COPY 2 of 3 COPIES

ARL 64-115
JULY 1964

180p.
HC $5.00
MF $1.00

AD 603667

Aerospace Research Laboratories

PRODUCTS AND QUOTIENTS OF RANDOM VARIABLES AND THEIR APPLICATIONS

JAMES D. DONAHUE
THE MARTIN COMPANY
DENVER, COLORADO

OFFICE OF AEROSPACE RESEARCH

United States Air Force

NOTICES

400 - July 1964 - 162-47-956

ARL 64-115

PRODUCTS AND QUOTIENTS OF RANDOM VARIABLES AND THEIR APPLICATIONS

JAMES D. DONAHUE

THE MARTIN COMPANY
DENVER, COLORADO

JULY 1964

Contract AF 33(615)-1023
Project 7071
Task 7071-01

AEROSPACE RESEARCH LABORATORIES
OFFICE OF AEROSPACE RESEARCH
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

FOREWORD

This monograph was prepared by James D. Donahue, Electronics and Mathematics Laboratory, The Martin Company, Denver, Colorado, on Contract AF33(615)-1023 for the Aerospace Research Laboratories, Office of Aerospace Research, United States Air Force. The research reported herein was accomplished on Task 7071-01, "Mathematical Techniques of Aeromechanics" of Project 7071, "Research in Mathematical Statistics & Probability" under the supervision of Dr. P.R. Krishnaiah of the Applied Mathematics Research Laboratory, ARL. This report covers work conducted during the period September 1963-March 1964.

The author wishes to express his sincere appreciation for the valuable assistance rendered by Dr. S. Birnbaum, Dr. R.S. Novosad, and Dr. G.W. Morgenthaler, Manager of the Electronics and Mathematics Laboratory, in compiling this monograph. He is also grateful to Dr. P.R. Rider for his suggestions on exposition.

ABSTRACT

In this report, the general techniques for determining the distributions of products and quotients of random variables are discussed. Some exact and asymptotic results pertaining to the distributions of the products and quotients of certain random variables which generally occur as measurement error are also presented together with their applications. An extensive bibliography is included at the end of the report.

TABLE OF CONTENTS

I. GENERAL INTRODUCTION

In the applied sciences, problems are frequently encountered concerning reliability analyses, measures of efficiency, biometry indices, etc., which properly belong in the framework of determining the distributions of various algebraic combinations of random variables. Many of the problems associated with this general area, and in particular with product and quotient forms of these variables, have been extensively investigated. Craig [21][1] and others, however, point out that these investigations have resulted in surprisingly little useful information which is generally available to the engineer, research scientist, etc. The scope of such problems is surprisingly broad and investigations into methods of analysis date as far back as to K. Pearson [97], 1910. As a collective entity, the literature concerning product and quotient forms of random variables is widely scattered and, unfortunately, is devoted almost exclusively to investigations utilizing very specialized quadrature methods. As a result, no publication has appeared which presents a general theory, methods of application, and useful tabular results pertaining to a variety of product and quotient forms of random variables.

Scattered and incomplete tables concerning exact distribution percentage points, approximating distributions, and other applied topics are available. To the engineer unschooled in the rudiments of random variable techniques, these are often meaningless and force him to use

1) In the body of this report, numbers in brackets, [ ], refer to the bibliography.

less desirable deterministic methods or even perhaps to make unwarranted assumptions so that the tabular results might be used.

The aim of this monograph is to present a useful expository manuscript on certain frequently occurring product and quotient forms of random variables which will present a general theory, bring together and discuss known results, and reference pertinent tabular results so as to provide a useful tool to the engineer or research scientist.

In this respect, the monograph is devoted to the following topics:

1) Examples of applied problems involving products and quotients. To demonstrate a wide range of applications to engineering problems, a number of varied examples are discussed, mainly from the viewpoint of recognizing and properly posing the problems in this general area.

2) General theoretical models for determining product and quotient distributions. The general techniques for determining the distributions of products and quotients of random variables are presented. This presentation treats the independent and dependent cases of continuous variables.

3) Exact results pertaining to products and quotients of random variables which generally occur as measurement error. Considerable attention is devoted to the quotient distributions; normal/normal, rectangular/normal, triangular/normal; to the product distribution, normal x normal, rectangular x normal, triangular x normal; and to certain discrete distributions of the form $\frac{1}{X}$.

4) Limiting distributions, approximations and asymptotic results. The general variable $Z = \prod_{i=1}^{n} X_i \Big/ \prod_{j=n+1}^{s} Y_j$ is discussed in respect to fitting the lognormal approximation to Z, along with other topics.

5) Characterizing properties of statistical distributions. Problems of the type: "If $X_1/X_2 = Y$ and Y follows a general F distribution, are X_1 and X_2 necessarily chi-square variables?" are treated.

6) An annotated bibliography. Products and quotients of random variables assume many different representations. Not only is one interested in ratios and products of random variables described by specific frequency functions, but also in functional forms of variances, ranges, proportions, etc.

The bibliography presented herein is very comprehensive in respect to articles pertaining to the theory of the distribution of random variables described by probability density functions. The remaining entries in the bibliography represent a small representative sample of the vast number of published articles pertaining to other forms of products and quotients.

The notation used herein conforms for the most part to what is believed to be a standard format in statistical literature. The symbols denoting random variables will be X, Y, Z and the values of the variates, x, y, z. Frequency or probability density functions, p.d.f.'s will be denoted by small letters $f(x)$, $g(y)$, . . . , and the respective cumulative distribution functions, c.d.f.'s, by $F(x)$, $G(y)$, . . . , . A random variable obeying a certain probability density function, say the Gaussian or normal probability law, will be denoted as $N(\mu, \sigma^2)$. Here μ and σ^2 respectively denote the expected value and the variance of the normal random variable.

II. ENGINEERING APPLICATIONS INVOLVING QUOTIENTS OR PRODUCTS OF RANDOM VARIABLES

The examples discussed in this chapter serve to emphasize the broad range of product and quotient forms arising in engineering applications of statistical distribution theory.

2.1 Cyclic Firing Rate of the T-160 (20 mm) Cannon in Korean Combat[2]

The need for establishing the reliability and functional suitability of weapons under combat conditions is apparent. The following account describes a preliminary model for a statistical analysis which was used to determine the combat suitability of the then newly manufactured T-160 gun installed on jet fighter aircraft during a period in the Korean conflict.

It was desired to establish the cyclic gun firing rate the T-160 achieved during combat by assessing sight reticle and scope camera film exposed during missions and from questionnaires filled out by pilots and armament technicians detailing the number of rounds expended on each sortie from belt counts. The data were used to obtain a probabilistic estimate of the typical cyclic rate of the T-160 for the Korean combat test.

The parameters r and Θ defined below refer to sortie averages. If

r = average cyclic gun rate per sortie in rounds/sec/gun,

Θ = average camera speed per sortie in frames/sec.,

---

2) George W. Morgenthaler, "Cyclic Rate of the T-160 in Korean Combat," Combat Film Analysis Newsletters, Institute of Air Weapons Research, University of Chicago, 1954.

then

$$(\text{No. of guns firing}) \frac{\text{No. firing frames}}{\Theta} r = (\text{No. of rounds expended}). \qquad (2\text{-}1)$$

There are, of course, certain errors in the input data for formula (2-1). The firing frame count may be in error because of the difficulty inherent in distinguishing a firing frame from a nonfiring frame, particularly with poor quality film. The nominal camera speed (Θ = 32 fr/sec.) is subject to a factory setting error, in addition to errors induced by mechanical wear, maintenance irregularities, and to errors arising from variations in temperature and the camera circuit voltage.

Test firings indicated that the average cyclic rate, r, on a sortie is a function of total time of fire, the number of bursts, the quality of the gun, etc. Being a function of random variables, r is a random variable and thus would generate a distribution over a large number of sorties. The expected value, E(r), is representative of the sortie cyclic rates experienced in Korea.

In the treatment of this problem, r was treated as a product of the dependent random variables Θ and r/Θ and E(r) was defined by the relationship:

$$E(r) = E\left(\Theta \bullet \frac{r}{\Theta}\right) = E(\Theta)\; E\left(\frac{r}{\Theta}\right) + \rho_{\Theta, \frac{r}{\Theta}} \sqrt{v(\Theta)\; V\left(\frac{r}{\Theta}\right)}\;; \qquad (2\text{-}2)$$

where:

$\rho_{\Theta, \frac{r}{\Theta}}$ = coefficient of correlation between Θ and $\frac{r}{\Theta}$,

$V(\Theta)$ = variance of Θ, and

$V\left(\frac{r}{\Theta}\right)$ = variance of $\frac{r}{\Theta}$.

In this case, the moments of the distribution of $r = (\Theta)\left(\frac{r}{\Theta}\right)$ were first obtained by making the assumption that Θ and $\frac{r}{\Theta}$ were normally distributed. This assumption was supported by experimental data. It was found that $1164 < E(r) < 1266$ rds./min. with a probability of .90. This estimate was surprisingly low in comparison with the T-160 designed cyclic rate of 1500 rds/min. It is apparent that with the percentage points of the distribution of $r = (\Theta)\left(\frac{r}{\Theta}\right)$, one could examine the probability of an E(r) such as was attained. This very low probability would cause one to search for a plausible explanation.

In this case, it was discovered that since the T-160 is a gas-operating, automatic weapon consisting of barrel and a rotating drum with five chambers, then the inertia of the drum would play an important role. At the outset, the inertia of the drum is overcome rather slowly and, hence, the gun does not have a cyclic rate but rather an average cyclic rate depending upon the duration of the burst of fire.

Computation of the expected T-160 cyclic rate for Korean type combat based on the Korean T-160 burst length later proved that values of $E(r) \cong 1200$ rds./min. were to be expected.

2.2 Selection of a Space "Workhorse" Booster.

Many experts agree that all long range space explorations to be attempted in the intermediate future will best be initiated from an earth or parking orbit. In this respect, the success of such space missions will then greatly depend upon the successful completion of a very important logistics operation. This operation will involve the transportation of all necessary equipment and supplies into the earth orbit and the assembly of it there.

