Selected Tables in Mathematical Statistics

Volume 7

Edited by the Institute of Mathematical Statistics

SELECTED TABLES IN MATHEMATICAL STATISTICS

Volume VII

THE PRODUCT OF TWO NORMALLY DISTRIBUTED RANDOM VARIABLES

by

WILLIAM Q. MEEKER, JR., LARRY W. CORNWELL, and LEO A. AROIAN

Edited by the Institute of Mathematical Statistics

Coeditors

W. J. Kennedy
Iowa State University

and

R. E. Odeh
University of Victoria

Managing Editor

J. M. Davenport
Texas Tech University

AMERICAN MATHEMATICAL SOCIETY
PROVIDENCE, RHODE ISLAND

This volume was prepared with the aid of:

J. E. Gentle, International Mathematical & Statistical Libraries, Inc.

K. Hinklemann, Virginia Polytechnic Institute and State University

R. L. Iman, Sandia Laboratories

D. B. Owen, Southern Methodist University

S. Pearson, Bell Laboratories

R. H. Wampler, National Bureau of Standards

1980 *Mathematics Subject Classification*
Primary 62Q05; Secondary 62E15, 60E05.

International Standard Serial Number 0094-8837
International Standard Book Number 0-8218-1907-0
Library of Congress Card Number 74-6283

PREFACE

This volume of mathematical tables has been prepared under the aegis of the Institute of Mathematical Statistics. The Institute of Mathematical Statistics is a professional society for mathematically oriented statisticians. The purpose of the Institute is to encourage the development, dissemination, and application of mathematical statistics. The Committee on Mathematical Tables of the Institute of Mathematical Statistics is responsible for preparing and editing this series of tables. The Institute of Mathematical Statistics has entered into an agreement with the American Mathematical Society to jointly publish this series of volumes. At the time of this writing, submissions for future volumes are being solicited. No set number of volumes has been established for this series. The editors will consider publishing as many volumes as are necessary to disseminate meritorious material.

Potential authors should consider the following rules when submitting material.

1. The manuscript must be prepared by the author in a form acceptable for photo-offset. This includes both the tables and introductory material. The author should assume that nothing will be set in type although the editors reserve the right to make editorial changes. The tables must be produced from computer output.

2. While there are no fixed upper and lower limits on the length of tables, authors should be aware that the purpose of this series is to provide an outlet for tables of high quality and utility which are too long to be accepted by a technical journal but too short for separate publication in book form.

3. The author must, wherever applicable, include in his introduction the following:

(a) He should give the formula used in the calculation, and the computational procedure (or algorithm) used to generate his tables. Generally speaking, FORTRAN or ALGOL programs will not be included but the description of the algorithm used should be complete enough that such programs can be easily prepared.

(b) A recommendation for interpolation in the tables should be given. The author should give the number of figures of accuracy which can be obtained with linear (and higher degree) interpolation.

(c) Adequate references must be given.

(d) The author should give the accuracy of the table and his method of rounding.

(e) In considering possible formats for his tables, the author should attempt to give as much information as possible in as little space as possible. Generally speaking, critical values of a distribution convey more information than the distribution itself, but each case must be judged on its own merits. The text portion of the tables (including column headings, titles, etc.) must be proportional to the size 5–1/4″ by 8–1/4″. Tables may be printed proportional to the size 8–1/4″ by 5–1/4″ (i.e., turned sideways on the page) when absolutely necessary; but this should be avoided and every attempt made to orient the tables in a vertical manner.

(f) The table should adequately cover the entire function. Asymptotic results should be given and tabulated if informative.

(g) An example or examples of the use of the tables should be included.

4. The author should submit as accurate a tabulation as he can. The table will be checked before publication, and any excess of errors will be considered grounds for rejection. The manuscript introduction will be subjected to refereeing and an inadequate introduction may also lead to rejection.

