STATISTICS: textbooks and monographs volume 40

HANDBOOK OF THE NORMAL DISTRIBUTION

JAGDISH K. PATEL CAMPBELL B. READ

Handbook of THE NORMAL DISTRIBUTION

Jagdish K. Patel • Campbell B. Read

Department of Mathematics and Statistics
University of Missouri-Rolla
Rolla, Missouri

Department of Statistics
Southern Methodist University
Dallas, Texas

MARCEL DEKKER, INC. New York and Basel

Library of Congress Cataloging in Publication Data

Patel, Jagdish K.
Handbook of the normal distribution.

(Statistics, textbooks and monographs; v. 40)
Includes bibliographical references and index.
1. Gaussian distribution. I. Read, Campbell B.
II. Title. III. Series.
QA273.6.P373 519.2 81-17422
ISBN 0-8247-1541-1 AACR2

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

Handbook of THE NORMAL DISTRIBUTION

STATISTICS: Textbooks and Monographs

A SERIES EDITED BY

D. B. OWEN, Coordinating Editor
Department of Statistics
Southern Methodist University
Dallas, Texas

Volume 22: Pocket Book of Statistical Tables, *compiled by R. E. Odeh, D. B. Owen, Z. W. Birnbaum, and L. Fisher*

Volume 23: The Information in Contingency Tables, *D. V. Gokhale and Solomon Kullback*

Volume 24: Statistical Analysis of Reliability and Life-Testing Models: Theory and Methods, *Lee J. Bain*

Volume 25: Elementary Statistical Quality Control, *Irving W. Burr*

Volume 26: An Introduction to Probability and Statistics Using BASIC, *Richard A. Groeneveld*

Volume 27: Basic Applied Statistics, *B. L. Raktoe and J. J. Hubert*

Volume 28: A Primer in Probability, *Kathleen Subrahmaniam*

Volume 29: Random Processes: A First Look, *R. Syski*

Volume 30: Regression Methods: A Tool for Data Analysis, *Rudolf J. Freund and Paul D. Minton*

Volume 31: Randomization Tests, *Eugene S. Edgington*

Volume 32: Tables for Normal Tolerance Limits, Sampling Plans, and Screening, *Robert E. Odeh and D. B. Owen*

Volume 33: Statistical Computing, *William J. Kennedy, Jr. and James E. Gentle*

Volume 34: Regression Analysis and Its Application: A Data-Oriented Approach, *Richard F. Gunst and Robert L. Mason*

Volume 35: Scientific Strategies to Save Your Life, *I. D. J. Bross*

Volume 36: Statistics in the Pharmaceutical Industry, *edited by C. Ralph Buncher and Jia-Yeong Tsay*

Volume 37: Sampling from a Finite Population, *J. Hájek*

Volume 38: Statistical Modeling Techniques, *S. S. Shapiro*

Volume 39: Statistical Theory and Inference in Research, *T. A. Bancroft and C.-P. Han*

Volume 40: Handbook of the Normal Distribution, *Jagdish K. Patel and Campbell B. Read*

Volume 41: Recent Advances in Regression Methods, *Hrishikesh D. Vinod and Aman Ullah*

OTHER VOLUMES IN PREPARATION

To my wife, Prabha, and children, Sanjay and Sonal--J.K.P.

To my mother and father--C.B.R.

PREFACE

This book contains a collection of results relating to the normal distribution. It is a compendium of properties, and problems of analysis and proof are not covered. The aim of the authors has been to list results which will be useful to theoretical and applied researchers in statistics as well as to students.

Distributional properties are emphasized, both for the normal law itself and for statistics based on samples from normal populations. The book covers the early historical development of the normal law (Chapter 1); basic distributional properties, including references to tables and to algorithms suitable for computers (Chapters 2 and 3); properties of sampling distributions, including order statistics (Chapters 5 and 8), Wiener and Gaussian processes (Chapter 9); and the bivariate normal distribution (Chapter 10). Chapters 4 and 6 cover characterizations of the normal law and central limit theorems, respectively; these chapters may be more useful to theoretical statisticians. A collection of results showing how other distributions may be approximated by the normal law completes the coverage of the book (Chapter 7).

Several important subjects are not covered. There are no tables of distributions in this book, because excellent tables are available elsewhere; these are listed, however, with the accuracy and coverage in the sources. The multivariate normal distribution other than the bivariate case is not discussed; the general linear model and regression models based on normality have been amply documented elsewhere; and the applications of normality in the methodology of statistical inference and decision theory would provide material for another volume on their own.

