Savitribai Phule Pune University

(Formerly University of Pune)

Faculty of Science & Technology

F.Y.B.A. (Applied Statistics)

Choice Based Credit System Syllabus

To be implemented from Academic Year 2019-2020

Title of the Course : B. A. (Applied Statistics)

Preamble of the syllabus:

The word ***Statistics*** is used in different ways in different contexts. To a cricket fan, Statistics is the information about runs scored or wickets taken by a player. To the manager of a manufacturing unit, Statistics may be the information about the process control. To a medical researcher investigating the effects of a new drug, Statistics are evidence of research efforts. For college student, Statistics are the grades or marks scored in a course. Thus, in all these illustrations Statistics word refers to quantitative data in the area under study. Statistics as a subject is an important branch of knowledge and is devoted to various techniques of collection, presentation, analysis and interpretation of data. It is a science of learning from data.

Statistics provides tools for making decisions when conditions of uncertainty prevail. Hence these tools and techniques are used in almost all fields. Statistics is indispensable for people working in fields like agriculture, business, management, economics, finance, insurance, education, biotechnology and medical science etc. Since last two decades, with the help of computers large amount of data can be handled and more sophisticated statistical techniques can be used in an effective manner. Knowledge of different aspects of Statistics has become crucial. There is a continuous demand for statisticians in every field – education, industry, software and research. The syllabus of the three Year B. A. degree course in Statistics is framed in such a way that the students at the end of the course can apply judiciously the statistical tools to a variety of data sets to arrive at some conclusions.

Statistics can be divided into two broad categories, (1) exploratory statistics or descriptive statistics, which is concerned with summarizing data and describing these data, and (2) confirmatory statistics or inferential statistics, which is concerned with making decisions about the population based on the sample.

Up to higher secondary school, students are mostly exposed to descriptive statistics. At the first year a student can take any one of the four subjects related statistics, such as Statistics, Applied Statistics, Mathematical Statistics and Statistical Prerequisites. If the student continues with these subjects at the second year and third year, it is expected that at the end of the degree course a student is able to apply the statistical tools to real life data.

Introduction: B. A. degree program is three years of duration, with semester pattern for the second and third year and annual examination pattern for the first year.

The structure of **Bachelor of Arts (B. A.) is as follows.** The student joining the First Year

B.A. Course has to take six subjects from 13 groups. The student cannot take more than one subject from one group. There are four subjects related to statistics. These are Statistics (Group L), Applied Statistics (Group L), Mathematical Statistics (Group J) and Statistical Prerequisites (Group K).

Structure of the Subject:

Structure of the subject for first and subsequent three years and the pattern of examination and question papers are as specified below.

a) Structure of F. Y. B. A. Statistics/Mathematical Statistics/Applied Statistics/ Statistical-Prerequisites

Semester	Subject	Subject code	Title	Credit	Marks
I	Statistics	ST- 13871	Descriptive Statistics I	3	100
	Mathematical Statistics	ST- 13271	Discrete Probability and Probability Distributions	3	100
	Applied Statistics	ST-14171	Descriptive Statistics I	3	100
	Statistical Pre-requisites	ST-13571	Descriptive Statistics I	3	100
II	Statistics	ST- 13872	Descriptive Statistics II	3	100
	Mathematical Statistics	ST- 13272	Discrete Probability Distributions	3	100
	Applied Statistics	ST-14172	Descriptive Statistics II	3	100
	Statistical Pre-requisites	ST-13572	Descriptive Statistics II	3	100

EQUIVALENCE FOR THEORY PAPERS: (From June 2019)

The 2013 pattern was an annual examination and 2019 pattern is CBCS hence equivalence cannot be given.

Detailed Syllabus for F.Y.B.A. Applied Statistics

Objectives: The main objective of this course is to acquaint students with some basic concepts in statistics. They will be introduced to some elementary statistical methods of analysis of data. At the end of this course students are expected to be able

(i) to compute various measurements of central tendency, dispersion, skewness and kurtosis.

(ii) to compute the correlation coefficient from ungrouped bivariate data and interpret them.

(iii) to analyze data pertaining to attributes and to interpret results.

(iv) to analyze data pertaining to Time Series and to interpret the results.

Note: Mathematical derivations and proofs are not expected

Semester I

ST-14171: Descriptive Statistics-I

1. Population and Sample: **(5L)**

1.1Notion of a statistical population: Finite population, infinite population, homogeneous population and heterogeneous population.

1.2Notion of a sample, random sample. Methods of sampling: Simple random sampling with and without replacement (SRSWR & SRSWOR), Stratified random sampling and Systematic sampling.

2. Types of Data and its Presentation: **(10L)**

2.1Attributes: Nominal scale and ordinal scale

2.2Variable: Interval scale, ratio scale. Discrete and continuous variables.