The selection of a suitable "workhorse" booster for this phase of the project represents only one of the myriad of complex decisions facing the project planners. But, since this decision must be made years ahead of most others and will involve great expenditures of money, manpower, and time, it perhaps represents the most important present day space industry problem.

Let us assume that, at the present time, the choice of a workhorse booster is restricted to typical systems such as the Nova, Saturn C-5, and the Titan III systems. The concepts of these boosters differ considerably;and any operational boosters forthcoming from the e projects will most assuredly differ in respect to payload capacity, costs, velocities, reliability, date of assembly of the first operational booster, etc. Thus the problem of an "optimal" selection of one booster system for the transportation job must be based on a very realistic evaluation model.

One meaningful index for an evaluation model of this type is the "dollar" cost per pound of equipment placed in orbit. This index would be obtained, of course, by dividing the total system cost, C_{system}, by the payload weight placed in orbit, $W_{payload}$, or

$$I = \frac{C_{system}}{W_{payload}} . \tag{2-3}$$

A closer investigation of C_{system} and $W_{payload}$ reveals that these components are functions of variable components. For example, total system costs include both research and development costs and operational costs - estimates of which are highly uncertain. Uncertainties in R & D cost estimates intervene in the form of unanticipated differences in estimated costs of component parts and in increased expenditures caused by modifications in the design to meet new or revised performance specifications. Operational costs are highly susceptible to the failure rates of the component parts and, more generally, to the level of "sophistication" in the supporting logistics system.

The weight and space of payload placed in orbit are functionally dependent upon various characteristics of the missile such as Specific Impulse, Isp, the weight-to-stage weight ratio, λ_1, and other aerodynamic characteristics, λ_i.

It becomes apparent from these considerations that equation (2-3) represents a complex random variable and would be better represented by

$$I = \frac{Cost_{R\&D} + C_{Operational}}{f(Isp, \lambda_1, \lambda_2, \ldots, v)} . \tag{2-4}$$

Additionally, each component of C_{system} and $W_{payload}$ is random in respect to measurement or estimate error. For instance, experimental data collected from operational booster systems to date show that the Isp of an "average" missile of the system varies from the parameter value the designer had intended.

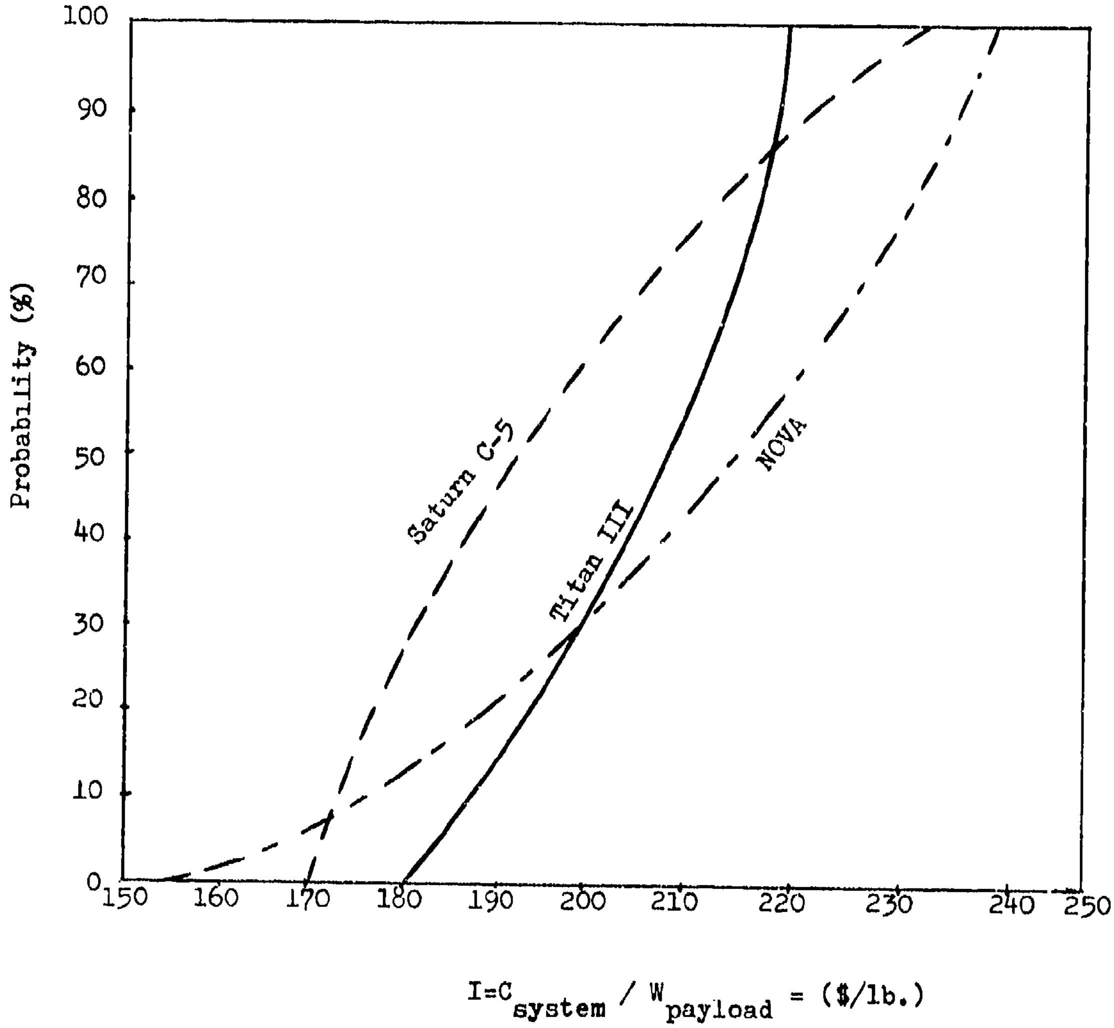
Without the knowledge of how to combine the functional forms of the components of I, only a point estimate of I will be possible. On this basis, the choice of a booster system would be greatly influenced by the relative differences in the point estimates of I for these systems. On the other hand, the c.d.f.'s of I, F(I), for each booster may be obtained with a knowledge of random variable techniques. Assume, for illustrative purposes, that the c.d.f.'s of Figure 1 are realistic.

Several additional, equally good, criteria for the selection of a booster system may be considered when the information of Figure 1 is available. For example, a system might be selected on the basis of: 1) the system expected to become operational soonest and for which I is greater than some specified level with probability $\geq$ 50%, say, or 2) the system for which I is largest at a certain prescribed level of probability, or 3) the system for which the maximum cost estimate is a minimum, etc.

More important , all of these criteria may be evaluated simultaneously with the aid of estimates of the respective c.d.f.'s. Such an analysis is certain to provide much more valuable insight into the problem of selecting an appropriate booster system.

2.3 Replenishment of a Life Support System.

With the advent of long-duration manned space flights, estimating the changes that will occur in the quantities of certain substances in the desired ecological system has become more complex.



The Cumulative Probability Distributions, F(I), of the Three Booster Systems.

Figure 1.

Ideally, a mathematical model for general system analysis of an ecological system for long-duration flights will provide for:

1) A formulation of the control problem from which optimal control functions can be determined;

2) A preliminary design for the ecological system in respect to system stability and control considerations;

3) A method for determining the time required to restore the system to a suitable balance in the case of a mishap and the time required initially to put the system into operation; and

4) The determination of resupply requirements.

One aspect of a preliminary model[3] designed to include the above considerations has resulted in the requirement for evaluating a product of random variables. It may be stated in the following manner:

Consider the amount of oxygen in a cabin atmosphere. The amount is affected by leakage, crew consumption, and resupply from a storage capacity. Let:

$X(t)$ = amount of oxygen (in moles) in the cabin atmosphere at time t,

L = proportion of rate of loss of oxygen due to leakage from the cabin atmosphere per time period,

---

3) Davis, Henry and Novosad, R. S., "Control, Replenishment, and Stability of Life Support Systems," Journal of Spacecraft and Rockets, Amer. Inst. of A. and A., January 1964.

Y_1 = rate of increase in oxygen content in the cabin atmosphere from storage per time period, and

K = rate of decrease in the oxygen content in the cabin atmosphere due to crew consumption per time period.

The estimated amount of oxygen at time t is

$$X(t) = \exp(-Lt)\left\{\int_0^t \exp(Lt)\,(Y_1 - K)\,dt + X(0)\right\}, \tag{2-5}$$

where

X(O) denotes the initial condition.

In this study, $W = \exp(Lt)$ and $Z = (Y_1 - K)$ represent random variables which are functions of time. A knowledge of combining random variables in product forms is required for solving this problem.

2.4 Stochastic Differential Equations with Product Coefficients.

In the space industry, stochastic differential equations whose coefficients are functional forms of products and quotients of random variables frequently play important roles. They appear in investigations of control theory problems, in analyses of stresses and material performances, and in estimations of the effects of hypervelocity impact.

Most frequently, the designer or engineer treats this type of equation as a deterministic model by using moments of the distribution as the constant coefficients in order to obtain a solution. If bounds are required for a particular solution, digital computer simulation of the effects of small perturbations in the coefficients usually affords a convenient method of analysis.

In the field of missile trajectory analysis, both very complicated models, such as a complete design optimization analysis, and very simple models, such as the path of motion in a plane of a point particle, are needed. The latter pro[illegible] analytic solutions in closed form from which valuable insight regar[illegible] trajectories can be obtained. They also furnish simple trajectory patterns that are of great value for problems involving the simultaneous optimization of design and trajectory. The more complicated models make it possible to take into account many effects that must of necessity be omitted from the simpler models and also provide a good framework for analyzing the complete system of simple models.

