population.
Sampling error	Variability of a statistic from sample to sample due to chance.
Scales of measurement	Characteristics of relations among numbers assigned to objects.
Scatter plot	A figure in which the individual data points are plotted in two-dimensional space.
Scatter diagram	A figure in which the individual data points are plotted in two-dimensional space.
Scattergram	A figure in which the individual data points are plotted in two-dimensional space.
Sigma	Symbol indicating summation.
Significance level	The probability with which we are willing to reject H_0 when it is in fact correct.
Simple effect	The effect of one independent variable at one level of another independent variable.
Skewness	A measure of the degree to which a distribution is asymmetrical.
Slope	The amount of change in Y for a one unit change in X.
Spearman's correlation coefficient for ranked data (r_s)	A correlation coefficient on ranked data.
Standard deviation	Square root of the variance.
Standard error	The standard deviation of a sampling distribution.
Standard error of differences between means	The standard deviation of the sampling distribution of the differences between means.
Standard error of estimate	The average of the squared deviations about the regression line.
Standard scores	Scores with a predetermined mean and standard deviation.
Standard normal	A normal distribution with a mean equal to 0 and variance equal to 1.

distribution	Denoted *N* (0, 1).
Statistics	Numerical values summarizing sample data.
Student's *t* distribution	The sampling distribution of the *t* statistic.
Subjective probability	Definition of probability in terms of personal subjective belief in the likelihood of an outcome.
Sufficient statistic	A statistic that uses all of the information in a sample.
Sums of squares	The sum of the squared deviations around some point (usually a mean or predicted value).
Symmetric	Having the same shape on both sides of the center.
T scores	A set of scores with a mean of 50 and a standard deviation of 10.
Test statistics	The results of a statistical test.
Type I error	The error of rejecting H_0 when it is true.
Type II error	The error of not rejecting H_0 when it is false.
Unconditional probability	The probability of one event *ignoring* the occurrence or nonoccurrence of some other event.
Unimodal	A distribution having one distinct peak.
Variables	Properties of objects that can take on different values.
Weighted average	The mean of the form: $(a_1X_1 + a_2X_2)/(a_1 + a_2)$ where a_1 and a_2 are weighting factors and X_1 and X_2 are the values to be average.

Reg. No ----------------

(16MMU603A)

Karpagam Academy of Higher Education

COIMBATORE-21

DEPARTMENT OF MATHEMATICS

Sixth Semester

First Internal Test – December 2018

Probability and Statistics

Date : 18.12.2018 (FN) **Class : III-B.Sc Mathematics**

Time: 2 Hours **Maximum: 50 Marks**

PART - A (20 x 1 =20 Marks)

Answer All the Questions

1. Which of the following represents Median?
   a) First Quartile b) 50th Percentile c) 6th Decile d) 3rd Quartile
2. Shoe size of most of the people in India is No. 7. Which measure of central value does it represent?
   a). Mean b) Second Quartile c) Eighth Decile d) Mode
3. Median can be located graphically with the help of
   a) Histogram b) Ogives c) Bar Diagram d) Scatter Diagram
4. The first quartile divides a frequency distribution in the ratio
   a) 4 : 1 b) 1 :4 c) 3 : 1 d) 1 : 3
5. The least value in a collection of data is 14.1. If the range of the collection is 28.4, then the greatest value of the collection is
   (a) 42.5 (b) 43.5 (c) 42.4 (d) 42.1
6. The greatest value of a collection of data is 72 and the least value is 28. Then the Coefficient of Range is
   (a) 44 (b) 0.72 (c) 0.44 (d) 0.28
7. If the Standard Deviation of a set of data is 1.6, then the Variance is
   (a) 0.4 (b) 2.56 (c) 1.96 (d) 0.04
8. If the Variance of a data is 12.25, then the S.D is
   (a) 3.5 (b) 3 (c) 2.5 (d) 3.25
9. Mean and Standard Deviation of a data are 48 and 12 respectively. The Coefficient of Variation is
   (a) 42 (b) 25 (c) 28 (d) 48
10. When tossing a coin, the Probability of getting tail is
    (a) 1/2 (b) 1/3 (c) 0 (d) 1
11. The probability of drawing a Queen from a pack of cards is
    (a) 1/4 (b) 1/12 (c) 1/11 (d) 1/13
12. A variable whose value is a number determined by the outcome of a random experiment is called a ……
    (a) Sample (b) Random variable
    (c) Outcome (d) Event
13. If a random variable takes only a finite or a countable number of values is called
    (a) Finite random space (b) Continuous random variable
    (c) Discrete random variable (d) Infinite random variable
14. The term **Statistic** refers to the statistical measures relating to
    (a) Population (b) Hypothesis
    (c) Sample (d) Parameter

