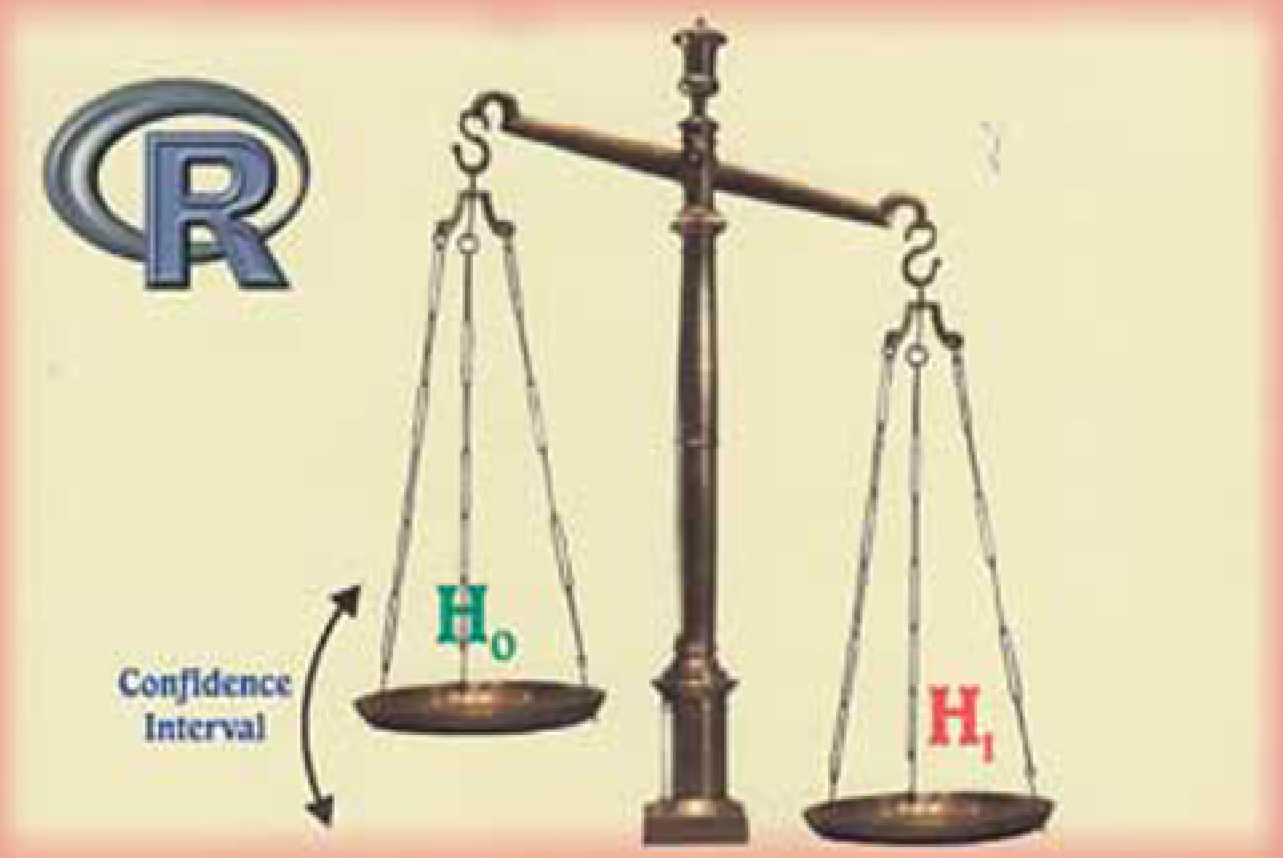
NEW SYLLABUS

A BOOK OF

SAMPLING DISTRIBUTIONS AND INFERENCE

S. Y. B. Sc. STATISTICS (ST 222) • PAPER II - SEMESTER II



Prof. P. G. DIXIT

Prof. P. S. KAPRE

Prof. V. R. PAWGI

NIRALI PRAKASHAN

A Book Of

SAMPLING DISTRIBUTION AND INFERENCE
STATISTICS

Paper – II
For
S.Y.B.Sc. (ST-222)
As Per Savitribai Phule Pune University's Revised Syllabus
Effective from June 2014

Prof. P. G. DIXIT
M.Sc., M.Phil. (Statistics)
Head of Statistics Department,
Modern College, Pune - 5.

