# Savitribai Phule Pune University

## (Formerly University of Pune)

## REVISED SYLLABUS OF S. Y. B. A.

## APPLIED STATISTICS (General)

## Choice Based Credit System Syllabus

## With Effect from June 2020

**Preamble:**

Statistics is used in different ways in different contexts. For a cricket fan, Statistics is the information about runs scored or wickets taken by a player. For the manager of a manufacturing unit, Statistics may be the information about the production/manufacturing process. For a medical researcher, investigating the effects of a new drug, Statistics is the evidence of research efforts. For a college student, Statistics shows the grades or marks scored in a course. Thus, in all these examples, Statistics 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. The subject provides tools for making decisions when conditions of uncertainty prevail. Hence Statistical tools and techniques are used in almost all fields such as agriculture, business, management, economics, finance, insurance, education, sports, biotechnology, medical science, etc. For the last two decades, large amount of data has been collected with the help of computers and more sophisticated statistical techniques are needed for the effective analysis and meaningful conclusions from these data. Knowledge of different aspects of Statistics has become crucial in the present technologically advanced scenario. There is a continuous demand for statisticians in fields of education, industry, software and research. The syllabi of three-year B.A. (special) degree course in Statistics are framed in such a way that the students at the end of the course are thorough in basic statistical techniques and are ready to pursue a Master degree and/or simultaneously able to seek jobs involving statistical analysis related to a variety of data sets in order to arrive at some valid conclusions.

**Note**: (1) A student of the Three-Year B.A. Degree Course offering 'Statistics' at the special level must offer 'Mathematical Statistics' as a General level subject in all the three years of the course. Further students of the three-year B.A. Degree Course are advised not to offer 'Statistics' as the General level unless they have offered 'Mathematical Statistics' as a General level subject in all the three years of the course.

(2) A student of three-year B.A. Degree Course offering 'Statistics' will not be allowed to offer 'Applied Statistics' and 'Statistical Pre-requisites' in any of the three years of the course.

(3) A student offering 'Statistics' at the Special level must complete all practicals in Practical Paper to the satisfaction of the teacher concerned.

(4) He/ She must produce the laboratory journal along with the completion certificate signed by the Head of the Department at the time of Practical Examination.

**(5) Structure of evaluation of practical paper at S.Y.B.A. Statistics:**

A) Continuous Internal Assessment (CIA):

| Section | Marks |
|---|---|
| i)  Journal | 20 |
| ii) Viva-voce | 05 |
| iii) Project | 05 |
| Total of CIA | 30 |

B) End of Semester Examination (ESE):

| Section | Nature | Marks | Time |
|---|---|---|---|
| I | On line examination: <br> Note: Question No.1 is compulsory. <br> Q. 1: Execute the commands and write the same in answer book along with answers using: <br> (A) $Ms - EXCEL$ (For Sem-III) <br> (B) $R - Software$  (For Sem-IV) | <br><br><br><br><br>10 <br>10 | Maximum 30 minutes |
| II | Using Calculator / Computer <br> Note: Attempt any two of the four questions (each of 25 marks): <br> Q2,  Q3, Q4 and  Q5. | 50 | *3 hours <br> And <br> ^2hrs 30 |
| III | Viva-voce | 10 | 10 minutes |
| | **Total of B** | **70** | # |

Foot note:
* For calculator user
^ For computer user
# Total examination time 3 hours 40 minutes for calculator user and   3hours 10 minutes for   computer user.

**(6) Structure of evaluation of theory paper at S.Y.B.A. Statistics (special and general), Applied statistics and Statistical pre-requisites:**

A) Continuous Internal Assessment (CIA) for theory subjects:

| Section | Marks |
|---|---|
| i) Theory examination | 20 |
| ii) Home assignment | 05 |
| iii) Seminar/class test etc. | 05 |
| Total of CIA | 30 |

B) End of Semester Examination (ESE) for theory subjects:

| Question | Nature | Marks |
|---|---|---|
| 1 | a) Choose correct alternative:<br>i)   ii)   iii)   iv)   v)<br>(each with four multiple choice answer A, B, C, D)<br>b) True or false:<br>i)   ii)   iii)   iv)   v) | 5<br><br><br><br>5 |
| 2 | Attempt any four of the following:<br>a)   b)   c)   d)   e)   f) | 20 |
| 3 | Attempt any four of the following:<br>a)   b)   c)   d)   e)   f) | 20 |
| 4 | Attempt any two of the following:<br>a)   b)   c)   d) | 20 |
| | **Total ESE** | **70\*** |

Foot note:
 * Numerical problem should not exceed 40% of total marks with option questions.