15. Larger group from which the sample is drawn is called ___

(a) Statistic (b) Sampling (c) Universe (d) Parameter

16. The distribution of means of all possible samples taken from a population is

(a) A sampling distribution (b) A Sample

(c) Population distribution (d) Parameter distribution

17. If P (A) = 0.25, P(B) = 0.50, P (A∩B) = 0.14 then P(neither A nor B) =

(a) 0.39 (b) 0.25 (c) 0.11 (d) 0.24

18. Two Dice are thrown simultaneously. The probability of getting a doublet is

(a) 1/36 (b) 1/3 (c) 1/6 (d) 2/3

19. Probability of sure event is

(a) 1 (b) 0 (c) 100 (d) 0.1

20. If φ is an impossible event, then *P(φ)* =

(a) 1 (b) 1/4 (c) 0 (d) 1/2

PART-B (3 x 2 =6 Marks)

Answer All the Questions

21. Define Statistics.
22. Write the Axioms of Probability.
23. Differentiate between Permutation and Combination.

PART-C (3 x 8 =24 Marks)

Answer All the Questions

24. a) Calculate the Mean and Median for the following data.

CI :	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
f:	4	14	20	51	32	17	6	4

(OR)

b) Calculate the Quartile Deviation and Mode.

Marks :	0 – 19	20 – 39	40 - 59	60 - 70	80 - 99
No. of students:	5	20	35	20	12

25. a) Find the Standard Deviation and CV for the following data.

391, 384, 591, 407, 672, 522, 777, 733, 1490, 2488

(OR)

b) Calculate the Standard Deviation and CV for the following data.

X :	0-10	10-20	20-30	30-40	40-50
f :	2	5	9	3	1

26. (a) Explain the Following

i) Event and Mutually Exclusive Event
ii) Exclusive and Exhaustive Events
iii) Dependent and Independent Events
iv) Simple and Compound Events

(OR)

(b) State and prove the Addition and Multiplication theorems of Probability.

Reg. No ----------------

(16MMU603A)

Karpagam Academy of Higher Education

COIMBATORE-21

DEPARTMENT OF MATHEMATICS

Sixth Semester

First Internal Test – December 2018

Probability and Statistics

Date : 18.12.2018 (FN) **Class : III-B.Sc Mathematics**

Time: 2 Hours **Maximum: 50 Marks**

PART - A (20 x 1 =20 Marks)

Answer All the Questions

1. Which of the following represents Median?
   a) First Quartile **b) 50th Percentile** c) 6th Decile d) 3rd Quartile
2. Shoe size of most of the people in India is No. 7. Which measure of central value does it represent?
   a). Mean b) Second Quartile c) Eighth Decile **d) Mode**
3. Median can be located graphically with the help of
   b) Histogram **b) Ogives** c) Bar Diagram d) Scatter Diagram
4. The first quartile divides a frequency distribution in the ratio
   a) 4 : 1 b) 1 :4 c) 3 : 1 **d) 1 : 3**
5. The least value in a collection of data is 14.1. If the range of the collection is 28.4, then the greatest value of the collection is
   (a) 42.5 (b) 43.5 (c) 42.4 (d) 42.1
6. The greatest value of a collection of data is 72 and the least value is 28. Then the Coefficient of Range is
   (a) 44 (b) 0.72 **(c) 0.44** (d) 0.28
7. If the Standard Deviation of a set of data is 1.6, then the Variance is
   (a) 0.4 **(b) 2.56** (c) 1.96 (d) 0.04
8. If the Variance of a data is 12.25, then the S.D is
   (a) 3.5 (b) 3 (c) 2.5 (d) 3.25
9. Mean and Standard Deviation of a data are 48 and 12 respectively. The Coefficient of Variation is
   (a) 42 **(b) 25** (c) 28 (d) 48
10. When tossing a coin, the Probability of getting tail is
    (a) 1/2 (b) 1/3 (c) 0 (d) 1
11. The probability of drawing a Queen from a pack of cards is
    (a) 1/4 (b) 1/12 (c) 1/11 **(d) 1/13**
12. A variable whose value is a number determined by the outcome of a random experiment is called a ……
    (a) Sample **(b) Random Variable**
    (c) Outcome (d) Event
13. If a random variable takes only a finite or a countable number of values is called
    (a) Finite Random Space (b) Continuous Random Variable
    (c) Discrete Random Variable (d) Infinite Random Variable
14. The term **Statistic** refers to the statistical measures relating to
    (a) Population (b) Hypothesis
    (c) **Sample** (d) Parameter