One of the simpler design models is that of electrical circuitry analysis. Such situations arise frequently in electrical circuits where the current flowing in one circuit is influenced by the current flowing in another through direct interconnection or through a mutual inductance. The differential equations for the currents are obtained by considering the potential drops across various elements of the circuit. This leads to the following pair of simultaneous equations:

$$R_1 I_1 + L_1 \frac{dI_1}{dt} + \frac{M\, dI_2}{dt} = E_0 \cos \omega t,$$

$$R_2 I_2 + L_2 \frac{dI_2}{dt} + \frac{1}{C} \int I_2 \, dt + M \frac{dI_1}{dt} = 0,$$

where the coefficients refer to standard circuitry notation. In the above form we have differential-integral equations. By substituting the definition of the current, $I = \frac{dQ}{dt}$, and rearranging terms, we obtain

$$\left(L_1 \frac{d^2}{dt^2} + R_1 \frac{d}{dt}\right) Q_1 + \left(M \frac{d^2}{dt^2}\right) Q_2 = E_o \cos \omega t,$$

$$\left(M \frac{d^2}{dt^2}\right) Q_1 + \left(L_2 \frac{d^2}{dt^2} + R_2 \frac{d}{dt} + \frac{1}{C}\right) Q_2 = 0.$$

The usual method of analysis of this model consists of using the expected values of the parameters to determine a solution. The expected values are obtained by measuring the parameters, such as resistance, etc. These parameter values permit a unique solution to the model.

As the parameter values obtained in this manner are subject to experimental or measurement error, a far more realistic model is obtained by considering the parameters as random variables. A "family of solutions" is generated by considering the density functions of the coefficients of the above set of equations. This family of solutions will allow many additional questions to be answered concerning the probability density function of the "solution".

2.5 Measurement of Radiation by Electronic Counters.

Proportional, Geiger, and scintillation counters are often used to detect X and γ radiation, as well as other charged particles such as electrons and α particles. Design of these counters and their associated circuits depends to some extent on what is to be detected. A device common to all counters is a scaler. This electronic device counts pulses produced by the counter. Once the number of pulses over a measured period of time is known, the average counting rate is obtained by simple division. If the rate of pulses is too high for a mechanical device, it is necessary to scale down the pulses by a known factor before feeding them to a mechanical counter. There are two kinds of scalers[4]: the binary scaler in which the scaler factor is some power of 2, and the decade scaler in which the scaling factor is some power of 10.

A typical binary scaler has several scaling factors ranging from 2^0 to 2^{14}. The scaling circuit is made up of a number of identical "stages" connected in series, the number of stages being equal to n, where 2^n is the desired scaling factor. Each stage is composed of a number of vacuum tubes, capacitors, and resistors, connected so that only one pulse of current is transmitted for every two pulses received. Since the output of one stage is connected to the input of another, this division by two is repeated as many times as there are stages. The output of the last stage is connected to a mechanical counter that will register one count for every pulse transmitted to it by the last stage. Thus, if N pulses from a counter are passed through a circuit of n stages,

4) Cullity, B. D., Elements of X-Ray Diffraction, Addison-Wesley Publishing Company, Reading, Massachusetts, 1956.

only $N/2^n$ will register on the mechanical counter.

Because arrival of X-ray quanta in the counter is random in time, the accuracy of a counting rate measurement is governed by the laws of probability. Two counts of the same X-ray beam for identical periods of time will not be precisely the same because of the random spacing between pulses, even though the counter and scaler are functioning perfectly. Clearly, the accuracy of a rate measurement of this kind improves as the time of counting is prolonged. It is therefore important to know how long to count in order to attain a specified degree of accuracy. This problem is complicated when additional background causes contamination in the counting process. This unavoidable background is due to cosmic rays and may be augmented, particularly in some laboratories, by nearby radioactive materials.

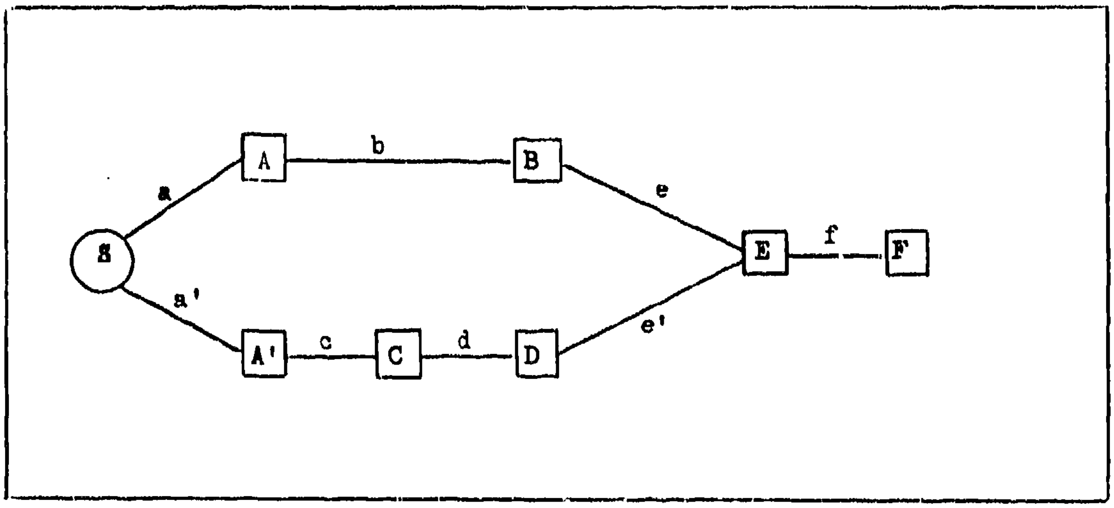
Suppose we want to estimate the diffraction background in the presence of a fairly large unavoidable background. Let N be the number of pulses counted in a given time from a radiation source; let N_b be the number counted in the same time with the radiation source removed. The N_b counts are due to unavoidable background and $(N-N_b)$ to the diffractable background being measured. The relative probable error in $(N-N_b)$ is

$$E_{N-N_b} = \frac{67\sqrt{N+N_b}}{N-N_b} \text{ percent.}$$

Since N and N_b are random variables, the desirability of obtaining the density function of the above quotient form of a random variable is apparent.

2.6 The Pert Model.

Project schedules of many kinds may be schematically presented as a network of nodes and connecting arcs. For illustrative purposes, consider the simple project schedule of Figure 2.



PERT Representation of a Simple Project Network.

Figure 2.

Each arc, a, a', b, . . . , f, in Figure 2 represents an "activity" which must be accomplished during the project schedule. The length of each arc denotes the time required to complete each activity. The nodes, A, A', B, . . . , F, represent events marking the beginning or the end of an activity. There are, of course, two special events or nodes: the initial event, S, and the terminal event, F, between which all project activity is contained.

All activities which must be accomplished during one phase of the project and which may be worked on simultaneously are shown in parallel.

in Figure 2, activities a and a' are such events. In addition, the chain of activities a, b, and e may be worked on concurrently with the series of activities a', b', d, and e'. Activities which may not begin before the termination of others, as f in Figure 2, are shown in series.

For fixed activity durations, a very simple algorithm gives the length of time required for the total project. "PERT" generalizes the method of approach to recognize uncertainties in activity durations by considering them as random variables. The usual assumption is that these durations are described by independent distributions, each with a finite range. Of critical interest in the PERT analysis is the distribution of the random variable describing the project's duration.

When the relevant activity durations are known with certainty, finding the project duration is a trivial matter even for very large networks. Unfortunately, in many space industry research and development projects, the time durations for various activities are known only with a high degree of uncertainty. For this reason, the PERT system was created to facilitate network planning.

The basic data required for PERT are the distributions of the activity durations. The data for these distributions are obtained from technicians who have had some experience with the type of activity involved. The distribution of the project's duration is a composite of these activity duration distributions. For all activities appearing in series, the total activity duration for that series is the sum of the random variables corresponding to each project activity. All activities appearing in parallel are treated by determining the distribution of the maximum times of the activities. As an example, the simple network in Figure 2 is treated in the following manner: Events A, B, and E mark the end of project activities a, b, and e. The total project duration time of these activities is a random variable determined from the sum of the three random variables, a, b, and e. Let U represent this sum so that

$$U = a + b + e.$$

Activities a', c, d, and e' are performed concurrently with the series U. Let V represent this sum of random variables so that

$$V = a' + c + d + e'.$$

It is easily seen that the total project time required for completion of this phase of the project is a random variable determined by the distribution of max $\left[U, V\right]$, the maximum of U and V.

As activity f must begin after the completion of all other events in U and V, the total project duration, T, is of the form

$$T = \max(U, V) + f.$$

Through this formulation, not only the distribution of project duration may be investigated, but certain other equally important topics as well. One such topic concerns the establishment of starting dates for various series of activities. The feasibility of starting the series of activities, say U and V, on the same date may be investigated by considering the random variable $Q = U/V$. This idea suggests that through examination of such probabilities as

$$\Pr(U/V \geq k) \text{ and } \Pr(k' \leq U/V \leq k),$$

the need for rescheduling U or V may be determined.

III. GENERAL THEORETICAL MODELS FOR QUOTIENTS AND PRODUCTS OF RANDOM VARIABLES

The notion of probability plays an important role in statistical theory; yet in a chronological sense, an adequate definition of "the probability of an event" has been subjected to an ensemble of varied approaches. Therefore, a brief mention of the measure-theoretic concept underlying the theorems and definitions presented herein is perhaps warranted. The measure theory approach to probability, popularized by Cramér [22] and others, features the embodiment of the fundamental notion of probability in measure-theoretic ideas through the concept of a theory of sets and the "measure" of a set. The theorems in this section are presented with the implication that their rigorous formulation may be established by utilizing certain measure-theoretic concepts.

3.1 Cumulative Distribution Functions.

In the univariate case, the cumulative distribution function, c.d.f., of a random variable X is defined by the following postulates:

$$\text{if } x_1 < x_2, \quad \text{then } F(x_2) - F(x_1) \geq 0, \tag{3-1}$$

$$F(-\infty) = 0, \; F(+\infty) = 1, \text{in the limit sense}, \tag{3-2}$$

$$\lim_{b \to x+} F(b) = F(x) . \tag{3-3}$$

The notation of (3-2) implies that the limit F(x) exists as $x \to (-\infty)$ or $(+\infty)$. Since (3-1) defines F(x) as being monotonic, it follows that F(x) has at most an enumerable set of discontinuities and that the limits $F(x+)$ and $F(x-)$ exist everywhere. The values of $F(x)$ at discontinuities are fixed by (3-3). It follows from (3-1) and (3-2) that $F(x)$ is non-negative.