# S.Y.B.A. STATISTICS Syllabus
## For Choice Based Credit System-2019 pattern
## To be implemented from the Academic year 2020-2021
## Structure of the course:

Table with code and Title for CBCS 2019 pattern:

| Subject | Semester III | | | Semester IV | | |
|---|---|---|---|---|---|---|
| | Code | Title | Credit | Code | Title | Credit |
| Statistics General-I | ST-23873 | Sampling Techniques(**CC1C**) | 3 | ST-23874 | Statistical Quality Control (**CC 1D**) | 3 |
| Statistics Special-I | ST-23883 | Continuous Prob. Distributions (**DSE 1A**) | 3 | ST-23884 | Sampling Distrib$^n$ and Inference (**DSE 1B**) | 3 |
| Statistics Special-II | ST-23893 | Statistics Practical (**DSE 2A**) | 3 | ST-23894 | Statistics Practical (**DSE 2B**) | 3 |
| Mathematical Statistics General-II | ST-23273 | Discrete Prob. Distributions and Time Series (**CC 2C**) | 3 | ST-23274 | Tests of Significance and Statistical Methods (**CC 2D**) | 3 |
| Skill Enhancement Course(**SEC**) | | **General subject other than Statistics (SEC 1A)** | 3 | | **General subject other than Statistics (SEC 1B)** | 3 |
| Skill Enhancement Course(**SEC**) | | Data Handling Through MS-Excel (**SEC 2A**) | 2 | | Data Handling Through R-software(**SEC2B**) | 2 |
| Statistical Prerequisites | ST-23573 | Applications of Statistics and Theory of Prob. (**SEC1A**) | 3 | ST-23574 | Theory of Prob. Distributions (**SEC 1B**) | 3 |
| Applied Statistics | ST-24173 | Applications of Statistics and Theory of Prob. (**SEC 1A**) | 3 | ST-24174 | Discrete Prob. Distributions and Demography (**SEC 1B**) | 3 |

Note: (1) Applied Statistics can be offered only as a General level subject.

(2) A student of Three-Year B.A. Degree course offering Applied Statistics will not be allowed to offer 'Mathematical Statistics', 'Statistical Pre-requisites' and / or 'Statistics' simultaneously in any of the three years of the course.

**SYBA**
**SEMESTER -III**
**Subject: Applied Statistics (General)**

**ST-24173: Applications of Statistics and Theory of Probability (SEC 1A)**

**Unit 1**. **Multiple regression plane, multiple and partial correlation coefficient**

**(using tri-variate data):** (10)

(No proof is required for the derivation of the equation of multiple regression plane)

1.1 Notion of multiple regression.

1.2 Equation of Multiple regression model: $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$

1.3 Obtaining the equation of plane of regression given:

i) the means, standard deviations and total correlation coefficients and obtaining the estimated value.

ii) sums, sums of squares and sum of squares of deviations from respective mean etc. and obtaining the estimated value.

1.4 Notion of multiple and partial correlation.

1.5 Definition of multiple correlation coefficient ( $R_{Y.X1X2}$ ) and partial correlation coefficient

( $r_{YX1.X2}$ and $r_{YX2.X1}$ )

1.6 Concept of multiple coefficient of determination and partial coefficient of determination with interpretation.

**Unit 2. Time Series:** (16)

2.1 Meaning of Time Series

2.2 Various components of a time series (Explanation and illustrations of each component)

2.3 Additive and multiplicative models of time series.

2.4 Meaning and usefulness of time series analysis.

2.4 Methods of estimating trends: (i) Freehand or graphical method (ii) Method of least square (line and second degree curve) (iii) Method of semi-averages (iv) Method of moving averages.

2.5 Methods of estimating seasonal components:

(i) Methods of simple averages. (ii) Ratio to trend obtained by moving averages.

2.6 Auto regression model, Fitting of AR(1)model.

## Unit 3. Permutations and Combinations: (6)

3.1 Definitions of permutation and combination for distinct objects.

3.2 Relation between permutation and combination. (Explain through illustrations)

(i) $\binom{n}{r} = \binom{n}{n-r}$    (ii) $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$

## Unit 4. Probability: (16)

**4.1 Revision of set theory. Concept and definition of union, intersection of two sets, complement of a set. (no problem should be asked on this topic in the examination)**

4.2 Concept of random experiment, sample space with its types.

4.3 Event and types of event: complementary event, elementary event, certain event, impossible event, mutually exclusive events and exhaustive events.

4.4 Classical definition of probability and its limitations.

4.5 Probability model.

4.6 Axioms of probability.

4.7 Theorems of Probability **(Without proof explain only through illustrations)**

(i) $P(A) + P(A') = 1$. (ii) $0 \leq P(A) \leq 1$. (iii) $P(\Phi)=0$. (iv) If $A \subset B$ then $P(A) \leq P(B)$.

(v) Total probability theorem/Addition theorem on two events

$P(A \cup B)=P(A)+P(B)-P(A \cap B)$. (vi) Statement for 3 events for total probability theorem

(vii) $P(A \cup B) \leq P(A)+P(B)$ ,(Boole's inequality).

