University of Pune

Three Year B. A. Degree Program

Syllabus for F. Y. B. A. Statistical Prerequisites

(With effect from Academic Year 2013-2014)

Submitted by

Prof. S. R. Deshmukh
Chairperson,
Board of Studies in Statistics

1) Title of the program: Three Year B. A. Degree

2) Preamble to 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. To a 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 decade, 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. Sc. 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.

3) Introduction: Three Year B. A. degree program is of three years 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).

4) Eligibility

First Year B.A.
Higher secondary school certificate examination of the Maharashtra State Board of Higher Secondary Education or an equivalent examination of any other statutory Board or University with English as a passing subject.

Detailed Syllabus for F.Y.B.A. Statistical Prerequisites
The course in 'Statistical Pre-requisites' may be taken only by candidates, taking one of the Social Sciences as their Special subject at the B. A. Degree Examination. The course in 'Statistical Pre-requisites' cannot be taken by those who take any of the courses in Statistics Group for their B.A Examination.

Note: Mathematical derivations and proofs are not expected:
Objective: 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 Index Numbers and to interpret the results.

First Term

1 Population and Sample: (5)

1.1 Types of characteristics
Attributes: Nominal scale and ordinal scale
Variable: Interval scale, ratio scale. Discrete and continuous variables,
1.2 Raw data,types of data – primary data and secondary data
1.3 Notion of a statistical population: Finite population, infinite population, homogeneous population and heterogeneous population. Notion of a
1.4 sample, random sample.
Methods of sampling: Simple random sampling with and without replacement (SRSWR & SRSWOR), Stratified random sampling, Systematic sampling.

2 Presentation of Data: (8)

2.1 Classification: Discrete frequency distribution, Inclusive and Exclusive methods of classification, Open-end classes, Cumulative frequency distributions and Relative frequency distribution.
2.2 Graphical presentation of data, Histogram, frequency curve, Ogive curves, Stem and leaf chart.

3. Measures of Central Tendency: (12)

3.1 Concepts of central tendency , statistical average, characteristics of a good statistical average.

3.2 Arithmetic Mean (A.M.) Definition, Effects of change of origin and scale, Combined mean of two groups, Merits and Demerits

3.3 Geometric Mean (G.M.) Definition, Formulae, Merits and Demerits

3.4 Harmonic Mean (H.M.) Definition, Formulae, Merits and Demerits

3.5 Weighted Mean: Weighted A.M., G.M.,H.M.

3.6 Mode: Definition, formula for computation, graphical method of determination of mode, Merits and Demerits.

3.7 Median: Definition, formula for computation, graphical method of determination of median, Merits and Demerits.

3.8 Partition values: Quartiles, Deciles and percentile, graphical method of determination of quartiles, deciles and percentiles for grouped frequency distributions, Box plot.

4. Measures of Dispersion: **(12)**

4.1 Concepts of dispersion, characteristics of good measure of dispersion

4.2 Range: definition, Merits and Demerits

4.3 Semi-interquartile range (Quartile deviation), Merits and Demerits

4.4 Mean deviation: definition, Merits and demerits

4.5 Mean square deviation: Definition,

4.6 Variance and Standard deviation (S.D.): Definition, Merits and Demerits

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

4.8 Simple numerical problems.

5. Moments (4)

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

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

5.3 Relation between central and raw moments upto 4^{th} order.(Only statement).

6. Skewness and Kurtosis (7)

6.1 Concept of skewness of frequency distribution, postive skewness, negative skewness, symmeric frequency distribution.

6.2 Bowley's coefficient of skewness, interpretation using Box Plot.

6.3 Karl Pearson's coefficient of skewness

6.4 Measures of skewness based on moments

6.5 Concepts of Kurtosis, Leptokurtic, Mesokurtic and Platykurtic frequency distribution

6.6 Measures of Kurtosis based on moments

6.7 Simple numerical problems

End of First Term

7. Index Numbers (16)

7.1 Introduction.
7.2 Definition and Meaning.
7.3 Brief discussion on construction of index numbers.
7.4 Simple and weighted price index numbers.
7.5 Laspeyre's, Paasche's and Fisher's Index numbers.
7.6 Simple average of price relative, weighted average of price relative.
7.7 Consumers price index numbers: Considerations in its construction: (i) family budget method (ii) aggregate expenditure method.
7.8 Shifting of base, splicing, deflating and purchasing power.
7.9 Examples and Problems.

8. Theory of Attributes (10)

8.1 Attributes: Classification, Notion of dichotomy and manifold classification, class-frequency, order of class, positive class-frequency, negative classfrequency, ultimate class-frequency, relationship among class-fequencies of different order (upto three attributes), dot operator, Fundamental set of class frequencies.
8.2 Consistency of data upto 3 attributes
8.3 Concepts of independence and association of two attribtes
8.4 Yule's coefficient of association (Q)
8.5 Simple numerical problems.

9. Correlation (8)

9.1 Bivariate data
9.2 Concepts of correlation between two variables, positive correlation, negative correlation
9.3 Scatter diagram, conclusion about the type of correlation from scatter diagram
9.4 Karl 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
9.5 Spearman'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.)
9.6 Simple numerical problems

10. Fitting of curves to the bivariate data (5)

10.1 Fitting of line ($Y = a + b X$),
10.2 Fitting of second degree curve ($Y = a + b X + c X^2$),
10.3 Fitting of exponential curves of the type $Y = a\, b^X$ and $Y = aX^b$.
In all these curves parameters are estimated by the method of least squares.

11. Linear Regression Model (9)

11.1 Meaning of regression, difference between correlation and regression,

11.2 Concept of error in regression, error modeled as a continuous random variable. Simple linear regression model: $Y = a + b X + \varepsilon$, where ε is a continuous random variable with $E(\varepsilon) = 0$, $V(\varepsilon) = \sigma^2$. Estimation of a, b by the method of least squares. Interpretation of parameters. Statement of the estimator of $\sigma^{2.}$

11.3 Concept of residual, plot of residual against X, concept of coefficient of determination.

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. Sarma, 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