The relation between the probability statements about the random variable X and its c.d.f. is expressed by

$$\Pr(X \leq x) = F(x). \tag{3-4}$$

Two important classes of c.d.f.'s may now be characterized:

i) a discontinuous c.d.f., $F(x)$, characterizes a relation such that each member x_i of an at most enumerable set of points $x_1, x_2, \ldots$ is associated with a respective probability $p_i \geqq 0$, such that $\sum_i p_i = 1$, and that the following condition holds:

$$\Pr(X \leq x) = F(x) = \sum_j p_j, \qquad \text{for } x_j \leq x. \tag{3-5}$$

ii) the second important class of c.d.f.'s is characterized by the existence of the function $f(x) \geqq 0$ such that

$$F(x) = \int_{-\infty}^{x} f(\eta)\, d\eta \ . \tag{3-6}$$

Equation (3-6) is referred to as the continuous c.d.f. and $f(x)$ as the probability density function, p.d.f., of random variable X. Obviously then

$$\Pr(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} f(\eta)\, d\eta \ . \tag{3-7}$$

3.1.2 The Bivariate Case: If R denotes a rectangular area in the (x, y) plane, $x_1 < x < x_2$, $y_1 < y < y_2$, and $\Delta_R^2 F(x, y)$ denotes the second difference,

$$\Delta_R^2 F(x,y) = F(x_2,y_2) + F(x_1,y_1) - F(x_1,y_2) - F(x_2,y_1), \tag{3-8}$$

then the joint c.d.f., $F(x, y)$, of random variables X and Y is subjected to the following postulates:

$$\Delta_R^2 F(x, y) \geqq 0 \tag{3-9}$$

and

$$F(-\infty, y) = F(x, -\infty) = 0, \; F(+\infty, +\infty) = 1. \tag{3-10}$$

By allowing $x_1, y_1 \longrightarrow -\infty$ in (3-9), we may conclude, using (3-10), that

$$F(x_1, y_2) - F(x_1, y_1) \geqq 0, \qquad \text{if } y_2 > y_1 . \tag{3-11}$$

Similarly,

$$F(x_2, y_1) - F(x_1, y_1) \geqq 0, \qquad \text{if } x_2 > x_1.$$

From these postulates, it is concluded that $F(x, y)$ is monotonic in each variable and the limits $F(x \pm, y)$, $F(x, y\pm)$ exist everywhere. It is easily shown that $F(x, y)$ is discontinuous on at most an enumerable set of lines, x = constant, (similar results for y). Finally, if $x_1 \longrightarrow -\infty$, and $y_1 \longrightarrow -\infty$ in (3-9), then $F(x, y) \geqq 0$ because of (3-10). The values of $F(x, y)$ at the discontinuities are fixed by (3-12) so that

$$F(x, y) = F(x+, y) = F(x, y+). \quad (3\text{-}12)$$

Again the connection between the probability statements about random variables X and Y and their joint c.d.f. is determined by

$$\Pr(X \leq x, Y \leq y) = F(x, y). \quad (3\text{-}13)$$

In the bivariate case, again the discontinuous and continuous cases are of particular interest.[5]

1. The discontinuous case is characterized by the existence of an at most enumerable set of points (x_i, y_i), $i = 1, 2, \ldots,$ and associated probabilities p_i such that $\sum_i p_i = 1$. So that $F(x,y)$ is

$$F(x,y) = \sum_j p_j, \qquad x_j \leq x, \qquad y_j \leq y. \quad (3\text{-}14)$$

2. The continuous case implies that there exists $f(x,y) \geq 0$ such that

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(\eta_1, \eta_2)\, d\eta_1\, d\eta_2 \quad (3\text{-}15)$$

and

$$\Pr(X, Y \in R) = \int_R \int f(x, y)\, dxdy. \quad (3\text{-}16)$$

5) The mixed case which is treated through application of the Stieltjes integral will not be covered.

3.2 Marginal Distributions.

The marginal distribution of X associated with the joint c.d.f. of random variables X and Y is defined by the relation

$$\Pr(X < x) = P(X < x,\ Y < +\infty) = F(x,\ +\infty). \tag{3-17}$$

Hence $F(x, +\infty)$ is the univariate c.d.f. of X and is called the marginal distribution function of X. Similarly $F(+\infty, y)$ is called the marginal distribution function of Y.

For the discontinuous joint c.d.f.,

$$F(x, +\infty) = \sum_j p_j \quad , \qquad x_j < x, \tag{3-18}$$

and for the continuous case

$$F(x, +\infty) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(\eta_1, \eta_2)\, d\eta_1\, d\eta_2 = \int_{-\infty}^{x} f_1(\eta_1)\, d\eta_1. \tag{3-19}$$

3.3 Statistical Independence.

If $F(x, y)$ is the joint c.d.f. of X and Y, then $F(x) = F(x, +\infty)$ and $F(y) = F(+\infty, y)$ as already noted. The random variables X and Y are said to be statistically independent if and only if

$$F(x, y) = F(x)\ F(y) \tag{3-20}$$

which in turn implies that the following relations hold:

$$f(x, y) = f(x)\ g(y) \text{ for the continuous case} \tag{3-21}$$

and

$$P(X = x,\ Y = y) = \Pr(X = x)\ \Pr(Y = y) \tag{3-22}$$

for the discrete case.

As a result, the following definition may be made:

Definition: The random variables X and Y are stochastically independent if and only if $f(x, y) \equiv f(x)\,g(y)$. If $f(x, y)$ cannot be expressed identically as the product of the marginal distributions, then X and Y are not statistically independent.

3.4 Transformation of Variables.

Change-of-variable integration techniques are used frequently in the study of algebraic combinations of random variables. In the following discussion, only the continuous random variable case is treated as the discrete case is analogous and presents no additional difficulties. At this point, it may be helpful to present a special problem to emphasize the ideas involved.

Consider a random variable X described by the p.d.f.,

$$f(x) = \tfrac{1}{2}x, \qquad 0 < x < 2,$$

$$= 0, \text{ elsewhere.}$$

The random variable X is defined on a set $\mathscr{A} = \{0 < x < 2\}$ where $f(x) > 0$. Define the random variable Y as an algebraic combination of X, say $Y = \frac{1}{64}X^3$, and consider the transformation $y = \frac{1}{64}x^3$. Under this transformation the set $\mathscr{A}$ is mapped into the set $\beta = \{0 < y < \frac{1}{8}\}$; moreover, the transformation is one-to-one. The one-to-one correspondence between the points of $\mathscr{A}$ and β insures that for every event

$0 < a < b < \frac{1}{8}$, the event $a < Y < b$ will occur when and only when the event $4\sqrt[3]{a} < X < 4\sqrt[3]{b}$ occurs.

Thus,

$$\Pr(a < Y < b) = \Pr(4\sqrt[3]{a} < X < 4\sqrt[3]{b})$$

$$= \int_{4\sqrt[3]{a}}^{4\sqrt[3]{b}} \tfrac{1}{2}\, x \, dx \, .$$

By changing the variable of integration in the above so that

$x = 4\sqrt[3]{y}$ and $\frac{dx}{dy} = \frac{4}{3\, y^{2/3}}$, the following result is obtained:

$$\Pr(a < Y < b) = \int_a^b \tfrac{1}{2}\left(4\sqrt[3]{y}\right)\left(\frac{4}{3\, y^{2/3}}\right) dy$$

$$= \int_a^b \frac{8}{3\, y^{1/3}} \, dy \, .$$

The p.d.f. of Y is

$$g(y) = \frac{8}{3\, y^{1/3}}, \quad 0 < y \leq \frac{1}{8},$$

$$= 0, \text{elsewhere.}$$

The change of variable technique in the univariate case is summarized by the following theorem:

Theorem 1: Let X be a continuous random variable described by the p.d.f.,

$f(x)$, defined on the set $\gamma = \{a \leq x \leq b\}$ and let $y = \varphi(x)$ be a monotonic transformation having a unique inverse $x = \varphi^{-1}(y)$ so that under the transformation, the set γ maps by one-to-one correspondence into $\beta = \{a' \leq y \leq b'\}$. Further, let $\varphi'(x)$ exist; then the p.d.f. of Y is given by

$$g(y) = f_X(\varphi^{-1}(y)) \left| \frac{d}{dy} \varphi^{-1}(y) \right|, \qquad \{a' \leq y \leq b'\}. \tag{3-23}$$

3.4.1 Bivariate Case: The method of finding the p.d.f. of one variable of a continuous type may be extended to a function of n random variables. Two-variable transformations are considered now as a special case.

Allow $y_1 = h(x_1, x_2)$ and $y_2 = g(x_1, x_2)$ to be a one-to-one transformation of random variables X_1 and X_2 with existing continuous first partial derivatives. This transformation will map a two-dimensional set γ defined in the (x_1, x_2) plane into a two-dimensional set β defined in the (y_1, y_2) plane. Under these conditions the following theorem holds:

Theorem 2: Let X_1 and X_2 be continuous random variables with joint p.d.f. $f(x_1, x_2)$ defined on the two-dimensional set $\gamma = \{a \leq x_1 \leq b,\ c \leq x_2 \leq d\}$. Let $y_1 = h(x_1, x_2)$ and $y_2 = g(x_1, x_2)$ be a transformation with a unique inverse $x_1 = h^{-1}(y_1, y_2)$ and $x_2 = g^{-1}(y_1, y_2)$ so that the first partial derivatives of the inverse functions exist and the set γ is mapped under the transformation into $\beta = \{a' \leq y_1 \leq b',\ c' \leq y_2 \leq d'\}$ in a one-to-one correspondence. Then the joint p.d.f. of random variables Y_1 and Y_2 is

$$g(y_1,y_2) = f_{X_1,X_2}\,(h^{-1}(y_1,y_2),g^{-1}(y_1,y_2))\ \ |J|\ ,\ \{a' \le y_1 \le b',\ c' \le y_2 \le d'\}, \qquad (3\text{-}24)$$

$$= 0\text{, elsewhere}$$

$$\text{where } J = \left|\frac{\partial(h^{-1}(y_1,y_2),g^{-1}(y_1,y_2))}{\partial(y_1,y_2)}\right| = \begin{vmatrix} \dfrac{\partial\ h^{-1}(y_1,y_2)}{\partial\ y_1} & \dfrac{\partial\ h^{-1}(y_1,y_2)}{\partial\ y_2} \\ \dfrac{\partial\ g^{-1}(y_1,y_2)}{\partial\ y_1} & \dfrac{\partial\ g^{-1}(y_1,y_2)}{\partial\ y_2} \end{vmatrix}.$$

Frequently in the treatment of the two-dimensional case, the random variables X_1, X_2 are treated as being independent so that $f_{X_1,X_2}(\,\bullet\,)$ becomes $f_{X_1}(h^{-1}(y_1,y_2))\cdot f_{X_2}(g^{-1}(y_1,y_2))$. In addition, the transformation function Y_1 is usually set equal to the algebraic combination of interest, say a quotient, so that $Y_1 = X_1/X_2$. The random variable Y_2 is defined as $Y_2 = X_2$. The p.d.f. of $Y_1 = X_1/X_2$, the quotient of interest, is then obtained by determining the marginal p.d.f. of Y_1. This procedure in most cases greatly simplifies the calculations involved in finding a desired p.d.f.