15. Larger group from which the sample is drawn is called ___

(a) Statistic (b) Sampling **(c) Universe** (d) Parameter

16. The distribution of means of all possible samples taken from a population is

(a) A sampling distribution (b) A Sample

(c) Population distribution (d) Parameter distribution

17. If P (A) = 0.25, P(B) = 0.50, P (A∩B) = 0.14 then P(neither A nor B) =

(a) 0.39 (b) 0.25 (c) 0.11 (d) 0.24

18. Two Dice are thrown simultaneously. The probability of getting a doublet is

(a) 1/36 (b) 1/3 (c) 1/6 (d) 2/3

19. Probability of sure event is

(a) 1 (b) 0 (c) 100 (d) 0.1

20. If φ is an impossible event, then *P(φ)* =

(a) 1 (b) 1/4 **(c) 0** (d) 1/2

PART-B (3 x 2 =6 Marks)

Answer All the Questions

21. Define Statistics.
22. Write the Axioms of Probability.
23. Differentiate between Permutation and Combination.

PART-C (3 x 8 =24 Marks)

Answer All the Questions

24. a) Calculate the Mean and Median for the following data.

CI :	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
f:	4	14	20	51	32	17	6	4

(OR)

b) Calculate the Quartile Deviation and Mode.

Marks :	0 – 19	20 – 39	40 - 59	60 - 70	80 - 99
No. of students:	5	20	35	20	12

25. a) Find the Standard Deviation and CV for the following data.

391, 384, 591, 407, 672, 522, 777, 733, 1490, 2488

(OR)

b) Calculate the Standard Deviation and CV for the following data.

X :	0-10	10-20	20-30	30-40	40-50
f :	2	5	9	3	1

26. (a) Explain the Following

v) Event and Mutually Exclusive Event
vi) Exclusive and Exhaustive Events
vii) Dependent and Independent Events
viii) Simple and Compound Events

(OR)

(b) State and prove the Addition and Multiplication theorems of Probability.

Reg. No ----------------

(16MMU603A)

Karpagam Academy of Higher Education
COIMBATORE-21
DEPARTMENT OF MATHEMATICS
Sixth Semester
Second Internal Test – February - 2019
Probability and Statistics

Date : 05.01.2019 (FN) **Class : III-B.Sc Mathematics**
Time: 2 Hours **Maximum: 50 Marks**

PART - A (20 x 1 =20 Marks)

Answer All the Questions

1. A numerical description of the outcome of an experiment is called a
   a. descriptive statistic b. probability function
   c. variance d. random variable
2. A random variable that can assume only a finite number of values is referred to as a(n)
   a. infinite sequence b. finite sequence
   c. discrete random variable
   d. discrete probability function
3. A probability distribution showing the probability of x successes in n trials, where the probability of success does not change from trial to trial, is termed a
   a. uniform probability distribution
   b. binomial probability distribution
   c. hypergeometric probability distribution
   d. normal probability distribution
4. A description of the distribution of the values of a random variable and their associated probabilities is called a
   a. probability distribution b. random variance
   c. random variable d. expected value
5. Which of the following is(are) required condition(s) for a discrete probability function?
   a. $\Sigma f(x) = 0$ b. $f(x) \geq 1$ for all values of x
   c. $f(x) < 0$ d. both a and b are correct
6. A measure of the average value of a random variable is
   a. variance b. standard deviation
   c. expected value d. range
7. An experiment consists of making 80 telephone calls in order to sell a particular insurance policy. The random variable in this experiment is a
   a. discrete random variable
   b. continuous random variable
   c. complex random variable
   d. None of the above answers is correct.
8. An experiment consists of determining the speed of automobiles on a highway by the use of radar equipment. The random variable in this experiment is a
   a. discrete random variable
   b. continuous random variable
   c. complex random variable
   d. None of the above answers is correct.
9. If you are conducting an experiment where the probability of a success is .02 and you are interested in the probability of 4 successes in 15 trials, the correct probability function to use is
   a. standard normal probability density function
   b. normal probability density function
   c. Poisson probability function
   d. binomial probability function
10. The Poisson probability distribution is a
    a. continuous probability distribution
    b. discrete probability distribution
    c. uniform probability distribution
    d. normal probability distribution
11. The binomial probability distribution is used with
    a. a continuous random variable
    b. a discrete random variable
    c. uniform probability distribution
    d. normal probability distribution