4.8 Definition of conditional probability.

4.9 Particular cases of $P(A \mid B)$ when $A \subset B$ or $B \subset A$ or $A \cap B=\Phi$.

4.10 Multiplication theorem on $P(A \cap B)$.

4.11 Concept and definition of independence of two events.

4.12 Pair-wise independence and complete independence in case of three events.

**References:**

1. G. Gupta and D. Gupta: Fundamental of Statistics, Vol.II, Shripati Bhattachrjee for the World Press Pvt. Ltd, Calcutta.

2. J.V.Uspensky: Introduction to Mathematical Probability.

3. Lipschutz : Probability and Statistics, Schaum's Outline Series,New York.

4. M.G.Kendall and Stuarr : Advanced theory of Statistics, Vol. I, Allan.

5. M. Siegleman : Introduction to Demography .

6. S.C.Gupta and V.K.Kapoor: Fundamentals of Applied Statistics, Sultan Chand and Sons, New Delhi.

7. Walpole and Myres: Probability and Statistics, Mcmillan Publishing Co. New York.

8. Introduction to Linear Regression Analysis, Douglas C. Montgomery, Elizabeth A. Peck, G.Geoffrey Vining, Wiley

9. Time Series Methods, Brockwell and Davis, Springer, 2006.

10. Time Series Analysis,4th Edition, Box and Jenkin, Wiley, 2008.

11. A First course in Probability, Sheldon Ross.Pearson Education Inkc.

12. Statistical Methods (An IntroductoryText), Medhi J., New Age International.

13. Modern Elementary Statistics, Freund J.E. 2005, Pearson Publication.

14. Fundamentals of Mathematical Statistics (3rd Edition), Gupta S. C. and Kapoor V. K.1987 S. Chand and Sons, New Delhi.

15. Mathematical Statistics (3rd Edition), Mukhopadhyay P. 2015, Books and Allied (P), Ltd.

16. Programmed Statistics, B.L. Agarwal, New Age International Publishers

**Subject: Applied Statistics(General)**

**ST-24174: Discrete Probability Distributions and Demography (SEC 1B)**

**Unit 1. Univariate discrete probability distributions:** (8)

1.1 Definition of a discrete random variable (r.v.)

1.2 Definition of probability mass function (p.m.f.) of a discrete r.v, cumulative distribution function(c.d.f.) and its properties.

1.3 Definition of expectation of a discrete r.v. and expectation of a linear function of discrete r.v. Mode and median of discrete r.v.

1.4 Definition of variance of discrete r.v.

**Unit 2. Bivariate Probability Distributions:** (15)

2.1 Why bivariate rv? Definition of two-dimentional discrete r.v. ,its joint probability distribution /p.m.f.

2.2 Computation of probabilities of events in bivariate probability distributions.

2.3 Concepts of marginal and conditional probability distributions.

2.4 Independence of two discrete r.vs.

2.5 Definition of mathematical expectation of function of two-dimensional discrete r.v.

2.6 Definitions of conditional mean and conditional variance.

2.7 Definition of covariance and correlation coefficient ($\rho$).

2.8 Effect of change of origin and scale on covariance and coefficient of correlation (only Statement).

2.9 Variance of linear combination of two-dimensional discrete r.v ,Var (aX+bY (only Statement).

**Unit 3. Special Discrete Probability Distributions:** (15)

3.1 Discrete uniform distribution: p.m.f. mean and variance. Illustrations of real life situations where this distribution can be applied.

3.2 Binomial distribution: Notation $X \sim B(n,p)$. p.m.f., mean and variance, additive property (derivations excluded). Illustrations of real life situations where the distribution can be applied. Computation of probabilities of events related to binomial r.v.

3.3 Poisson distribution: Notation X~P(m) p.m.f., mean and variance, additive property (derivations excluded), Illustrations of real life situations where the distribution can be applied. Number of vacancies that arises in US Supreme Court, Computation of probabilities of events related to a Poisson r.v.

**Unit 4. Elements of Demography:** (10)

4.1 Introduction to demography, need of vital statistics and methods of obtaining vital statistics.

4.2 Mortality Rates: Crude Death Rate (CDR), Age-specific death rates(ASDR), Standardized Death Rate(STDR) ( Direct and indirect method).

4.3 Fertility and Reproduction Rates: Crude Birth Rate (CBR), General Fertility rate(GFR), Age-specific Fertility Rate(ASFR). Total Fertility Rate(TFR), Gross Reproduction rate(GRR), Net Reproduction Rate(NRR).

**References:**

1. G. Gupta and D. Gupta: Fundamental of Statistics, Vol.II , Shripati Bhattachrjee for the World Press Pvt. Ltd, Calcutta.

2. J.V.Uspensky: Introduction to Mathematical Probability.

3. Lipschutz : Probability and Statistics, Schaum's Outline Series,New York

4. M.G.Kendall and Stuart : Advanced theory of Statistics, Vol. I, Allan.

5. M. Siegleman : Introduction to Demography .

6. S.C.Gupta and V.K.Kapoor: Fundamentals of Applied Statistics, Sultan Chand and Sons,New Delhi.

7. Walpole and Myres : Probability and Statistics , Mcmillan Publishing Co. New York.