2.3Raw data and types of data: primary data and secondary data. Classification of attribute type raw data, classification of variable type data (inclusive and exclusive methods). Open end classes, ungrouped frequency distribution, Sturge's rule, Grouped frequency distribution, cumulative frequency distribution and relative frequency distribution.

2.4Graphical presentation of data: Bar Diagrams, Pie Diagram, Histogram, frequency curve, Ogive curves, stem and leaf chart.

3. Statistical Averages: **(10L)**

3.1Concept of central tendency of statistical data: Statistical average, characteristics of a good statistical average.

3.2Arithmetic Mean (A.M.) Definition, effect of change of origin and scale (Only statement), combined mean of a number of groups (Only formula), trimmed arithmetic mean.

3.3Mode : Definition, formulae. Median : Definition, formulae. Empirical relation between mean, median and mode.

3.4Partition Values : Quartiles, Deciles and Percentiles

3.5Geometric Mean (G.M.): Definition and formulae. Harmonic Mean (H.M.): Definition and Formulae.Order relation between arithmetic mean, geometric mean, harmonic mean.

3.6Weighted Mean : Weighted A.M., G.M. and H.M.Situations where one kind of average is preferable to others.

4.Measures of Dispersion: (10L)

4.1Concept of dispersion.

4.2Range : Definition.Semi-interquartile range (Quartile deviation).

4.3Mean deviation: Definition, minimality property (without proof).

4.4Mean square deviation: Definition, minimality property of mean square deviation (without proof) ,

4.5Variance and standard deviation: Definition, effect of change of origin and scale on variance(Only statement), Combined variance and Combined standard deviation (Only formula).

4.6Measures of dispersion for comparison: coefficient of range, coefficient of quartile deviation and coefficient of mean deviation, coefficient of variation (C.V.)

5.Moments: (5L)

5.1Raw moments (m'_r) for ungrouped and grouped data upto 4^{th} order.

5.2Central moments (m_r) for ungrouped and grouped data upto 4^{th} order. Effects of change of origin and scale (Only statement).Relation between central and raw moments upto 4^{th} order (Only statement).

6.Skewness and Kurtosis: (8L)

6.1Concept of skewness of frequency distribution, positive skewness, negative skewness, symmetric frequency distribution.

6.2Bowley's coefficient of skewness, Box Plot and its interpretation. Karl Pearson's coefficient of skewness. Measures of skewness based on moments ($\hat{\beta}_1$, $\hat{\gamma}_1$).

6.3Concepts of Kurtosis, Leptokurtic, Mesokurtic and Platykurtic frequency distribution. Measures of Kurtosis based on moments ($\hat{\beta}_2$, $\hat{\gamma}_2$).

Recommended Books:

1. Agarwal, B. L. (2003). Programmed Statistics, Second Edition, New Age International Publishers, New Delhi.
2. Draper, N. R. and Smith, H. (1998). Applied Regression Analysis Third Edition, John Wiley and Sons
3. Goon, A.M., Gupta, M.K. and Dasgupta, B. (1983). Fundamentals of Statistics, Vol. 1, Sixth Revised Edition,The World Press Pvt. Ltd., Calcutta.
4. Gupta, S.C. and Kapoor, V. K. (1983). Fundamentals of Mathematical Statistics, Eighth Edition, Sultan Chand and Sons Publishers, New Delhi.
5. Gupta, S.C. and Kapoor, V. K. (1997). Fundamentals of Applied Statistics, Third Edition, Sultan Chand and Sons Publishers, New Delhi.

6. Freund, J.E. (1977). Modern Elementary Statistics. Fourth Edition, Prentice Hall of India Private Limited, New Delhi.
7. Montgomery, D. C; Peck, E. A.; Vining, G. G. (2006). Introduction to Linear Regression Analysis, John Wiley and Sons
8. Purohit, S. G., Gore S. D., Deshmukh S. R. (2008). Statistics Using R, Narosa Publishing House, New Delhi.
9. Sharma, K. V. S. (2001). Statistics Made it Simple: Do it yourself on PC. Prentce Hall of India, New Delhi.
10. Snedecor G.W. and Cochran W.G.(1989), Statistical Methods, Eighth Ed. East- West Press.
11. Mukhopadhyay, P(1996). Mathematical Statistics, New Central Book Agency, Calcutta

Semester II

ST-14172: Descriptive Statistics-II

1.Correlation: (10L)

1.1Concept and objective of studying Bivariate data. Concepts of correlation between two variables, positive correlation, negative correlation. Scatter diagram, conclusion about the type of correlation from scatter diagram.