12. Which of the following is a characteristic of a binomial experiment?
    a. at least 2 outcomes are possible
    b. the probability changes from trial to trial
    c. the trials are independent
    d. All of the above answers are correct.
13. Assume that you have a binomial experiment with p = 0.5 and a sample size of 100. The expected value of this distribution is
    a. 0.50 b. 0.30 c. 100 d. 50
14. Which of the following is not a property of a binomial experiment?
    a) the experiment consists of a sequence of n identical trials
    b) each outcome can be referred to as a success or a failure
    c) the probabilities of the two outcomes can change from one trial to the next
    d) the trials are independent
15. The Poisson probability distribution is used with
    a) a continuous random variable
    b) a discrete random variable
    c) either a continuous or discrete random variable
    d) any random variable
16. The standard deviation of a binomial distribution is
    a. Square root of Pn(1 - n)
    b. Square root of P(1 - P)
    c. Square root of nP
    d. Square root of nP(1 - P)
17. The expected value for a binomial probability distribution is
    a. E(x) = Pn(1 - n) b. E(x) = P(1 - P)
    c. E(x) = nP d. E(x) = nP(1 - P)
18. The variance for the binomial probability distribution is
    a. var(x) = P(1 - P) b. var(x) = nP
    c. var(x) = n(1 - P) d. var(x) = nP(1 - P)
19. Assume that you have a binomial experiment with p = 0.4 and a sample size of 50. The variance of this distribution is
    a. 20 b. 12 c. 3.46
    d. Not enough information is given to answer this question.
20. A continuous variable can be described as:
    a) Able to take only certain discrete values within a range of scores
    b) Able to take any value within a range of scores
    c) Being made up of categories
    d) Being made up of variables

PART-B (3 x 2 =6 Marks)

Answer All the Questions

21. Define moments.
22. Formula for Mean and Variance of Binomial Distribution.
23. Write any two conditions of a Poisson Distribution.

PART-C (3 x 8 =24 Marks)

Answer All the Questions

24. a) A test paper containing 10 problems is given to three students A, B, C. It is considered that student A can solve 60% problems, student B can solve 40% problems and student C can solve 30% problems. Find the probability that the problem chosen from the test paper will be solved by all the three students.

(OR)

b) Describe about cumulative distribution function and probability density functions.

25. a) If 20% of the electric bulbs manufactured by a certain company are defective. Find the probability that in a sample of 200 bulbs i) less than 2 bulbs ii) more than 3 bulbs are defective [Given: e-4 = 0.0183].

(OR)

b) Briefly explain the properties of Binomial distribution with example.

26. a) Define and write the properties of Poisson Distribution.

(OR)

b) Define and explain Marginal and Conditional distributions.

Reg. No ----------------

(16MMU603A)

Karpagam Academy of Higher Education

COIMBATORE-21

DEPARTMENT OF MATHEMATICS

Sixth Semester

Second Internal Test – February - 2019

Probability and Statistics

Date : 05.01.2019 (FN) **Class : III-B.Sc Mathematics**

Time: 2 Hours **Maximum: 50 Marks**

PART - A (20 x 1 =20 Marks)