Theorem 2 is essentially a corollary to a more general theorem dealing in several variables. This general theorem is summarized by stating that if a set of transformation functions, $y_i = u_i(x_1,\ldots,x_n)$, $i = 1,\ldots,n$, exists with inverse functions, $x_i = w_i(y_1,y_2,\ldots,y_n)$, and if the conditions of one-to-one transformations and the existence of first partial derivatives hold; then

$$g(y_1,\ldots,y_n) = |J| \left[\phi\,(w_1(y_1,\ldots,y_n),\ldots,w_n(y_1,\ldots,y_n)) \right] \qquad (3\text{-}25)$$

where

$$J = \begin{vmatrix} \frac{\partial w_1}{\partial y_1} & \frac{\partial w_1}{\partial y_2} & \cdots & \frac{\partial w_1}{\partial y_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial w_n}{\partial y_1} & \cdots & \cdots & \frac{\partial w_n}{\partial y_n} \end{vmatrix}$$

As an example of the use of (3-24), we shall consider the following textbook type example:

Allow: $f(x_1,x_2)\, dx_1 dx_2 = \frac{1}{2\pi}\, e^{-\frac{1}{2}(x_1^2 + x_2^2)}\, dx_1 dx_2$ and

consider the transformation:

$$y_1 = \sqrt{x_1^2 + x_2^2}\,,$$

$$y_2 = \tan^{-1} \frac{x_2}{x_1}\,.$$

The domain of the transformation is $y_1 \geq 0$, $0 < y_2 \leq 2\pi$. The inverse of the transformation is

$$x_1 = y_1 \cos y_2\,,$$

$$x_2 = y_1 \sin y_2.$$

Thus,

$$\left|\frac{\partial\,(x_1,x_2)}{\partial\,(y_1,y_2)}\right| = \begin{vmatrix}(\cos y_2) & (-y_1 \sin y_2)\\ (\sin y_2) & (y_1 \cos y_2)\end{vmatrix} = |y_1| = y_1;$$

hence, from (3-24), the joint probability density function of $\{Y_1, Y_2\}$ is

$$f(y_1,y_2) = \frac{1}{2\pi} e^{-\frac{1}{2} y_1^2} y_1 \, dy_1 \, dy_2.$$

3.4.2 <u>Transformations Not One-to-One</u>: Let $f(x_1,x_2, \ldots, x_n)$ be the n-dimensional p.d.f. of continuous random variables $X_1, X_2, \ldots, X_n$. Under the transformation $y_1 = u_1(x_1, x_2, \ldots, x_n)$, $y_2 = u_2(x_1, x_2, \ldots, x_n), \ldots, y_n = u_n(x_1, x_2, \ldots, x_n)$, the n-dimensional space γ where $f(x_1, x_2, \ldots, x_n) > 0$ is mapped into β in the $(y_1, y_2, \ldots, y_n)$ space. Under transformations which are not one-to-one, each point of γ will correspond to one point in β, but to certain points in β there will correspond more than one point in γ. The difficulty presented by this circumstance is diminished in the following manner.

If the set γ may be represented as a union of r mutually exclusive sets $\gamma_1, \gamma_2, \ldots, \gamma_r$ so that the transformation $y_1, y_2, \ldots, y_n$ defines a one-to-one transformation of each subset γ_i into β. Then the groups,

$$\begin{aligned} x_1 &= u_{1i}\,(y_1, y_2, \ldots, y_n),\\ x_2 &= u_{2i}\,(y_1, y_2, \ldots, y_n), \qquad (i = 1, 2, \ldots, r), \qquad (3\text{-}26)\\ x_n &= u_{ni}\,(y_1, y_2, \ldots, y_n),\end{aligned}$$

represent r groups of n inverse functions. If the first partial derivatives exist and are continuous, and if

$$J_i = \begin{vmatrix} \frac{\partial u_{1i}}{\partial y_1} & \frac{\partial u_{1i}}{\partial y_2} & \cdots & \frac{\partial u_{1i}}{\partial y_n} \\ \frac{\partial u_{2i}}{\partial y_1} & \frac{\partial u_{2i}}{\partial y_2} & & \frac{\partial u_{2i}}{\partial y_n} \\ \vdots & \vdots & & \\ \frac{\partial u_{ni}}{\partial y_1} & \frac{\partial u_{ni}}{\partial y_2} & & \frac{\partial u_{ni}}{\partial y_n} \end{vmatrix}, \quad (i = 1, 2, \ldots, r),$$

are not identically zero, the change-of-variable technique may be applied to the union of the r mutually exclusive subsets of $\mathcal{X}$. The n-dimensional p.d.f. of random variables $Y_1, Y_2, \ldots, Y_n$ is then given by

$$g(y_1, y_2, \ldots, y_n) = \sum_{i=1}^{r} |J_i| \; f_{X_1, \ldots, X_n}(u_{1i}(y_1, \ldots, y_n), \ldots, u_{ni}(y_1, \ldots, y_n)). \quad (3\text{-}27)$$

The marginal p.d.f. of any one Y, say Y_1, becomes

$$g_1(y_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(y_2, \ldots, y_n)\, dy_2 \cdots dy_n .$$

A frequently used example will serve to illustrate this result. Let X follow a normal p.d.f. with parameters $(\mu = 0, \sigma^2 = 1)$. Consider the random variable $Y = X^2$. The transformation $y = x^2$ maps the set $\gamma = \{-\infty < x < \infty\}$ into $\beta = \{0 \leq y < \infty\}$. This transformation is not one-to-one, however, as to each point $y \neq 0$ in β there correspond two points, namely $-\sqrt{y}$ and $\sqrt{y}$, in the set γ.

The set γ on which $f(x) > 0$ may be represented as the union of two mutually exclusive sets,

$$\gamma_1 = \{-\infty < x < 0\} \quad \text{and} \quad \gamma_2 = \{0 < x < \infty\} .$$

This is accomplished by observing that in the case of a continuous random variable, the $Pr(X = b) = 0$; we may define the p.d.f. of X at any point b, or for that matter any set of points α with measure zero, i.e., the property that $Pr(X \in \alpha) = 0$, without affecting the distribution of X. Accordingly, two mutually exclusive sets, γ_1 and γ_2, are obtained by defining the normal p.d.f. as

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} x^2}, \quad -\infty < x < 0, \quad 0 < x < \infty, \\ &= 0 , \qquad\qquad\qquad x = 0 . \end{aligned}$$

The function $y = x^2$ now defines a one-to-one transformation which maps each γ_1 and γ_2 into $\beta = \{0 < y < \infty\}$. The two inverse functions of $y = x^2$ are $x = -\sqrt{y}$ and $x = \sqrt{y}$ so that $J_1 = -\frac{1}{2}\sqrt{y}$ and $J_2 = \frac{1}{2}\sqrt{y}$. Using (3-27), we have the well known chi-square distribution withone degree of freedom,

$$g(y) = \left| - \frac{1}{2\sqrt{y}} \right| \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}} + \left| \frac{1}{2\sqrt{y}} \right| \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}},$$

$$= \frac{1}{\sqrt{2\pi}} y^{\frac{1}{2}-1} e^{-\frac{y}{2}}, \quad 0 < y < \infty \quad .$$

The fundamental theorems involving the algebraic product, quotient, sum, and the difference of random variables X_1 and X_2 may now be stated. These theorems are a consequence of (3-24) and (3-27). (See footnote 8, p. 42.)

Theorem 3: The random variables product $Y_1 = X_1 X_2$ will be distributed as

$$\varphi(y_1) = \int_{\beta} f_{X_1,X_2}(y_1/y_2, y_2) \left| \frac{1}{y_2} \right| dy_2, \tag{3-28}$$

$$= 0, \text{elsewhere}$$

provided the random variables X_1 and X_2 are distributed in accordance with the joint p.d.f. $f_{X_1,X_2}(x_1, x_2)$ which has been defined on the set γ so that the transformation $y_1 = u_1(x_1, x_2)$ and $y_2 = u_2(x_1, x_2)$ maps the set γ into the set β in a one-to-one correspondence.

A very important result pertaining to the product distribution of two independent normally distributed variables may be obtained by using Theorem 3. Assume X_1 is $N(0, \sigma_1^2)$ and X_2 is $N(0, \sigma_2^2)$ and defined so that $Pr(X_2 = 0) = 0$. The set $\gamma = \{-\infty < x_1 < \infty;\ -\infty < x_2 < 0,\ 0 < x_2 < \infty\}$ is mapped into the set $\beta = \{-\infty < y_1 < \infty;\ -\infty < y_2 < 0,\ 0 < y_2 < \infty\}$ in a one-to-one correspondence under the transformation $y_1 = x_1 x_2$, $y_2 = x_2$.