Prof. P. S. KAPRE
M.Sc., M.Phil. (Statistics)
Department of Statistics,
Modern College, Pune - 5.

Prof. V. R. PAWGI
M. Sc., M. Phil., P.G.D.C.A.
Head, Department of Statistics,
Abasaheb Garware College,
Pune – 4.

Price ₹ 140.00

NIRALI PRAKASHAN
ADVANCEMENT OF KNOWLEDGE

N1430

S.Y.B.Sc. STATISTICS P – II (S-II)

Second Edition : January 2017 **ISBN 978-93-5164-283-1**

© : Authors

The text of this publication, or any part thereof, should not be reproduced or transmitted in any form or stored in any computer storage system or device for distribution including photocopy, recording, taping or information retrieval system or reproduced on any disc, tape, perforated media or other information storage device etc., without the written permission of Authors with whom the rights are reserved. Breach of this condition is liable for legal action.

Every effort has been made to avoid errors or omissions in this publication. In spite of this, errors may have crept in. Any mistake, error or discrepancy so noted and shall be brought to our notice shall be taken care of in the next edition. It is notified that neither the publisher nor the authors or seller shall be responsible for any damage or loss of action to any one, of any kind, in any manner, therefrom.

Published By :
NIRALI PRAKASHAN
Abhyudaya Pragati, 1312, Shivaji Nagar,
Off J.M. Road, PUNE – 411005
Tel - (020) 25512336/37/39, Fax - (020) 25511379
Email : niralipune@pragationline.com

Polyplate

Printed By :
YOGIRAJ PRINTERS AND BINDERS
Works: Sr. No. 10\1,Ghule Industrial Estate,
Nanded Village Road,
Tal-Haveli, Dist-Pune 411041.
Mobile - 9850046517, 9404233041

☞ DISTRIBUTION CENTRES

PUNE

Nirali Prakashan : 119, Budhwar Peth, Jogeshwari Mandir Lane, Pune 411002, Maharashtra
Tel : (020) 2445 2044, 66022708, Fax : (020) 2445 1538
Email : bookorder@pragationline.com, niralilocal@pragationline.com

Nirali Prakashan : S. No. 28/27, Dhyari, Near Pari Company, Pune 411041
Tel : (020) 24690204 Fax : (020) 24690316
Email : dhyari@pragationline.com, bookorder@pragationline.com

MUMBAI

Nirali Prakashan : 385, S.V.P. Road, Rasdhara Co-op. Hsg. Society Ltd.,
Girgaum, Mumbai 400004, Maharashtra
Tel : (022) 2385 6339 / 2386 9976, Fax : (022) 2386 9976
Email : niralimumbai@pragationline.com

☞ DISTRIBUTION BRANCHES

JALGAON

Nirali Prakashan : 34, V. V. Golani Market, Navi Peth, Jalgaon 425001,
Maharashtra, Tel : (0257) 222 0395, Mob : 94234 91860

KOLHAPUR

Nirali Prakashan : New Mahadvar Road, Kedar Plaza, 1st Floor Opp. IDBI Bank
Kolhapur 416 012, Maharashtra. Mob : 9850046155

NAGPUR

Pratibha Book Distributors : Above Maratha Mandir, Shop No. 3, First Floor,
Rani Jhanshi Square, Sitabuldi, Nagpur 440012, Maharashtra
Tel : (0712) 254 7129

DELHI

Nirali Prakashan : 4593/21, Basement, Aggarwal Lane 15, Ansari Road, Daryaganj
Near Times of India Building, New Delhi 110002, Mob : 08505972553