Answer All the Questions

1. A numerical description of the outcome of an experiment is called a
   a. descriptive statistic b. probability function
   c. variance **d. random variable**
2. A random variable that can assume only a finite number of values is referred to as a(n)
   a. infinite sequence b. finite sequence
   c. discrete random variable
   d. discrete probability function
3. A probability distribution showing the probability of x successes in n trials, where the probability of success does not change from trial to trial, is termed a
   a. uniform probability distribution
   b. binomial probability distribution
   c. hypergeometric probability distribution
   d. normal probability distribution
4. A description of the distribution of the values of a random variable and their associated probabilities is called a
   a. probability distribution b. random variance
   c. random variable d. expected value
5. Which of the following is(are) required condition(s) for a discrete probability function?
   a. $\Sigma f(x) = 0$ **b. $f(x) \geq 1$ for all values of x**
   c. $f(x) < 0$ d. both a and b are correct
6. A measure of the average value of a random variable is
   a. variance b. standard deviation
   c. expected value d. range
7. An experiment consists of making 80 telephone calls in order to sell a particular insurance policy. The random variable in this experiment is a
   a. discrete random variable
   b. continuous random variable
   c. complex random variable
   d. None of the above answers is correct.
8. An experiment consists of determining the speed of automobiles on a highway by the use of radar equipment. The random variable in this experiment is a
   a. discrete random variable
   b. continuous random variable
   c. complex random variable
   d. None of the above answers is correct.
9. If you are conducting an experiment where the probability of a success is .02 and you are interested in the probability of 4 successes in 15 trials, the correct probability function to use is
   a. standard normal probability density function
   b. normal probability density function
   c. Poisson probability function
   d. binomial probability function
10. The Poisson probability distribution is a
    a. continuous probability distribution
    b. discrete probability distribution
    c. uniform probability distribution
    d. normal probability distribution
11. The binomial probability distribution is used with
    a. a continuous random variable
    b. a discrete random variable
    c. uniform probability distribution
    d. normal probability distribution

12. Which of the following is a characteristic of a binomial experiment?
    a. at least 2 outcomes are possible
    b. the probability changes from trial to trial
    c. the trials are independent
    d. All of the above answers are correct.
13. Assume that you have a binomial experiment with p = 0.5 and a sample size of 100. The expected value of this distribution is
    a. 0.50 b. 0.30 c. 100 **d. 50**
14. Which of the following is <u>not</u> a property of a binomial experiment?
    a) the experiment consists of a sequence of n identical trials
    b) each outcome can be referred to as a success or a failure
    c) the probabilities of the two outcomes can change from one trial to the next
    d) the trials are independent
15. The Poisson probability distribution is used with
    a) a continuous random variable
    b) a discrete random variable
    c) either a continuous or discrete random variable
    d) any random variable
16. The standard deviation of a binomial distribution is
    a. Square root of Pn(1 - n)
    b. Square root of P(1 - P)
    c. Square root of nP
    d. Square root of nP(1 - P)
17. The expected value for a binomial probability distribution is
    a. E(x) = Pn(1 - n) b. E(x) = P(1 - P)
    c. E(x) = nP d. E(x) = nP(1 - P)
18. The variance for the binomial probability distribution is
    a. var(x) = P(1 - P) b. var(x) = nP
    c. var(x) = n(1 - P) **d. var(x) = nP(1 - P)**
19. Assume that you have a binomial experiment with p = 0.4 and a sample size of 50. The variance of this distribution is
    a. 20 **b. 12** c. 3.46
    d. Not enough information is given to answer this question.
20. A continuous variable can be described as:
    a) Able to take only certain discrete values within a range of scores
    b) Able to take any value within a range of scores
    c) Being made up of categories
    d) Being made up of variables

PART-B (3 x 2 =6 Marks)

Answer All the Questions

21. Define moments.
22. Formula for Mean and Variance of Binomial Distribution.
23. Write any two conditions of a Poisson Distribution.

PART-C (3 x 8 =24 Marks)

Answer All the Questions

24. a) A test paper containing 10 problems is given to three students A, B, C. It is considered that student A can solve 60% problems, student B can solve 40% problems and student C can solve 30% problems. Find the probability that the problem chosen from the test paper will be solved by all the three students.

    (OR)

    b) Describe about cumulative distribution function and probability density functions.
25. a) If 20% of the electric bulbs manufactured by a certain company are defective. Find the probability that in a sample of 200 bulbs i) less than 2 bulbs ii) more than 3 bulbs are defective [Given: e-4 = 0.0183].

    (OR)

    b) Briefly explain the properties of Binomial distribution with example.
26. a) Define and write the properties of Poisson Distribution.