By Theorem 3, the p.d.f. of the product $Y_1 = X_1 X_2$ is

$$\varphi(y_1) = \int_\beta \frac{1}{2\pi\,\sigma_1\,\sigma_2\,|y_2|} \exp\left\{ -\tfrac{1}{2}\left(\frac{y_1^2}{y_2^2\,\sigma_1^2} + \frac{y_2^2}{\sigma_2^2} \right) \right\} dy_2 \,.$$

Since $\varphi(y_1)$ is an even function in y_2, $\varphi(y_1)$ may be written as

$$\varphi(y_1) = \frac{1}{\pi\,\sigma_1\,\sigma_2} \int_0^\infty \frac{e^{-\frac{1}{2}\left(\frac{y_1^2}{y_2^2\,\sigma_1^2} + \frac{y_2^2}{\sigma_2^2} \right)}}{y_2}\, dy_2 \,.$$

The following substitution simplifies the above expression. Let

$$u = \frac{y_2^2}{|y_1|}\left(\frac{\sigma_1}{\sigma_2}\right);$$

then $\varphi(y_1)$ becomes

$$\varphi(y_1) = \frac{1}{2\pi\,\sigma_1\,\sigma_2} \int_0^\infty \frac{e^{-\frac{1}{2\,\sigma_1\,\sigma_2}\left\{ |y_1|\frac{1}{u} + u\,|y_1| \right\}}}{u}\, du.$$

Hence, substituting $u = e^t$, $\varphi(y_1)$ may be written as [6]

6) Watson, G. N., Theory of Bessel Functions, Cambridge University Press, Cambridge, 1922, p. 181.

$$\wp(y_1) = \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty} e^{-\frac{|y_1|}{\sigma_1\sigma_2}\left(\frac{e^t + e^{-t}}{2}\right)}dt = \frac{1}{2\pi\sigma_1\sigma_2}\int_{-\infty}^{\infty} e^{-\frac{|y_1|}{\sigma_1\sigma_2}\cosh t}\,dt$$

$$= \frac{1}{\pi\sigma_1\sigma_2} K_o\left(\frac{|y_1|}{\sigma_1\sigma_2}\right), \quad y_1 \neq 0, \qquad (3\text{-}29)$$

where $K_o(\cdot)$ is a modified Bessel function of the second kind of zero order.

This result is well known and is discussed in considerable detail in Chapter IV. A tabulation of $K_o(\cdot)$ is presented in [141] and [142].

Theorem 4: The random variable quotient $Y_1 = X_1/X_2$ will be distributed as

$$\wp(y_1) = \int_{\beta} f_{X_1,X_2}(y_1 y_2, y_2)\,|y_2|\,dy_2 \qquad (3\text{-}30)$$

provided that the random variables X_1 and X_2 are distributed in accordance with the joint p.d.f., $f_{X_1, X_2}(x_1, x_2)$ which is defined on the set γ so that the transformation $y_1 = u_1(x_1, x_2)$, $y_2 = u_2(x_1, x_2)$ maps the set γ into the set β in a one-to-one correspondence.

Theorem 4 may be used to derive the p.d.f. of the ratio of two correlated chi variates[7]. Consider a bivariate normal p.d.f., in

7) Several properties of the bivariate chi p.d.f., have been extensively investigated by P. R. Krishnaiah, et al. [69], [70]. N. L. Johnson [62] and D. J. Finney [34] present certain of the many important applications of this statistic.

random variables X_1 and X_2 with parameters ($\mu_{x_1} = \mu_{x_2} = 0$, $\sigma_{x_1} = \sigma_1$, $\sigma_{x_2} = \sigma_2$, ρ). S. Bose [12] has shown that the statistics

$$s_1 = \frac{U\,\sigma_1}{n} \quad \text{and} \quad s_2 = \frac{V\,\sigma_2}{n},$$

where:

$$U = \left[\sum_{j=1}^{n+1} (x_{1j} - \bar{x}_1)^2 / \sigma_1^2\right]^{1/2}, \quad V = \left[\sum_{j=1}^{n+1} (x_{2j} - \bar{x}_2) / \sigma_2^2\right]^{1/2},$$

are jointly distributed in what is known as the central bivariate correlated chi distribution. P. Krishnaiah, et al. [70], beginning with Bose's expression, have shown that the joint p.d.f. of U and V is

$$f(U,V) = 4\,(1-\rho^2)^{n/2} \sum_{i=0}^{\infty} \frac{\Gamma\left[(n/2)+i\right]\ \rho^{2i}(U\,V)^{n+2i-1}\ \exp\left[-(U^2+V^2)/2\ (1-\rho^2)\right]}{\Gamma(n/2)i!\ \left\{2^{(n/2)+i}\ \Gamma\left[(n/2)+i\right]\ (1-\rho^2)^{n/2+i}\right\}^2}.$$

By Theorem 4, the marginal p.d.f. of $Y_1 = U/V$, under the transformation $y_1 = U/V$, $y_2 = V$, may be derived from (3-31).

Thus

$$f(y_1,y_2) = 4(1-\rho^2)^{n/2} \sum_{i=0}^{\infty} \frac{\Gamma\left[(n/2)+i\right]\rho^{2i}\ (y_1 y_2)^{n+2i-1}(y_2)^{n+2i-1}}{\Gamma\ (n/2)\ \ i!\left\{2^{(n/2)+i}\ \Gamma\left[(n/2)+i\right]\ (1-\rho^2)^{(n/2)+i}\right\}^2}.$$

$$\cdot \exp\left\{-((y_1 y_2)^2 + y_2^2)/2(1-\rho^2)\right\} |y_2|.$$

The marginal p.d.f. of Y_1 satisfies

$$f(y_1) = \frac{4(1-\rho^2)^{n/2}}{\Gamma\left[\frac{n}{2}\right]} \sum_{i=0}^{\infty} \frac{\Gamma\left[(n/2)+i\right] \rho^{2i} \; y_1^{\,n+2i-1}}{i! \left\{ 2^{(n/2)+i}\Gamma\left[(\frac{n}{2})+i\right] (1-\rho^2)^{(n/2)+i} \right\}^2} \,.$$

$$\cdot \int_0^{\infty} y_2^{\,2(n+2i-\frac{1}{2})} \; \exp\left[- y_2^{\,2}(y_1^{\,2}+1)/2(1-\rho^2)\right] \; dy_2 \,. \tag{3-31}$$

With the substitution of $\zeta = y_2^{\,2}$, the integral expression in (3-31) reduces to the Gamma function,

$$\tfrac{1}{2}\int_0^{\infty} \zeta^{\,n+2i-1} \; e^{-a\zeta} \; d\zeta = \frac{\Gamma[k+1]}{2\,a^{k+1}} \,.$$

Since $k = \{n+2i-1\}$ is a positive integer, then $\Gamma(k+1) = k!$. These substitutions simplify (3-31) so that $f(y_1)$ may be expressed as in [12, 69],

$$f(y_1) = \frac{2(1-\rho^2)^{n/2}}{\sqrt{\pi}} \sum_{i=0}^{\infty} \frac{\Gamma\left[(n+2i)\right] \rho^{2i} \; y_1^{\,n+2i-1}}{\Gamma\left[(n/2)+i\right] \; i! \;\; (y_1^{\,2}+1)^{n+2i}} \,.$$

Theorem 5: If random variable X is stochastically distributed in accordance with the p.d.f. $f(x)$ and Y in accordance with the p.d.f. $g(y)$, then the random variable sum $S = X + Y$ will be distributed in accordance with

$$\int_{-\infty}^{\infty} \psi(s)dx = 1, \text{ where } \quad \psi(s) = \int_{-\infty}^{\infty} f(s-y)\ g(y)\ dy. \tag{3-32}$$

The well known fact that the sum of two normally distributed random variables is also normally distributed is shown by using Theorem 5. Let $X = N(\mu_x, \sigma_x^2)$ and $Y = N(\mu_y, \sigma_y^2)$. The random variable $S = X + Y$ is distributed in accordance with (3-32) by Theorem 5. Thus

$$g(s) = \frac{1}{2\pi\sigma_1\ \sigma_2} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2} e^{-\frac{1}{2}\left(\frac{s-x-\mu_y}{\sigma_y}\right)^2} dx. \tag{3-33}$$

It follows with some manipulation that

$$g(s) = \frac{1}{\sqrt{2\pi}\,\sigma^{**}}\ e^{-\frac{1}{2}\left(\frac{y-m^{**}}{\sigma^{**}}\right)^2} \cdot \left\{ \frac{1}{\sqrt{2\pi}\,\sigma^{*}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-m^{*}}{\sigma^{*}}\right)^2} \right\} dx \tag{3-34}$$

where: $m^* = \dfrac{\mu_x \sigma_x + (y-\mu_y)\ \sigma_x^2}{\sigma_x^2 + \sigma_y^2}$, $\quad \sigma^{*2} = \dfrac{\sigma_x^2\ \sigma_y^2}{\sigma_x^2 + \sigma_y^2}$,

$m^{**} = \mu_x + \mu_y$, $\quad \sigma^{**} = \sigma_x^2 + \sigma_y^2$.

The expression in the braces in (3-34) is equal to 1 by the properties of the normal distribution. Therefore, it follows that g(s) is normally distributed with parameters $\mu_s = \mu_x + \mu_y$ and $\sigma_x^2 + \sigma_y^2 = \sigma_s^2$

Theorem 6: If random variable X is distributed in accordance with the p. d. f. f(x) and Y in accordance with the p. d. f. g(y), and if X and Y are statistically independent, then the random variable difference W = X-Y is distributed as

$$\int_{-\infty}^{\infty} R(w)dw = 1,$$

where

$$R(w) = \int_{-\infty}^{\infty} f(w+y)g(y)dy.$$

In Theorems 3-6, the fact that X_1 and X_2 are assumed to be statistically independent allows a convenient integral expression for each algebraic combination of X_1 and X_2 to be obtained. The difficulties are compounded when this assumption is improper. In this case, one is faced with the necessity of obtaining a holomorphic expression for the joint density function of the two random variables.