In citing references, the authors have tried to balance the aim of giving historical credit where it is due with the desirability of citing easily obtainable sources which may be consulted for further detail. In the latter case, we do not aim to cite every such work, but only enough to give the researcher or student a readily available source to which to turn.

We would like to thank the following persons for reviewing parts of the manuscript and giving helpful suggestions: Lee J. Bain, Herbert A. David, Maxwell E. Engelhardt, C. H. Kapadia, C. G. Khatri, Samuel Kotz, Lloyd S. Nelson, Donald B. Owen--who also gave editorial guidance, Stephen M. Stigler, Farroll T. Wright, and a referee. For assistance in typing the manuscript and for their infinite patience, we thank Connie Brewster, Sheila Crain, Millie Manley, and Dee Patterson; we would like to thank Dr. Maurits Dekker and the staff at MDI for their work in taking the manuscript through production. We would also like to thank Southern Methodist University for giving one of us leave for a semester in order to do research for the manuscript.

J.K.P. C.B.R.

CONTENTS

chapter 1

GENESIS: A HISTORICAL BACKGROUND

> I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by the "Law of Frequency of Error." The law would have been personified by the Greeks and deified, if they had known of it. It reigns with serenity and in complete self-effacement amidst the wildest confusion.

So wrote Sir Francis Galton (1889, p. 66) about the normal distribution, in an age when the pursuit of science was tinged with the romanticism of the nineteenth century. In this age of computers, it is hard to find enthusiasm expressed with the sense of wonder of these men of letters, so much do we take for granted from modern technology.

In the seventeenth century Galileo (trans. 1953; 1962, pp. 303-309) expressed his conclusions regarding the measurement of distances to the stars by astronomers (Maistrov, 1974, pp. 31-34). He reasoned that random errors are inevitable in instrumental observations, that small errors are more likely to occur than large ones, that measurements are equally prone to err in one direction (above) or the other (below), and that the majority of observations tend to cluster around the true value. Galileo revealed here many of the characteristics of the normal probability distribution law, and also asserted that (random) errors made in observation are distinct from (systematic) final errors arising out of computation.

Although the study of probability began much earlier, modern statistics made its first great stride with the publication in 1713 of Jacob Bernoulli's *Ars Conjectandi*, in which Bernoulli proved the weak law of large numbers. The normal distribution first appeared in 1733 as an approximation to the probability for sums of binomially distributed quantities to lie between two values, when Abraham de Moivre communicated it to some of his contemporaries. A search by Daw and Pearson (1972) confirmed that several copies of this note had been bound up with library copies of de Moivre's *Miscellanea Analytica* which were printed in 1733 or later.

The theorem appeared again in de Moivre's book *The Doctrine of Chances* (1738, 1756; 1967; see also David, 1962, where it appears as an appendix). Although the main result is commonly termed "the de Moivre-Laplace limit theorem" (see [3.4.8]), the same approximation to binomial probabilities was obtained by Daniel Bernoulli in 1770-1771, but because he published his work through the Imperial Academy of Sciences in St. Petersburg, it remained there largely unnoticed until recently (Sheynin, 1970). Bernoulli also compiled the earliest known table of the curve $y = \exp(-\mu^2/100)$; see Table 1.1.

The natural development of probability theory into mathematical statistics took place with Pierre Simon de Laplace, who "was more responsible for the early development of mathematical statistics than any other man" (Stigler, 1975, p. 503). Laplace (1810, 1811; 1878-1912) developed the characteristic function as a tool for large sample theory and proved the first general central limit theorem; broadly speaking, central limit theorems show how sums of random variables tend to behave, when standardized to have mean zero and unit variance, like standard normal variables as the sample size becomes large; this happens, for instance, when they are drawn as random samples from "well-behaved" distributions; see [6.1].

Laplace showed that a class of linear unbiased estimators of linear regression coefficients is approximately normally distributed if the sample size is large; in 1812 (Laplace, 1812, chap. VIII) he proved that the probability distribution of the expectation of life at any specified age tends to the normal (Seal, 1967, p. 207).