BENGALURU

Pragati Book House : House No. 1, Sanjeevappa Lane, Avenue Road Cross,
Opp. Rice Church, Bengaluru – 560002.
Tel : (080) 64513344, 64513355,Mob : 9880582331, 9845021552
Email:bharatsavla@yahoo.com

CHENNAI

Pragati Books : 9/1, Montieth Road, Behind Taas Mahal, Egmore,
Chennai 600008 Tamil Nadu, Tel : (044) 6518 3535,
Mob : 94440 01782 / 98450 21552 / 98805 82331,
Email : bharatsavla@yahoo.com

niralipune@pragationline.com | www.pragationline.com

Also find us on f www.facebook.com/niralibooks

Note : Every possible effort has been made to avoid errors or omissions in this book. In spite this, errors may have crept in. Any type of error or mistake so noted, and shall be brought to our notice, shall be taken care of in the next edition. It is notified that neither the publisher, nor the author or book seller shall be responsible for any damage or loss of action to any one of any kind, in any manner, therefrom. The reader must cross check all the facts and contents with original Government notification or publications.

Statistical Thinking will one day be necessary for effective citizenship as the ability to read and write

– H.G. Wells

Preface ...

We feel indeed very happy to present this text-book of 'Statistics Paper-II' **Sampling Distributions and Inference (ST-222)** to the students of S.Y.B.Sc. The book is written according to the revised syllabus of Savitribai Phule, Pune University with effect from June, 2014.

The main purpose of the book is to provide foundation as well as comprehensive background of Probability Theory and statistical methods to beginners in simple and interesting manner. In order to make the contents of the book easier to comprehend, we have included a requisite number of illustrations, remarks, figures, diagrams etc. to elucidate statistical concepts. Application of Statistics in real life situations is emphasized through illustrative examples. Ample number of graded problems, theoretical as well as numerical are provided at the end of each chapter along with hints and answers. The numerical problems will also be useful for the F.Y.B.Sc. students computer science to prepare for Paper – III : Practicals. A list of practicals is given in the syllabus. Appendix A compiles some important mathematical results which are needed during the entire course. Values of individual terms of binomial probabilities are given in Appendix B. A specimen paper is set for student's self assessment. We have included MS-EXCEL and R-software in finding probabilities and tests of hypotheses. It is an additional feature of the book.

This book will also serve the purpose of reference book for M.B.A., C.A., M.P.M., classes.

We are thankful to Mr. D. K. Furia and the staff of Nirali Prakashan for bringing out this book in short time. Mrs. Anagha Medhekar and Mr. Santosh Bare deserve special thanks for the work done with atmost care and sincerely. Finally, our families deserve special thanks for their support, encouragement and tolerance.

We request our colleagues, teaching Statistics to offer their criticism and suggestions, for further improvement of the book.

– Authors

Kartiki Ekadashi

Syllabus ... ST-222 : Sampling Distributions and Inference

1 Chi-square (χ_n^2) Distribution **(10 L)**

1.1 Definition χ^2 r.v. as sum of squares of i.i.d. standard normal variables, derivation of p.d.f. of χ^2 with n degrees of freedom (d.f.) using M.G.F., nature of p.d.f. curve, computations of probabilities using tables of χ^2 distribution. Mean, variance, M.G.F., C.G.F., central moments, β_1, β_2, γ_1, γ_2, mode, additive property.

1.2 Normal approximation : $\frac{\chi_n^2 - n}{\sqrt{2n}}$ with proof.

1.3 Distribution of $\frac{X}{X + Y}$ and $\frac{X}{Y}$, where X and Y are two independent chi-square random variables.

2. Student's t-distribution **(06 L)**

2.1 Definition of T r.v. with n d.f. in the form $T = \frac{U}{\sqrt{\chi_n^2/n}}$, where $U \to N(0, 1)$ and χ_n^2 is a χ^2 r.v. with n d.f. and U and χ_n^2 are independent r.v.s.