This difficulty may be illustrated by considering the following example: Let random variable $X_1 = N(\mu_x = \mu_1, \sigma_x^2 = \sigma_1^2)$ and $X_2 = N(\mu_{x_2} = \mu_2, \sigma_x^2 = \sigma_2^2)$ be stochastically dependent. This dependence is characterizied by the coefficient of correlation, ρ. The joint p.d.f. of X_1, X_2 may be expressed as

$$\varphi(x_1,x_2) = \frac{1}{2\pi\sigma_1\ \sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\}.$$

Suppose the p.d.f. of the random variable quotient $Y_1 = X_1/X_2$, where $Pr(X_2 = 0) = 0$, is desired. We may utilize the methods of Section 3.2.4 by defining the transformation,

$$y_1 = x_1/x_2\,,$$
$$y_2 = x_2.$$

The $|J| = |y_2|$ and by Theorem 2, the joint p.d.f. of (y_1,y_2) is

$$\varphi(y_1,y_2) = \frac{|y_2|}{2\pi\sigma_1\ \sigma_2\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho)^2}\left[\left(\frac{y_1\ y_2-\mu_1}{\sigma_1}\right)^2\right.\right.$$

$$\left.\left.-2\rho\left(\frac{y_1\ y_2-\mu_1}{\sigma_1}\right)\left(\frac{y_2-\mu_2}{\sigma_2}\right)+\left(\frac{y_2-\mu_2}{\sigma_2}\right)^2\right]\right\}. \quad (3\text{-}35)$$

By integrating (3-35) with respect to y_2, the marginal p.d.f. of the random variable $Y_1 = X_1/X_2$ is obtained. After much detail, the following cumbersome result is obtained;

$$\varphi(y_1) = \frac{1}{\pi}\,\frac{\sigma_1\,\sigma_2\,\sqrt{1-\rho^2}}{\left(\sigma_2^2\,y_1^2 - 2\rho\,y_1\sigma_1\sigma_2 + \sigma_1^2\right)}\exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{\mu_1^2}{\sigma_1^2} - \frac{2\rho\mu_1\mu_2}{\sigma_1\,\sigma_2} + \frac{\mu_2^2}{\sigma_2^2}\right)\right\} +$$

$$+\exp\left\{-\tfrac{1}{2}\,\frac{(\mu_1 - y_1\,\mu_2)^2}{(\sigma_1^2\,y_1^2 - 2\rho\,y_1\,\sigma_1\sigma_2 + \sigma_1^2)}\right\}\;\frac{\sigma_1(\rho\,\mu_1\sigma_2 - \mu_2\,\sigma_1) + y_1\sigma_2(\rho\,\mu_2\sigma_1 - \mu_1\sigma_2)}{\pi(\sigma_2^2\,y_1^2 - 2\rho\,y_1\,\sigma_1\,\sigma_2 + \sigma_1^2)^{3/2}}\;\cdot$$

$$\cdot\int_0^{\frac{\sigma_1\,(\rho\,\mu_1\,\sigma_2 - \mu_2\,\sigma_1) + y_1\,\sigma_2(\rho\,\mu_2\,\sigma_1 - \mu_1\,\sigma_2)}{\sigma_1\,\sigma_2\{(1-\rho^2)(\sigma_2^2\,y_1^2 - 2\rho\,y_1\,\sigma_1\,\sigma_2 + \sigma_1^2)\}^{1/2}}} e^{-\frac{1}{2}u^2}\,du, \qquad (3\text{-}36)$$

where the integral function is the error integral.

This example serves to illustrate that although in theory the change-of-variable integration techniques are always applicable; they often result in very unwieldy integration problems. In the next chapter, then, some of the more frequently occurring forms of products and quotients of random variables will be investigated. Ways and methods of obtaining the desired marginal distributions which avoid certain integration difficulties will be investigated.

---

[8] Theorems 3-6 are widely known and are frequently quoted in statistical literature. No one person is credited with their derivation. However, Huntington [60] presents a geometric proof for each of these theorems in his paper.

IV. PRODUCT AND QUOTIENT FORMS OF MEASUREMENT ERROR.

4.1 Measurement Error

Physical and economic limitations, the ever present constraints in any industrial environment, often require that the formulation of an estimate of a ratio or of a product of two quantities be based on individual measurements of the two components. It is commoh, for example, to determine the proportion of a reactant which has reacted in a chemical process by measuring the residue after the process has completed, or to estimate the efficiency of a rocket engine by determining the ratio of fuel energy input to thrust output. Estimates of efficiency, velocity, and proportions are examples of indices which are necessarily ratios. Product forms frequently arise in assessing the probability of a successful event, such as the launching of a missile, which depends on the successful operation of several independent component subevents.

As is frequently the case, further complications arise when the individual measurement of each component of the product or ratio is subject to error. This error most often is due either to what in reality are true random fluctuations in the measurable quantity or to instrument error caused by the inherent limitations of an imperfect measuring device. These two sources of error are often treated alike statistically and are called measurement error. When the two quantities are subject to measurement error, their product or quotient is distributed about its expected value. In the following, two examples are presented which demonstrate two commonly occurring situations which require the use of approximate methods to establish the precision of an estimate of measurement error.

Whenever economically feasible, the distribution of error of each component quantity may be determined from repeated experimentation. On other occasions, repeated trials are impossible or undesirable; whereupon ,the precision of a specific trial must be established by theoretical arguments. This requirement is frequently necessary in the acceptance trials of certain types of heavy industrial machinery such as steam boilers, petroleum processing equipment, blast furnaces, generators, large rocket engines, etc. Such trials are lengthy and expensive processes which usually require that an estimate of the efficiency of the equipment be made from a single trial operation. It is necessary from the point of establishing confidence limits to obtain the probability density function of the estimate of efficiency. Approximate methods, such as the assumption of normality,must suffice for this purpose since more desirable exact methods are clearly unattainable.

The second situation involves instrument error. Measurements taken with several different measuring devices, all of a specific type, generally are different. Repeated measurement taken with any one of these instruments is physically possible, yet is often pointless; the instrument has been manufactured to certain tolerances and is incapable of recording with a greater or lesser degree of precision on any successive trial. Other measuring devices will record with the same degree of precision on any specified trial but in most cases slight differences will be noted in the expected values and dispersion of measurements taken with the individual instruments. The experimenter will usually have at his disposal only one or two measurement instruments of a specified type; thus he will be unable to conduct a large number of experiments with which to calibrate his own equipment. As a result, the tolerance associated with a measurement accomplished with a "typical" measuring instrument is best considered as being random in nature and should be treated from a statistical point of view.

In many industrial applications, the errors of measurement are considered to be described by one of three p.d.f.'s: the rectangular, the triangular, or the normal density function.

Instrument error is often approximated by the rectangular p.d.f. Certain measurements such as length, time, and weight are usually recorded in terms of deviations from some preselected value, X'. These deviations $y = (x_i - x')$ are assumed to be in an interval $S = \{a \leq y \leq b\}$ so that all deviations are equally likely. It is implied that the deviation between the true value of the measurement and the preselected value is equally likely to be anywhere in the interval S. A measurement which has been rounded off from a more precise measurement is an example of an estimate which is subject to rectangular error.

The triangular p.d.f. describes measurement error resulting from summing or taking the difference of two readings which are subject to rectangular measurement error.

The p.d.f. describing sampling error and the distribution of the expected value of a sample of measurements is often taken to be normal or Gaussian. Product and quotient forms of these three common p.d.f.'s describing measurement error will be discussed in the following sections. These three r.v.'s will be denoted in an obvious notation as R, T, and N hereafter.

4.2 The Normal Density Function.

4.2.1 The Product of Two Normally Distributed Random Variables: A wide variety of interesting approaches and techniques has been applied to the problem of finding the p.d.f. of a product of n normally distributed random variables. In fact, as a collective group, these investigations represent a rather extensive study. An interest in this problem, as indicated by published articles, first became apparent in the early 1930's, with the bulk of important derivations occurring only a few years later.

In certain cases, the quadrature method used in treating the example under Theorem 3, may be used to derive an expression for $Y_1 = N_1 N_2$, the product of two normally distributed random variables. These special cases involve the dependent and independent cases of N_1 and N_2 described by the normal p.d.f.'s with parameters N(0, 1) or $N(0, \sigma_1^2)$ and $N(0, \sigma_2^2)$. These four results are enumerated in Table I Appendix A.

The dependent and independent cases of $Y_1 = N_1 N_2$ in which N_1, N_2 are normally distributed with arbitrary means ($\mu_i \neq 0$) and variances do not readily lend themselves to the ordinary integral methods, and as a result have undergone extensive investigation.

A chronological history of the important results in the study of this problem is outlined in order to suggest certain "difficulties" in the applications of these results.

In 1932, Wishart and Bartlett [133] considered the problem of determining the p.d.f. of the product $(N_1 N_2)$. The problem was posed in the following framework.

Let X_1 and X_2 be two normally distributed random variables, $N_1(0, \sigma_1^2)$, $N_2(0, \sigma_2^2)$, which satisfy the joint p.d.f.,

$$\varphi(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x_1^2}{\sigma_1^2} - \frac{2\rho x_1 x_2}{\sigma_1\sigma_2} + \frac{x_2^2}{\sigma_2^2}\right)\right\} \tag{4-1}$$

The characteristic function of the product $X_1 X_2$ is by definition $E\left[e^{it X_1 X_2}\right]$, or

$$\phi_{X_1X_2}(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{it X_1X_2} \varphi(x_1, x_2)\, dx_1\, dx_2$$

$$= \left\{1-2i\rho\sigma_1\sigma_2 t + (1-\rho^2)\sigma_1^2\sigma_2^2 t^2\right\}^{½} \tag{4-2}$$

The generating function of the semi-invariants of $y_1 = X_1X_2$ is[9]

$$K = -½ \log\left\{1-it\sigma_1\sigma_2(\rho+1)\right\} -½ \log\left\{1-it\sigma_1\sigma_2(\rho-1)\right\}$$

9) The semi-invariants were so named by Thiele and were later called cumulants by Fisher. If the logarithm of the characteristic function (as in (4-2)), a function of t, is expanded in a power series of (it) which converges for some range of t containing the origin as an interior point, the semi-invariants k_v are the coefficients of $(it)^v/v!$ in the expansion. A simple relation exists between the $k_v(y_1)$ and the moments of Y_1; the first two being $k_1(y_1) = E\left[Y_1\right]$, $k_2(y_1) = \mathrm{Var}\left[Y_1\right]$.

The semi-invariants of y_1, $k_v(y_1)$, are the coefficients of $(it)^v/v!$ in the power series expansion of K. Thus

$$k_v(y_1) = \tfrac{1}{2}(v-1)!\ \sigma_1^{\,v}\ \sigma_2^{\,v}\ \left\{(\rho+1)^v + (\rho-1)^v\right\}. \qquad (4-3)$$

The moments of the product Y_1 may be investigated through the application of (4-3).