TABLE 1.1 Some Early Tables of Normal Functions, with Coverage and Accuracy

Source	Function	Coverage	Accuracy
D. Bernoulli (1770-1771)	$\exp(-x^2/100)$	$x = 1(1)5(5)30$	4 sig.
Kramp (1799)	$\log_{10}\{\sqrt{\pi}[1 - \Phi(\sqrt{2}x)]\}$	$x = 0(0.01)3$	7 dec.
Legendre (1826)	$2\sqrt{\pi}[1 - \Phi(\sqrt{2}x)]$	$x = 0(0.01)5$ $\exp(-x^2) = 0(0.01)0.8$	10 dec.
de Morgan (1837) de Morgan (1838)	$2\Phi(\sqrt{2}x) - 1$	$x = 0(0.01)2$ $x = 0(0.01)3$	7 dec. 7 dec.
Glaisher (1871) Markov (1888)	$\sqrt{\pi}[1 - \Phi(\sqrt{2}x)]$	$x = 3(0.01)4.5$ $x = 0(0.001)3(0.01)4.8$	11 dec. 11 dec.
Burgess (1898)	$2\Phi(\sqrt{2}x - 1)$	$x = 0(0.0001)1(0.002)3$ $(0.1)5(0.5)6$	15 dec.
Sheppard (1898)	$\sqrt{2\pi}\phi(x)$	$x = 0(0.05)4(0.1)5$	5 dec
Sheppard (1898)	Z_α	$2\alpha - 1 = 0.1(0.1)0.9$ $2\alpha - 1 = 0(0.01)0.99$	10 dec. 5 dec.
Sheppard (1903)[a]	$\Phi(x)$, $\phi(x)$	$x = 0(0.01)6$	7 dec.
Sheppard (1907)[a]	Z_α	$\alpha = 0.5(0.001)0.999$	4 dec.

[a]Incorporated into Pearson and Hartley (1958).

Source: Greenwood and Hartley, 1962.

He derived the asymptotic distribution of a single-order statistic in a linear regression problem as normal, when the parent distribution is symmetric about zero and well behaved. In 1818 he showed that when the parent distribution is normal, the least squares estimator (LSE) has smaller variance than any linear combination of observations (Stigler, 1973). In the course of deriving this result, Laplace showed that the asymptotic joint distribution of the LSE and his order statistic estimator is bivariate normal (see Chapter 9), and obtained the minimum variance property of the LSE under normality while trying to combine the two estimators to reduce the variance.

These results pertain to Laplace's work as it relates to the normal distribution. The scope of his work is much fuller; for further discussion, see Stigler (1973, 1975a).

Problems arising from the collection of observations in astronomy led Legendre in 1805 to state the least squares principle, that of minimizing the sum of squares of "errors" of observations about what we would call in modern terms a regression plane; Legendre also obtained the normal equations. In 1809, Carl Friedrich Gauss published his *Theoria Motus Corporum Coelestium*, stating that he had used the least squares principle since 1795. This led to some controversy as to priority, involving Gauss, Laplace, Legendre, and several colleagues of Gauss (Plackett, 1972), but it all hinged upon whether publication should be the criterion for settling the issue or not. In the nineteenth century, research was often done independently, without knowledge of the achievements of others, as we shall see later. It comes as no surprise, then, that Gauss knew nothing of Legendre's earlier work when he published his *Theoria Motus*.

In this work, Gauss showed that the distribution of errors, assumed continuous, must be normal if the location parameter has (again in modern terminology) a uniform prior, so that the arithmetic mean is the mode of the posterior distribution (Seal, 1967). Gauss's linear least squares model was thus appropriate when the "errors" come from a normal distribution. An American mathematician, Robert Adrain (1808), who knew nothing of Gauss' work but who may have seen Legendre's book, derived the univariate and bivariate normal distributions as distributions of errors, and hence the method of least squares (Stigler, 1977), but his work did not influence the development of the subject.

The study of least squares, or the theory of errors, was to proceed for several decades without much further interaction with developing statistical theory. The normal distribution had not yet found its place in either theoretical or applied branches of the subject, and Gauss gave little further consideration to it (Seal, 1967). However, he points out (see Maistrov, 1974, pp. 155-156) that under the normal law, errors of any magnitude are possible. Once the universality of the normal law was accepted and then assumed, as it was later for some time, scientists also assumed that all observations should therefore be retained, resulting in a delay in developing methods for identifying and discarding outliers. For a good summary of Gauss' contributions to statistics and the theory of least squares, see Sprott (1978) or Whittaker and Robinson (1924, 1926).

The astronomer Friedrich Wilhelm Bessel (1818) published a comparison of the observed residuals and those expected from Gauss' normal law of errors and found a remarkably close agreement, from sets of 300 or more measurements of angular coordinates of stars. The publication of a book by Hagen (1837), which contained a derivation of the normal law as an approximation to the distribution of the total error, when that error is assumed to be the resultant of an infinitely large number of equal but equally likely positive or negative elementary errors, may have led Bessel in 1838 to develop the hypothesis of elementary errors. Bessel thus derived the normal law as an approximation for the total error, assumed now to be the sum of a large number of mutually independent, but not identically distributed elementary errors with well-behaved properties, including symmetrical distribution about zero.

The hypothesis of elementary errors became firmly established, particularly among astronomers like G. B. Airy (1861), who