2.2 Derivation of p.d.f., nature of probability curve, mean, variance, moments, mode, use of tables of t-distribution for calculation of probabilities, statement of normal approximation.

3. Snedecore's F-distribution **(06 L)**

3.1 Definition of F r.v. with n_1 and n_2 d.f. as $F_{n_1, n_2} = \frac{\chi_{n_1}^2/n_1}{\chi_{n_2}^2/n_2}$ where $\chi_{n_1}^2$ and $\chi_{n_2}^2$ are independent chi-square r.v.s with n_1 and n_2 d.f. respectively.

3.2 Derivation of p.d.f., nature of probability curve, mean, variance, moments, mode.

3.3 Distribution of $1/F_{n_1, n_2}$, use of tables of F-distribution of calculation of probabilities.

3.4 Interrelations among, χ^2, t and F variates.

4. Sampling Distributions **(08 L)**

4.1 Random sample from a distribution as i.i.d. r.v.s. $X_1, X_2, \ldots, X_n$.

4.2 Notion of a statistic as function of $X_1, X_2, \ldots, X_n$ with illustrations.

4.3 Sampling distribution of a statistic. Distribution of sample mean $\overline{X}$ from normal, exponential and gamma distribution, Notion of a standard error of a statistic.

4.4 Distribution of $\frac{nS^2}{\sigma_2} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \overline{X})_2$ for a sample from a normal distribution using orthogonal transformation. Independence of $\overline{X}$ and S^2.

5. Exact Tests **(18 L)**

5.1 Tests based on chi-square distribution :

(a) Test for independence of two attributes arranged in 2×2 contingency table. (With Yates' correction).

(b) Test for independence of two attributes arranged in $r \times s$ contingency table, McNemar's test.

(c) Test for 'Goodness of Fit'. (Without rounding-off the expected frequencies).

(d) Test for $H_0 : \sigma^2 = \sigma_0^2$ against one-sided and two-sided alternatives when (i) mean is known, (ii) mean is unknown.

5.2 Tests based on t-distribution :

(a) t-tests for population means : (i) one sample and two sample tests for one-sided and two-sided alternatives, (ii) $100 (1 - \alpha)$ % two sided confidence interval for population mean (μ) and difference of means ($\mu_1 - \mu_2$) of two independent normal population.

(b) Paired t-test for one-sided and two-sided alternatives.

5.3 Tests based on F-distribution :

(a) Tests for $H_0 : \sigma_1^2 = \sigma_2^2$ against one-sided and two-sided alternatives when (i) means are known, (ii) means are unknown.

●●●

LIST OF PRACTICE

Sr. No.	Title of the experiment	No. of practicals
1.	Fitting of negative binomial distribution, testing goodness of fit.	**1**
2.	Fitting of normal distribution, testing goodness of fit (also using qq-plot)	**1**
3.	Applications of normal, negative binomial and multinomial distribution	**2**
4.	Model sampling from exponential, normal distribution using (i) distribution function, (ii) Box- Muller transformation.	**1**
5.	Time series : Estimation and forecasting of trend by fitting of AR (1) model, exponential smoothing, moving averages.	**1**
6.	Estimation of seasonal indices by ratio to trend.	**1**
7.	Test for means and construction of confidence interval. (Also using MS-EXCEL) (i) $H_0 : \mu = \mu_0$, σ^2 known and σ^2 unknown (ii) $H_0 : \mu_1 = \mu_2$, σ_1, σ_2 known (iii) $H_0 : \mu_1 = \mu_0$, $\sigma_1 = \sigma_2 = \sigma$ unknown (iv) $H_0 : \mu_1 = \mu_2$, paired t test	**2**
8.	Tests for proportions and construction of confidence interval for $P = P_1 – P_2$. $H_0 : P = P_0$, $H_0 : P_1 = P_2$	**1**
9.	Tests based on χ^2 distribution (i) Goodness of fit. (ii) Independence of attributes (2×2, $m \times n$ contingency table), (iii) Mc Nemar's test. (iv) $H_0 : \sigma^2 = \sigma_0^2$, μ unknown, confidence interval for σ^2	**2**
10.	Tests based on F-distribution $H_0 : \sigma_1^2 = \sigma_2^2$. (i) means known, (ii) means unknown.	**1**
11.	Fitting of multiple regression plane using MS-EXCEL.	**1**
12.	Fitting of normal distribution using MS-EXCEL.	**1**
13.	Exponential smoothing using MS-EXCEL.	**1**
14.	Computations of probabilities of Normal, Exponential gamma χ^2, t, F using R.	**1**
15.	Use of basic R software commands, finding summary statistics using R software	**1**
16.	Tests using R software	**1**
17.	**Project :** Project based on analysis of data collected by students in groups of maximum 6 students. (Project is equivalent to five practicals)	**5**