    (OR)

    b) Define and explain Marginal and Conditional distributions

Reg. No ----------------

(16MMU603A)

Karpagam Academy of Higher Education

Coimbatore-21

DEPARTMENT OF MATHEMATICS

Sixth Semester

Third Internal Test – March - 2019

Probability and Statistics

Date : 12.03.2019 (FN) **Class : III-B.Sc Mathematics**

Time: 2 Hours **Maximum: 50 Marks**

PART - A (20 x 1 =20 Marks)

Answer All the Questions

1. An experiment consists of determining the speed of automobiles on a highway by the use of radar equipment. The random variable in this experiment is a
   a) discrete random variable b) continuous random variable
   c) complex random variable d) None of the above answers.
2. Which one of the following prefers the term correlation?
   a) Relationship between two values
   b) Relationship between two variables
   c) Average relationship between two variables
   d) Relationship between two things
3. Rank correlation was discovered by
   a) R.A Fisher b) Sir Francis Galton
   c) Karl Pearson d) Spearman
4. The regression line cut each other at the point of
   a) average of X only b) average of Y only
   c) average of X and Y d) the median of X on Y
5. ------------ is independent of origin and scale.
   a) correlation coefficient b) regression coefficients
   c) coefficient of range d) coefficient of variation
6. Small sample test is also known as --------------------.
   a) Exact test b) t – test c) Normal test d) F-test
7. Using a 98% confidence interval instead of a 95% confidence interval for the same population and sample, will give a confidence interval which is:
   a) Narrower b) Wider c) The same width d) No interval
8. While drawing a scatter diagram if all points appear to form a straight line getting downward from left to right, then it is inferred that there is
   a) perfect positive correlation b) simple positive correlation
   c) perfect negative correlation d) no correlation
9. Regression coefficient is independent of
   a) origin b) scale
   c) both origin and scale d) neither origin nor scale.
10. --------- is used to measure closeness of relationship between variables.
    a) regression b) mean
    c) rank correlation d) correlation
11. When b_{xy}=0.85 and b_{yx}= 0.89, then correlation coefficient r =
    a) 0.98 b) 0.5 c) 0.68 d) 0.87
12. If the sample size is less than 30, then the sample is called
    a) Large sample b) small sample c) population d) parameter
13. If the computed value is less than the critical value, then -------------
    a) Null hypothesis is accepted
    b) Null hypothesis is rejected
    c) Alternative hypothesis is accepted
    d) Alternative hypothesis is rejected
14. Coefficient of correlation value lies between
    a) 1 and –1 b) 0 and 1 c) 0 and ∞ d) 0 and –1.

15. If r =1, then the angle between two lines of regression is
    a) Zero degree   b) sixty degree
    c) ninety degree   d) thirty degree
16. Which of the following measurement scales is required for the valid calculation of Karl Pearson's correlation coefficient?
    a) Ordinal   b) interval   c) Ratio   d) nominal
17. When the two regression lines coincide, then r is
    a) 0   b) 1   c) -1   d) either 1 or -1
18. If the sample size is greater than 30, then the sample is called --------
    a) Large sample   b) small sample   c) population   d) statistic
19. Suppose we have two populations, one with a smaller standard deviation than the other. We take two samples of the same size, one from each population, and work out a 95% confidence interval for each mean. The confidence interval for the population with the smaller standard deviation will be:
    a) Narrower   b) Wider   c) The same width   d) No Interval
20. A researcher wishes to find out about the mean of a population. She takes a sample, calculates the sample mean and works out the 95% confidence interval for the population mean. Which of the following results would she prefer?
    a) The confidence interval is very wide
    b) The confidence interval is very narrow
    c) It makes no difference, what the width of the confidence interval is
    d) The same width

PART-B (3 x 2 =6 Marks)

Answer All the Questions

21. Write the Spearman's Rank Correlation formula for tied Scores.
22. Define Chebyshev's inequality.
23. Write a brief note on classification of state.

PART-C (3 x 8 =24 Marks)

Answer All the Questions

24. a) Describe about Joint probability distribution function and its properties.

(OR)

b) Calculate Karl Pearson's Coefficient of Correlation from the following data

Wages in Rs.	110	107	102	105	100	99	97	98
Cost of Living	98	99	95	97	96	92	90	91

25. a) From the data given below find the two regression lines.

X	10	12	13	12	16	15
Y	40	38	43	45	37	43

Also estimate Y when X = 21.

b) Describe how to interpret the weak and strong law of large numbers.

26. Explain the Central Limit theorem for independent and identically distributed random variables with finite variance.

(OR)

(b) State and explain the Chapman-Kolmogorov equation.