5. Authors having tables they wish to submit should send two copies to:

Dr. Robert E. Odeh, Coeditor
Department of Mathematics
University of Victoria
Victoria, B. C., Canada V8W 2Y2

At the same time, a third copy should be sent to:

Dr. William J. Kennedy, Coeditor
117 Snedecor Hall
Statistical Laboratory
Iowa State University
Ames, Iowa 50011

Additional copies may be required, as needed for the editorial process. After the editorial process is complete, a camera-ready copy must be prepared for the publisher.

Authors should check several current issues of *The Institute of Mathematical Statistics Bulletin* and *The AMSTAT News* for any up-to-date announcements about submissions to this series.

ACKNOWLEDGMENTS

The tables included in the present volume were checked at the University of Victoria. Dr. R. E. Odeh arranged for, and directed this checking with the assistance of Mr. Bruce Wilson. The editors and the Institute of Mathematical Statistics wish to express their great appreciation for this invaluable assistance. So many other people have contributed to the instigation and preparation of this volume that it would be impossible to record their names here. To all these people, who will remain anonymous, the editors and the Institute also wish to express their thanks.

To: **KAREN, SARA, and ARMINÉ**

Contents of VOLUMES I, II, III, IV, V, and VI of this Series

TABLE OF CONTENTS

THE PRODUCT OF TWO NORMALLY DISTRIBUTED RANDOM VARIABLES

William Q. Meeker, Jr., Iowa State University

Larry W. Cornwell, Western Illinois University

Leo A. Aroian, Union College and University

ABSTRACT

Tables for the fractiles of the distribution of the product of two normal random variables are presented. The numerical methods used to compute and check the tables are described and some of the applications of this distribution are reviewed. Interpolation in the tables is discussed and some examples are given.

1. INTRODUCTION

The product of two normally distributed random variables occurs frequently in applications, for example, in the fields of physical, engineering, and social sciences, and also in biometry, sampling, auditing, and other business applications. Some specific applications are outlined in Section 6. The important properties of this distribution were first investigated by Craig (1936). Some further mathematical results are given by Aroian et al. (1978). Aroian (1947) showed that under certain conditions the distribution is asymptotically normal. For situations when these conditions are not met, the tabulations of the fractiles, presented here can be used.

Tables of the distribution of the product of two <u>independent</u> normally distributed random variables were originally given by Aroian (1959). The present tables contain fractiles of the distribution of the

---

Received by the editors January 1980 and in revised form May 1980.

product of two possibly dependent normally distributed random variables.

The important properties of this distribution are reviewed in Section 2 and numerical methods for computing the distribution function are given in Section 3. The tables of the distribution function and the fractiles are explained in Sections 4 and 5 respectively. Interpolation in the tables is explained in Section 7 and some examples of interpolation are given in Section 8.

2. NOTATION AND PROPERTIES OF THE DISTRIBUTION

Let X_1 and X_2 be two normally distributed random variables with means μ_i, variances σ_i^2, i = 1, 2, and correlation coefficient ρ. Define $Z = X_1X_2/(\sigma_1\sigma_2)$, $\delta_1 = \mu_1/\sigma_1$, and $\delta_2 = \mu_2/\sigma_2$. For the random variable Z, Craig (1936) derived the distribution function, the moment generating function, the cumulants, mean, variance and measures of skewness and kurtosis. Briefly, his results are as follows.

The mean and variance of Z are

$$\begin{aligned} \mu_Z &= \delta_1\delta_2 + \rho \\ \sigma_Z^2 &= \delta_1^2 + \delta_2^2 + 2\delta_1\delta_2\rho + 1 + \rho^2. \end{aligned} \tag{2.1}$$

Measures of skewness and kurtosis are

$$\xi_3 = \mu_3/\sigma_Z^3 = \{6[\rho(\delta_1^2 + \delta_2^2) + \delta_1\delta_2(1 + \rho^2)] + 2\rho(3 + \rho^2)\}/\sigma_Z^3$$

$$\begin{aligned} \xi_4 = \mu_4/\sigma_Z^4 = [&12(\delta_1^2 + \delta_2^2)(1 + 3\rho^2) + 24\delta_1\delta_2\rho(3 + \rho^2) \\ &+ 6(1 + 6\rho^2 + \rho^4) + 3\sigma_Z^4]/\sigma_Z^4 \end{aligned}$$

where μ_3 and μ_4 are the third and fourth central moments, respectively, of Z.