1.2Karl Pearson's coefficient of correlation (r): Definition, computation for ungrouped data and interpretation. Statement of properties:

(i) $-1 \leq r \leq +1$; (ii) Effects of change of origin and scale

1.3Spearman's rank correlation coefficient: Definition, computation and interpretation (without ties), In case of ties, compute Karl Pearson's correlation coefficient between ranks. (Spearman's rank correlation coefficient formula with correction for ties not expected.)

2.Linear Regression: (12L)

2.1Concept of regression, lines of regression, fitted equations of lines of regression by the least squares method (No derivation), interpretation of slope and intercept.

2.2 Regression coefficient (byx, bxy): Definition, computation, properties (without proof): i) byx*bxy = r2, ii) byx*bxy ≤1, iii) algebraic sign of regression coefficients are same, iv) Effect of change of origin and scale, v) Angle between the two lines of regression (only statement),vi) byx=rσyσx, vii) bxy=rσxσy.

2.3Concept of residual, plot of residual against X, concept of coefficient of determination with uses. Explained and unexplained variation.

3.Non-linear Regression: (8L)

3.1Concept of non-linear regression. Necessity and importance of drawing second degree curve.

3.2Fitting of second degree curve $Y = a + bX + cX^2$ by the method of least squares(No derivation).

3.3Fitting of exponential curves of the type $Y = ab^X$ *and* $Y = aX^b$ by the method of least squares(No derivation only procedure).

4.IndexNumbers: (10L)

4.1Definition and Meaning. Brief discussion on construction of index numbers.

4.2Simple and weighted price index numbers. Laspeyre's, Paasche's and Fisher's price index numbers.

4.3 Simple average of price relatives, weighted average of price relatives.

4.4Consumers price index numbers: Considerations in its construction, Computation of consumer's price index numbers by (i) family budget method (ii) aggregate expenditure method.

4.5Shifting of base, splicing, deflating and purchasing power.

5.Theory of Attributes: (8L)

5.1Attributes: Classification, Notion of dichotomy and manifold classification, class-frequency, order of class, positive class-frequency, negative class frequency, ultimate class-frequency, relationship among class-frequencies of different order (upto three attributes), dot operator, Fundamental set of class frequencies. Consistency of data upto 3 attributes.

5.2Concepts of independence and association of two attributes. Yule's coefficient of association (Q), interpretation on values of (Q).

Recommended Books:

1. Agarwal, B. L. (2003). Programmed Statistics, Second Edition, New Age International Publishers, New Delhi.
2. Draper, N. R. and Smith, H. (1998). Applied Regression Analysis Third Edition, John Wiley and Sons
3. Goon, A.M., Gupta, M.K. and Dasgupta, B. (1983). Fundamentals of Statistics, Vol. 1, Sixth Revised Edition,The World Press Pvt. Ltd., Calcutta.
4. Gupta, S.C. and Kapoor, V. K. (1983). Fundamentals of Mathematical Statistics, Eighth Edition, Sultan Chand and Sons Publishers, New Delhi.
5. Gupta, S.C. and Kapoor, V. K. (1997). Fundamentals of Applied Statistics, Third Edition, Sultan Chand and Sons Publishers, New Delhi.
6. Freund, J.E. (1977). Modern Elementary Statistics. Fourth Edition, Prentice Hall of India Private Limited, New Delhi.
7. Montgomery, D. C; Peck, E. A.; Vining, G. G. (2006). Introduction to Linear Regression Analysis, John Wiley and Sons
8. Purohit, S. G., Gore S. D., Deshmukh S. R. (2008). Statistics Using R, Narosa Publishing House, New Delhi.
9. Sharma, K. V. S. (2001). Statistics Made it Simple: Do it yourself on PC. Prentce Hall of India, New Delhi.
10. Snedecor G.W. and Cochran W.G.(1989), Statistical Methods, Eighth Ed. East- West Press.
11. Mukhopadhyay, P(1996). Mathematical Statistics, New Central Book Agency, Calcutta

Reference Websites:

1. www.stats.unipune.ac.in (100 Data sets for Statistics Education by Dr. Anil P. Gore, Dr. Mrs. S. A. Paranjpe and Madhav B. Kulkarni available in ISPS folder).
2. www.freestatistics.tk (National Statistical Agencies)
3. www.psychstat.smsu.edu/sbk00.htm (Online book)
4. www.bmj.bmjournals.com/collections/statsbk/index.shtml
5. www.statweb.calpoly.edu/bchance/stat-stuff.html
6. www.amstat.org/publications/jse/jse-data-archive.html (International journal on teaching and learning of statistics)
7. www.amstat.org/publications/chance (Chance magazine)
8. www.statsci.org/datasets.html (Data sets)
9. www.math.uah.edu/stat (Virtual laboratories in Statistics)
10. www.amstat.org/publications/stats (STATS : the magazine for students of Statistics)
11. www.stat.ucla.edu/cases (Case studies in Statistics).