*****

Reg. No ----------------

(16MMU603A)

Karpagam Academy of Higher Education
Coimbatore-21
DEPARTMENT OF MATHEMATICS
Sixth Semester
Third Internal Test – March - 2019
Probability and Statistics

Date : 12.03.2019 (FN) **Class : III-B.Sc Mathematics**
Time: 2 Hours **Maximum: 50 Marks**

PART - A (20 x 1 =20 Marks)

Answer All the Questions

1. An experiment consists of determining the speed of automobiles on a highway by the use of radar equipment. The random variable in this experiment is a
   a) discrete random variable **b) continuous random variable**
   c) complex random variable d) None of the above answers.
2. Which one of the following prefers the term correlation?
   a) Relationship between two values
   b) Relationship between two variables
   c) Average relationship between two variables
   d) Relationship between two things
3. Rank correlation was discovered by
   a) R.A Fisher b) Sir Francis Galton
   c) Karl Pearson **d) Spearman**
4. The regression line cut each other at the point of
   a) average of X only b) average of Y only
   c) average of X and Y d) the median of X on Y
5. ------------ is independent of origin and scale.
   a) correlation coefficient b) regression coefficients
   c) coefficient of range d) coefficient of variation
6. Small sample test is also known as --------------------.
   a) Exact test b) **t – test** **c)** Normal test d) F-test
7. Using a 98% confidence interval instead of a 95% confidence interval for the same population and sample, will give a confidence interval which is:
   a) Narrower **b) Wider** c) The same width d) No interval
8. While drawing a scatter diagram if all points appear to form a straight line getting downward from left to right, then it is inferred that there is
   a) perfect positive correlation b) simple positive correlation
   c) perfect negative correlation d) no correlation
9. Regression coefficient is independent of
   a) origin b) scale
   c) both origin and scale d) neither origin nor scale.
10. --------- is used to measure closeness of relationship between variables.
    a) regression b) mean
    c) rank correlation **d) correlation**
11. When b_{xy}=0.85 and b_{yx}= 0.89, then correlation coefficient r =
    a) 0.98 b) 0.5 c) 0.68 **d) 0.87**
12. If the sample size is less than 30, then the sample is called
    a) Large sample b) **small sample** **c)** population d) parameter
13. If the computed value is less than the critical value, then -------------
    a) **Null hypothesis is accepted** b) Null hypothesis is rejected
    c) Alternative hypothesis is accepted
    d) Alternative hypothesis is rejected
14. Coefficient of correlation value lies between
    a) 1 and –1 b) 0 and 1 c) 0 and ∞ d) 0 and –1.
15. If r =1, then the angle between two lines of regression is
    a) Zero degree b) sixty degree
    c) ninety degree d) thirty degree
16. Which of the following measurement scales is required for the valid calculation of Karl Pearson's correlation coefficient?
    a) Ordinal **b) interval** c) Ratio d) nominal
17. When the two regression lines coincide, then r is
    a) 0 b) 1 c) -1 d) either 1 or -1
18. If the sample size is greater than 30, then the sample is called --------
    a) **Large sample** b) small sample c) population d) statistic

19. Suppose we have two populations, one with a smaller standard deviation than the other. We take two samples of the same size, one from each population, and work out a 95% confidence interval for each mean. The confidence interval for the population with the smaller standard deviation will be:
    a) Narrower b) Wider c) The same width d) No Interval
20. A researcher wishes to find out about the mean of a population. She takes a sample, calculates the sample mean and works out the 95% confidence interval for the population mean. Which of the following results would she prefer?
    a) The confidence interval is very wide
    b) The confidence interval is very narrow
    c) It makes no difference, what the width of the confidence interval is
    d) The same width

PART-B (3 x 2 =6 Marks)

Answer All the Questions

21. Write the formula for Spearman's Rank Correlation for tied Scores.
22. Define Chebyshev's inequality.
23. Write a brief note on classification of state.

.

PART-C (3 x 8 =24 Marks)

Answer All the Questions

24. a) Describe about Joint probability distribution function and its properties.

(OR)

b) Calculate Karl Pearson's Coefficient of Correlation from the following data

Wages in Rs.	110	107	102	105	100	99	97	98
Cost of Living	98	99	95	97	96	92	90	91

25. a) From the data given below find the two regression lines.

X	10	12	13	12	16	15
Y	40	38	43	45	37	43

Also estimate Y when X = 21.

b) Describe how to interpret the weak and strong law of large numbers.

26. Explain the Central Limit theorem for independent and identically distributed random variables with finite variance.

(OR)

(b) State and explain the Chapman-Kolmogorov equation.