Let $F_Z(z) = F_Z(z;\delta_1,\delta_2,\rho) = P(Z \leq z)$ be the cumulative distribution function of Z. Evident relationships for the distribution are

$$\begin{aligned} F_Z(z;\delta_1,\delta_2,\rho) &= F_Z(z;\delta_2,\delta_1,\rho) \\ F_Z(-z;-\delta_1,\delta_2,-\rho) &= 1 - F_Z(z;\delta_1,\delta_2,\rho). \end{aligned} \tag{2.2}$$

Also, there is a singularity in the pdf $f_Z(z)$ at $z = 0$. The characteristic function of Z is

$$\varphi_Z(t) = [\exp(-N/D)]/E, \text{ where}$$

$$N = (\delta_1^2 + \delta_2^2 - 2\rho\delta_1\delta_2)t^2 - 2\delta_1\delta_2 it$$

$$D = 2[1 - (1 + \rho)it][1 + (1 - \rho)it]$$

$$E = (D/2)^{1/2}$$

For further information, see Craig (1936) and Aroian, Taneja, and Cornwell (1978).

Aroian et al. (1978) discuss several special cases for the distribution of Z. For $\rho = 1$ and $\delta_1 = \delta_2 = 0$, the distribution is chisquare with 1 degree of freedom. When $\rho = 1$ and $\delta_1 = \delta_2 = \delta$, the distribution is noncentral chisquare with 1 degree of freedom and noncentrality parameter $\lambda = \delta^2/2$. A further result for $\rho = 1$ (-1) is that the distribution of $Z - \mu_Z$ depends only on $\delta_1 + \delta_2$ ($\delta_1 - \delta_2$) and is thus related to the noncentral chisquare distribution in the obvious way. This can be seen in the tables as well as by manipulation of the characteristic function of Z.

3. NUMERICAL METHODS

The distribution function and fractiles were initially computed as described in Sections 4 and 5 using the algorithm of Cornwell, et al. (1978). This was done on the CDC Cyber 170 model 730 using single precision (60 bit words). The entire fractile table was checked with the Iowa State University ITEL AS/6 system using double precision (64 bit words) with a more accurate algorithm due to Meeker and Miller (1979). This algorithm is described here.

Assume that X_1 and X_2 follow a bivariate normal distribution with means δ_1 and δ_2 respectively, unit variances and correlation ρ. Then $Z = X_1X_2$ follows the distribution of the product of two normally distributed random variables with parameters δ_1, δ_2 and ρ.

The distribution of Z can be obtained by integrating

$$F_Z(z) = P(Z \leq z) = \iint_{x_1 x_2 < z} g(x_1, x_2; \delta_1, \delta_2, \rho) dx_1 dx_2 \tag{3.1}$$

where

$$g(x_1, x_2; \delta_1, \delta_2, \rho) = [2\pi(1-\rho^2)^{1/2}]^{-1} \exp(-Q/2)$$

and

$$Q = \frac{1}{1-\rho^2} \{(x_1 - \delta_1)^2 - 2\rho(x_1-\delta_1)(x_2-\delta_2) + (x_2-\delta_2)^2\} .$$

A numerical algorithm for computing this integral is developed here for $z \geq 0$. For $z < 0$, the relationship in (2.2) can be used.

The probability (volume) represented by (3.1) is evaluated by dividing the $x_1 x_2$ plane into 8 parts and integrating over each of these parts, thus avoiding the problem of singularities and improving computational efficiency (see Figure 1). The task is somewhat simplified because the integrands of the interated integrals can be expressed as functions of the standard normal cumulative distribution and density functions which can be efficiently computed to near machine accuracy.