•••

Contents ...

1.	**Chi-Square (χ^2) Distribution**	**1.1 – 1.30**
2.	**Student's t-Distribution**	**2.1 – 2.18**
3.	**Snedecore's F-Distribution**	**3.1 – 3.22**
4.	**Sampling Distributions**	**4.1 – 4.16**
5.	**Exact Test**	**5.1 – 5.66**
	Appendix	**A.1 – A.6**
	Specimen Question Paper	**S.1 – S.2**
	University Question Papers	**P.1 – P.4**

●●●

Chapter 1...

Chi-Square (χ^2) Distribution

Yates

Born on May 12 1902 in Manchester England he was the eldest of five children. He began his mathematic education at a private school were he was encouraged and tought by an advanced mathematics teacher. Yates recieved a scholarship to Clifton College and was later awarded another scholarship at St. Johns College in Cambridge. He taught mathematics for a short period of time but wanted to use mathematics to help others instead of just teaching it. A few years later he went to Africa where he was part of the gold coast survey. A few years later he moved back and was appointed assistant statistician at Rothamsted Experimental Station. When a friend, R.A Fisher became chairman at the University College London in 1933, Yates was appointed the Head of Statistics at Rothamsted. He held that postition until 1968 when he retired. During the time he was working, he worked often with Fisher on experimental design, block designs, theory of variance and operational research.

Contents ...

1.0 Introduction

1.1 Derivation of P.D.F. of χ^2 with n Degrees of Freedom

1.2 Moments of Chi-square Distribution

1.3 Recurrence Relation among the Central Moments

1.4 Additive Property

1.5 Mode of the Chi-square Distribution

1.6 Use of χ^2 Tables for Calculation of Probabilities

1.7 Limiting behaviour of χ^2 as $n \to \infty$. [Normal Approximation]

1.8 Results Based on Bivariate Transformation

Key Words :

Gamma distribution, chi-square distribution, degrees of freedom (d.f.), normal approximation.

Objectives :

(1) To understand chi-square distribution as a particular case of gamma distribution.

(2) To study the relationship between chi-square distribution and other distributions. Also limiting behaviour of chi-square distribution as $n \to \infty$ (normal approximation).

(3) Computation of probabilities using χ^2 tables and also using MS-EXCEL, R-software.

(4) To study the properties of χ^2 distribution.

1.0 Introduction

The gamma distribution is related with the theory of independent normal variates in natural way. Because we have shown that if $X_1, X_2, \ldots X_i, \ldots, X_n$ are n independent standard normal variates then the distribution of $\sum_{i=1}^{n} X_i^2$ has $G\left(\frac{1}{2}, \frac{n}{2}\right)$ distribution. This particular case of gamma distribution is called as *chi-square* distribution with n degrees of freedom (d.f.). The random variable involved in this distribution is denoted as χ^2_n. The degrees of freedom (d.f.) represents the number of independent variables used in the construction of χ^2. Thus χ^2_n variate can be looked-upon as the sum of squares of n independent standard normal variates.

Tests based on chi-square distribution like chi-square test of goodness of fit, chi-square test of independence of attributes etc. are widely used. It also plays an important role in statistical inference.

1.1 Derivation of P.D.F. of χ^2 with n Degrees of Freedom (using M.G.F.)

Let $X_1, X_2, \ldots X_i, \ldots, X_n$ be i.i.d. N (0, 1) variates. Then m.g.f. of X_i^2 is given by,

$$M_{X^2}(t) = \int_{-\infty}^{\infty} e^{tx^2} f(x)\, dx \text{ where, } f(x) \text{ is p.d.f. of standard normal distribution.}$$

$$\therefore \quad M_{X^2}(t) = \int_{-\infty}^{\infty} e^{tx^2} \cdot \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}x^2}\, dx$$

$$= \frac{2}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-\left(\frac{1}{2} - t\right)x^2}\, dx \quad \left[\because \text{Integrand is even function of x, Also } a = \frac{1}{2} - t, b = 1\right]$$

$$= \frac{2}{\sqrt{2\pi}} \cdot \frac{1}{2} \frac{\Gamma\left(\frac{1}{2}\right)}{\left(\frac{1}{2} - t\right)^{1/2}} \quad ; \quad \frac{1}{2} - t > 0$$

$$= \frac{2}{\sqrt{2\pi}} \cdot \frac{1}{2} \frac{\Gamma\left(\frac{1}{2}\right)}{(1-2t)^{1/2}} \; 2^{1/2} = \left(\frac{1}{1-2t}\right)^{1/2}$$

$$= (1-2t)^{-1/2}$$

$$\therefore \qquad M_{X^2}(t) = \left(1 - \frac{t}{1/2}\right)^{\frac{-1}{2}} \qquad ; \qquad t < \frac{1}{2}$$

$\because$ If X_i is N (0, 1) variate $M_i^2(t) = \left(1 - \frac{t}{1/2}\right)^{-\frac{1}{2}}$; $t < \frac{1}{2}$

Suppose, $Y = \sum_{i=1}^{n} X_i^2$. Then Y has χ^2 distribution with n d.f. $\left(\chi_n^2\right)$. The m.g.f. of Y is as follows :

$$M_Y(t) = M_{\sum_{i=1}^{n} X_i^2}(t)$$

$$= M_{X_1^2}(t) . M_{X_2^2}(t) \ldots M_{X_i^2}(t) \ldots M_{X_n^2}(t)$$

$$= \left(1 - \frac{t}{1/2}\right)^{\frac{-n}{2}} \qquad \left[\because X_i\text{'s are independent}\right]$$

It is m.g.f. of $G\left(\frac{1}{2}, \frac{n}{2}\right)$. Thus $\sum_{i=1}^{n} X_i^2 = Y$ has $G\left(\frac{1}{2}, \frac{n}{2}\right)$. Hence p.d.f. Y is given as,

$$f(y) = \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} e^{\frac{-y}{2}} y^{\frac{n}{2}-1} \qquad ; \qquad y \geq 0$$

$$= 0 \qquad ; \qquad \text{otherwise}$$

OR

$$f(y) = \frac{1}{\left(\sqrt{2}\right)^n \Gamma\left(\frac{n}{2}\right)} e^{\frac{-y}{2}} y^{\frac{n}{2}-1} \qquad ; \qquad y \geq 0 \qquad \ldots (1)$$

Thus a continuous variable Y is said to follow chi-square distribution with n d.f. if its p.d.f. is given by relation (1).

Another method :

Let $X_1, X_2, \ldots, X_i \ldots X_n$ be i.i.d. N(0, 1) variates.