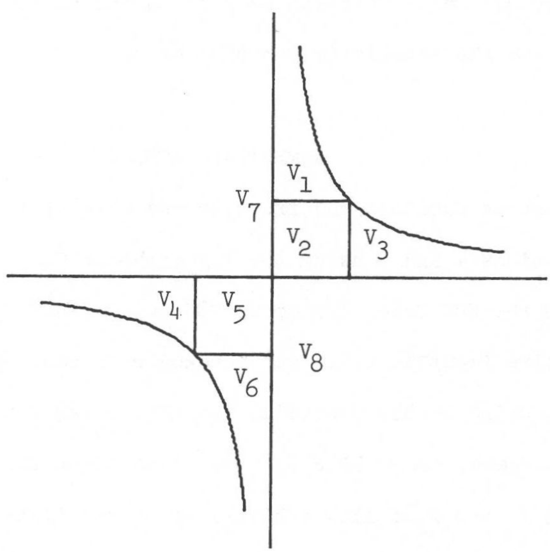

Figure 1. Regions of Integration

For $z \geq 0$, the probability in (3.1) is then computed as the sum of the following integrals.

$$
\begin{aligned}
V_1 &= \int_{\sqrt{z}}^{\infty} G_E(x_2, z; \delta_2, \delta_1, \rho)dx_2 \\
V_2 &= \int_{0}^{\sqrt{z}} G_R(x_1, 0, \sqrt{z}; \delta_1, \delta_2, \rho)dx_1 \\
V_3 &= \int_{\sqrt{z}}^{\infty} G_E(x_1, z; \delta_1, \delta_2, \rho)dx_1 \\
V_4 &= -\int_{-\infty}^{-\sqrt{z}} G_E(x_1, z; \delta_1, \delta_2, \rho)dx_1 \\
V_5 &= \int_{-\sqrt{z}}^{0} G_R(x_1, -\sqrt{z}, 0; \delta_1, \delta_2, \rho)dx_1 \qquad (3.2) \\
V_6 &= -\int_{-\infty}^{-\sqrt{z}} G_E(x_2, z; \delta_2, \delta_1, \rho)dx_2 \\
V_7 &= \int_{-\infty}^{0} G_R(x_1, 0, \infty; \delta_1, \delta_2, \rho)dx_1 \\
V_8 &= \int_{0}^{\infty} G_R(x_1, -\infty, 0; \delta_1, \delta_2, \rho)dx_1
\end{aligned}
$$

where

$$
\begin{aligned}
&G_E(y, z; \delta_y, \delta_x, \rho) = \\
&\qquad \{\Phi[R_1(z/y-\delta_x-\rho(y-\delta_y))]-\Phi[R_1(-\delta_x-\rho(y-\delta_y))]\}\varphi(y-\delta_y) \\
&G_R(y, x_L, x_u; \delta_y, \delta_x, \rho) = \qquad (3.3) \\
&\qquad \{\Phi[R_1(x_u-\delta_x-\rho(y-\delta_y))]-\Phi(R_1(x_L-\delta_x-\rho(y-\delta_y))]\}\varphi(y-\delta_y)
\end{aligned}
$$

and $\Phi(\cdot)$ and $\varphi(\cdot)$ are the standard normal cumulative distribution and density functions respectively and $R_1 = (1-\rho^2)^{-1/2}$.

The normal cumulative distribution functions were computed with the IBM FORTRAN supplied double precision complementary error function (DERFC). Integration was done with the IMSL (1975) subroutine DCADRE for cautious, adaptive Romberg integration, with a specified absolute error tolerance of 10^{-10}. Several values of $\delta_i + \gamma$ were used to replace ∞ in the limits of integration. It was determined that $\delta_i + \gamma = 8$ yields accuracy of at least 10 decimal places in $F_Z(z)$.