Suppose $Y_i = X_i^2$ we shall obtain distribution function of Y. It is given by

$$G_Y(y) = P[Y \leq y] = P[X_i^2 \leq y]$$

$$= P(-\sqrt{y} \leq X_i \leq \sqrt{y})$$

$$= F(\sqrt{y}) - F(-\sqrt{y}) \qquad [F(\cdot) \text{ is distribution function of } X_i]$$

In order to obtain p.d.f. of y, we take derivatives of both sides w.r.t. y,

$$g(y) = \frac{d\,G(y)}{dy} = \frac{d}{dy}[F(\sqrt{y}) - f(-\sqrt{y})]$$

$$= \frac{f(\sqrt{y})}{2\sqrt{y}} - \frac{f(-\sqrt{y})}{2\sqrt{y}}(-1)$$

$$= \frac{f(\sqrt{y}) + f(-\sqrt{y})}{2\sqrt{y}} \qquad [\because X \to N(0, 1), f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}]$$

$$= \frac{\frac{1}{\sqrt{2\pi}} e^{-y/2} + \frac{1}{\sqrt{2\pi}} e^{-y/2}}{2\sqrt{y}}$$

$$\therefore \qquad g(y) = \frac{\frac{1}{\sqrt{2\pi}} e^{-y/2}}{\sqrt{y}} \quad ; \quad y > 0$$

$$= \frac{\left(\frac{1}{2}\right)^{1/2}}{\Gamma\left(\frac{1}{2}\right)} e^{-y/2} y^{\frac{1}{2}-1} \; ; \; y > 0 \qquad [\because \Gamma\frac{1}{2} = \sqrt{\pi}]$$

which p.d.f. of gamma distribution with parameters $\left(\frac{1}{2}, \frac{1}{2}\right)$. Thus X_i^2 follows $G\left(\frac{1}{2}, \frac{1}{2}\right)$ distribution i= 1, 2 … n.

$\because$ Each X_i^2 follows $G\left(\frac{1}{2}, \frac{1}{2}\right)$ distribution. Using additive property of gamma variates.

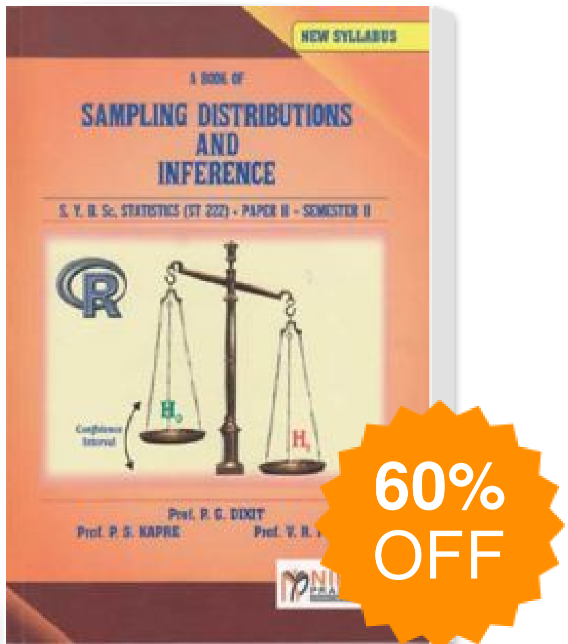
$T = \sum_{i=1}^{n} X_i^2$ follows $G\left(\frac{1}{2}, \frac{n}{2}\right)$ distribution. The p.d.f. of T which has χ_n^2 probability distribution is given by

$$f(t) = \frac{1}{(\sqrt{2})^n \Gamma\left(\frac{n}{2}\right)} e^{-t/2} t^{n/2-1} \; ; \; t \geq 0$$

Remarks :

1. The probability density curves for n = 1, 2, 3, 4, 5, and 6 are as shown in figure 1.1.

Sampling Distribution And Inference Statistics Paper – II



Publisher : Nirali Prakashan

ISBN : 9789351642831

Author : Prof. P. G. Dixit, Prof. P. S. Kapre, Prof. V. R. Pawgi

Type the URL : http://www.kopykitab.com/product/21710

Get this eBook