When $\rho = 1$, the distribution is degenerate and the following scheme can be used. Again, let $X_i \sim N(\delta_i, 1)$, i=1,2 and now, without loss of generality, let $\rho = 1$. Then $X_1 - \delta_1 = X_2 - \delta_2$ or $X_2 = X_1 + a$ where $a = \delta_2 - \delta_1$. The distribution of $Z = X_1 X_2 = X_1^2 + aX_1$ is

$$F_Z(z;\delta_1,\delta_2,1) = P(Z < z) =$$

$$P(X_1^2 + aX_1 - z < 0) = \begin{cases} \Phi(r_U - \delta_1) - \Phi(r_L - \delta_1) & z > -a^2/4 \\ 0 & z \le -a^2/4 \end{cases}$$

where

$$r_L = [-a - (a^2 + 4z)^{1/2}]/2$$

$$r_U = [-a + (a^2 + 4z)^{1/2}]/2.$$

For $\rho = -1$, use (2.2).

4. THE STANDARDIZED DISTRIBUTION FUNCTION

Let $W = (Z-\mu_Z)/\sigma_Z$, where μ_Z and σ_Z are defined in (2.1). As shown by Aroian (1947), W is asymptotically N(0,1) for either δ_1 or δ_2 or both large in absolute value, except when $\rho = 1$ and $\delta_1 = -\delta_2$ or when $\rho = -1$ and $\delta_1 = \delta_2$ in which case the asymptotic distribution is related to the noncentral chisquare distribution, as described in Section 2. The standardized distribution function $F_W(w) = P(W \le w)$ has been computed for values of $\delta_1 \le \delta_2$ with $\delta_1, \delta_2 = 0(.4)4, 6, 12$, $-6 \le w \le 6$ and $\rho = 0.00, \pm 0.20, \pm 0.40, \pm 0.60, \pm 0.80, \pm 0.90, \pm 0.95$, and ± 1.00. The interval in w is 0.2 from -3.2 to 3.2 and 0.4 outside this interval. A computer tape of these tabulations can be obtained from the authors.

5. TABLES OF THE FRACTILES w_α

The present tables contain the fractiles of the standardized distribution function $F_W(w)$. These are solutions of the equations $F_W(w) = \alpha$ and are tabled here for 39 values of α between 0.0005 and 0.9995 for the same values of δ_1, δ_2, and ρ used in the tabulation of $F_W(w)$ described in Section 4. The choice of $\delta_1 \leq \delta_2 = 4, 6, 12, \infty$ was made to facilitate harmonic interpolation but this must be used judiciously.

The fractiles were computed using the algorithm of Cornwell et al. (1978) and an inverse iterative interpolation scheme. Iterations were stopped when a value of w_α was found such that $|F_W(w_\alpha)-\alpha| < .5\text{x}10^{-7}$ or $|W_{\alpha_i}-W_{\alpha_{i-1}}| < .5\text{x}10^{-6}$ where W_{α_i} is the i^{th} iterate of W_α.

The fractiles were checked with the algorithm described in Section 3 using the following procedure. For particular values of δ_1, δ_2, ρ, and α, the fractile estimate of w_α was rounded to exactly five places after the decimal and then incremented (if $F_W(w_\alpha) < \alpha$) or decremented (if $F_W(w_\alpha) > \alpha$) by 0.00001 to see if $|F_W(w_\alpha) - \alpha|$ could be reduced.

The asymptotic normal distribution fractiles (for $\delta_1 = \infty$ or $\delta_2 = \infty$ or $\delta_1 = \delta_2 = \infty$) are given for all values of ρ except -1. When $\rho = -1$ and as $\delta = \delta_1 = \delta_2$ increases, the asymptotic distribution is the same as that for $\rho = -1$, $\delta_1 = \delta_2 = 0$. When δ_1 is finite and fixed and δ_2 increases (or vice versa) the asymptotic distribution is again normal.

6. APPLICATIONS

Craig (1936) mentions two inquiries concerning the distribution of a product, one from an investigator in business statistics and another from a psychologist. The third author's own interest arose from his use of sampling methods in auditing while at the Metropolitan Life Insurance Company, particularly for efficient verification of policy reserves. Problems such as the following are typical.

Let $A = \sum_{i=1}^{T} v_i$ be a population total which needs to be estimated. The mean of the population $\mu = \sum_{i=1}^{T} v_i/T$ is unknown. Also, it is often too costly to determine T exactly. However, both μ and T can be estimated economically by sampling. Estimates $\hat{\mu}$ and $\hat{T}$ and their estimated standard errors $s_{\hat{\mu}}$ and

$s_{\hat{T}}$ are thus obtained. If the estimates $\hat{\mu}$ and $\hat{T}$ are independent, $\rho = 0$; otherwise ρ must be determined or estimated. It is also assumed that $\hat{T}$ and $\hat{\mu}$ approximately follow a normal distribution. In practice, these usually are not restrictive assumptions. In some situations, T is estimated with different sample units than those used to estimate μ, often because T is more readily estimated than μ.

Under these assumptions, the estimated total

$$\hat{A} = \hat{\mu}\,\hat{T}$$

follows the distribution of the product tabled here with estimated standard error

$$s_{\hat{A}} = (\hat{T}^2 s_{\hat{\mu}}^2 + \hat{\mu}^2 s_{\hat{T}}^2 + 2\hat{T} s_{\hat{T}} \hat{\mu} s_{\hat{\mu}} \hat{\rho} + s_{\hat{\mu}}^2 s_{\hat{T}}^2 (1 + \hat{\rho}^2))^{1/2}$$

For the distribution of $\hat{A}$, $\delta_\mu = \mu/\sigma_{\hat{\mu}}$ and $\delta_T = T/\sigma_{\hat{T}}$. If either δ_μ or δ_T is large, one can take advantage of the asymptotic normality of $\hat{A}$ (Aroian (1947)); in any case, the tables presented here can be used to find the fractiles of $\hat{A}$ and, for example, to find approximate confidence intervals for the true value of A in the obvious manner. Inspection of the tables gives an indication of how large the δ's need to be to use the asymptotic distribution. It is seen, for example, that at least one of the δ's must be large to use the approximation in the tails of the distribution.

The distribution of the product is a function of δ_1, δ_2 and ρ. Use of the distribution in inferential applications assumes that these quantities and the σ's are known. In practice, the δ's and/or the σ's are unknown and sample estimates must be used, yielding approximate results.

An example of the above situation is the estimation of the total accounts receivable (A) for a large company, the total number of accounts (T) being unknown. First, a random sample is taken from a known number of files (or offices) in order to estimate the total number of accounts. Then a random sample of accounts is taken to estimate μ, the average account size. The estimate of the total accounts receivable is then the product of these two estimates. Similar problems arise in studies of wildlife population, and in the estimation of total production, crop yield, etc. For example, Cordeiro and Rathie (1979) use the distribution to estimate total soybean and wheat yields for a particular area.

Sampson and Bruening (1971) have applied the distribution of the product to the uniformity specification of drugs in tablet form. Let tablet weights be designated by y and drug substance by p, then the problem is to estimate μ_{yp} by $\hat{\mu}_{yp} = \bar{y}\,\bar{p}$ where $\bar{y}$ and $\bar{p}$ are sample estimates, which they assume are independent. They emphasize that y and p need not be distributed normally since for moderately large samples $\bar{y}$ and $\bar{p}$ will be.

The distribution of the product is useful in economic theory for supply and demand situations involving prices and quantities. In reliability problems the product distribution has been used to find confidence intervals in a series system to estimate $p_1 p_2$, where p_i is the reliability of system i and the individual estimates are approximately normally distributed. Better methods are given by Mann (1972). Other applications are given by Broadbent (1956) and Barnard (1962) who use the lognormal to approximate the distribution of the product.

Donohue (1964) notes uses of the product in the cyclic firing rate of cannons and in the distribution of measurement errors. His report includes a large list of references as do the reports of Springer et al. (1964, 1966). Both discuss some more general problems concerning products and quotients. Aroian used the distribution of the product in some undocumented random noise problems at Hughes Aircraft Company, and Deutsch (1962) shows how the distribution of the product can be applied to the study of multiplicative random processes which occur in certain physical systems. We have been told that the distribution of the product is useful in problems associated with reliability assessment of nuclear reactors.

In quite a different example, Devlin et al. (1974) use the distribution of the product in their researches on robust estimation and outlier detection