In the case of a sum of independent products taken from the joint p.d.f. $\varphi(x_1, x_2)$, denoted by $\Sigma = y_1^{(1)} + y_1^{(2)} + \ldots + y_1^{(n)}$, this relation,

$$K_\Sigma = n\,K,$$

holds so that from (4-3), the semi-invariants of $\left(\frac{\Sigma}{n}\right)$ may be obtained from

$$k_v\left(\frac{\Sigma}{n}\right) = \tfrac{1}{2}\,(v-1)!\ \sigma_1^{\,v}\ \sigma_2^{\,v}\ n^{1-v}\ \left\{(\rho+1)^v + (\rho-1)^v\right\}.$$

The relation of $\phi_\Sigma\ (t)$ to $f(\Sigma)$ satisfies

$$\phi_\Sigma\ (t) = \left[\phi_{Y_1}(t)\right]^n = \left\{1-2\,i\,\rho\ \sigma_1\ \sigma_2\ t + (1-\rho^2)\ \sigma_1^{\,2}\ \sigma_2^{\,2}\ t^2\right\}^{-\frac{1}{2}n}$$

$$= \int_{-\infty}^{\infty} e^{it\Sigma}\ f(\Sigma)\ d\Sigma\ .$$

By inverting this characteristic function, it follows that

$$f(\Sigma) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{e^{-it\Sigma}}{\left\{1-2i\ \rho\ \sigma_1\ \sigma_2\ t + (1-\rho^2)\ \sigma_1^{\,2}\ \sigma_2^{\,2}\ t^2\right\}^{\frac{1}{2}n}}\ dt. \qquad (4-4)$$

When $n = 1$, then $\sum = y_1$, so that $f(\Sigma)\,d\Sigma = \varphi(y_1)dy_1$ and (4-4) becomes

$$\varphi(y_1) = \frac{(1-\rho^2)^{\frac{1}{2}n} e^{\rho y_1}}{2\pi} \int_{\rho i - \infty}^{\rho i + \infty} \frac{e^{iy_1 w}}{(1+w^2)^{\frac{1}{2}n}} dw , \qquad (4\text{-}5)$$

after these substitutions are made,

$$(1-\rho^2)\, \sigma_1 \sigma_2 t = \rho i - w, \ (1-\rho^2)\, \sigma_1 \sigma_2 y_1 = \Sigma \quad .$$

From the theory of the complex variable, it is clear that the contour may be deformed into a real axis as $|\rho| < 1$; thus (4-5) may be written as

$$\varphi(y_1) = \frac{(1-\rho^2)^{\frac{1}{2}n} e^{\rho y_1}}{\pi} \int_0^{\infty} \frac{\cos y_1 w}{(1+w^2)^{\frac{1}{2}n}} dw$$

$$= \frac{(1-\rho^2)^{\frac{1}{2}n} e^{\rho y_1}}{\pi^{\frac{1}{2}} 2^{\frac{1}{2}(n-1)} \Gamma(\frac{1}{2}n)} \left\{ y_1^{\frac{1}{2}(n-1)} K_{\frac{1}{2}(n-1)}(y_1) \right\} , \ -\infty < y_1 < \infty \qquad (4\text{-}6)$$

where $K_{\frac{1}{2}(n-1)}(y_1)$ is a modified Bessel function of the second kind of $\left[\frac{1}{2}(n-1)\right]$ order. For $n = \sigma_x = \sigma_y \doteq 1$ and $r = 0$ the p.d.f. of y_1 reduces to

$$\varphi(y_1) = \frac{1}{\pi} K_0 \quad (|y_1|) , \qquad y_1 \neq 0 . \qquad (4\text{-}7)$$

Wishart and Bartlett's method may be used to show that Z, the product of two normal variables where $Z = \frac{x_1}{\sigma_1} \cdot \frac{x_2}{\sigma_2}$ is described by

$$\varphi(z) = \frac{1}{\pi} K_0(z)$$

which possesses a singularity at $z = 0$.

Shortly after the publication of this result, P. T. Yuan [134] demonstrated that if X_1 and X_2 are independently and lognormally distributed, the product $Z = (X_1 - a)(X_2 - b)$, where a and b are the upper (lower) limits of the range of X_1 and X_2, is distributed as (4-7).

The analysis of C. C. Craig [20], [21], following in 1936, is perhaps the most notable concerning the product of two normally distributed random variables. Craig considered the bivariate normal p.d.f. with parameters (μ_{x_1}, μ_{x_2}, σ_{x_1}, σ_{x_2} and coefficient of correlation, ρ). By Theorem 2, this joint normal p.d.f., under the transformation $w = x_1 x_2$, $y_1 = y_2$, may be written as

$$\varphi(w, y_2) = \frac{1}{2\pi\sigma_{x_1}\sigma_{x_2}\sqrt{1-\rho^2}\ |y_2|} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{w/y_2 - \mu_{x_1}}{\sigma_{x_1}}\right)^2 - 2\rho\left(\frac{w/y_2 - \mu_{x_1}}{\sigma_{x_1}}\right)\left(\frac{y_2 - \mu_{x_2}}{\sigma_{x_2}}\right) + \left(\frac{y_2 - \mu_{x_2}}{\sigma_{x_2}}\right)^2\right]\right\}.$$

In an effort to generalize this result, Craig introduced

$V_1 = \dfrac{\mu_{x_1}}{\sigma_{x_1}} = \dfrac{\mu_1}{\sigma_1}$ and $V_2 = \dfrac{\mu_{x_2}}{\sigma_{x_1}} = \dfrac{\mu_2}{\sigma_2}$, the reciprocals of the coefficients of variation, and considered the transformation $y_1 = \dfrac{x_1\ x_2}{\sigma_1\ \sigma_2} = \dfrac{w}{\sigma_1\ \sigma_2}$,

$x = \frac{x_1}{\sigma_1}$. The marginal p.d.f. of y_1, under this transformation, may be expressed as the difference of two integrals. Thus,

$$\varphi(y_1) = I_1(y_1) - I_2(y_1), \tag{4-8}$$

in which

$$I_1(y_1) = \frac{e^{-\frac{v_1^2 - 2\rho v_1 v_2 + v_2^2}{2(1-\rho^2)}}}{2\pi\sqrt{1-\rho^2}} \; \frac{\rho y_1}{1-\rho^2} \int_0^\infty \exp\left\{ -\frac{1}{2(1-\rho^2)}\left(x^2 + \frac{y_1^2}{x^2}\right) + \right.$$

$$\left. + \frac{1}{1-\rho^2}\left[(v_1 - \rho v_2)\, x + (v_2 - \rho v_1)\frac{y_1}{x}\right]\right\} \frac{dx}{x},$$

and $I_2(y_1)$ is the integral of the same function defined on the interval $(-\infty, 0)$.

In an effort to simplify any numerical calculation, Craig reformulated (4-9) as an infinite series. The infinite series expansion is derived from the joint p.d.f. of variables U and Z, where

$$\frac{x}{\sqrt{1-\rho^2}} = u \quad \text{and} \quad \frac{y_1}{1-\rho^2} = \frac{x_1 \, x_2}{\sigma_1 \, \sigma_2 (1-\rho^2)} = z.$$

Under this transformation, Equation (4-8) becomes

$$\phi(z) = \frac{\sqrt{1-\rho^2}}{2\pi} \exp \left[- \frac{v_1^2 - 2\rho v_1 v_2 + v_2^2}{2(1-\rho^2)} + \frac{\rho z}{(1-\rho^2)^2} \right] \int_0^\infty e^{-\frac{1}{2}\left(u^2 + \frac{z^2}{u^2} \right)} \cdot$$

$$\cdot \exp \left\{ \frac{v_1 - \rho v_2}{\sqrt{1-\rho^2}} u + \frac{v_2 - \rho v_1}{\sqrt{1-\rho^2}} \cdot \frac{z}{u} \right\} \frac{du}{u}. \qquad (4\text{-}9)$$

The term $\exp \left\{ \dfrac{\frac{v_1 - \rho v_2}{\sqrt{1-\rho^2}} u + \frac{v_2 - \rho v_1}{\sqrt{1-\rho^2}} \frac{z}{u}}{u} \right\}$ may be expanded in a Laurent series in powers of u for all u, $u \neq 0$. This expansion is simplified to some extent by substituting

$$\frac{v_1 - \rho v_2}{\sqrt{1-\rho^2}} = R_1 \quad \text{and} \quad \frac{v_2 - \rho v_2}{\sqrt{1-\rho^2}} = R_2 .$$

In the expansion, the coefficient of u^{r-1}, $r \geq 1$, is $\frac{R_1^{\,r}}{r!} \sum_r (R_1 R_2 z)$, in which $\sum_r(\cdot)$, the confluent hypergeometric function of order r, is[10]

$$\sum\nolimits_r (R_1 R_2 z) = 1 + \frac{R_1 R_2 z}{r+1} + \frac{(R_1 R_2 z)^2}{(r+2)^{(2)} 2!} + \frac{(R_1 R_2 z)^3}{(r+3)^{(3)} 3!} + \ldots ,$$

with $(r+k)^{(k)} = (r+k)(r+k-1) \ldots (r+1)$.

By this expansion, the p.d.f. of $Z = X_1 X_2 / \sigma_1 \sigma_2(1-\rho^2)$ may be expanded in an infinite series involving confluent hypergeometric functions

10) These functions are discussed in detail in Whitaker, E. T., and G. N. Watson, A Course in Modern Analysis, Cambridge University Press, Cambridge, 1958.

and powers of z, v_1, and v_2. This series is

$$\wp(z) = \sqrt{\frac{1-\rho^2}{\pi}} \exp\left\{-\frac{v_1^2 - 2\rho v_1 v_2 + v_2^2}{2(1-\rho^2)} + \frac{\rho z}{(1-\rho^2)^2}\right\}\left[\sum\nolimits_0 (R_1 R_2 z) K_o(z) + \right. \quad (4\text{-}10)$$

$$+ (R_1^2 + R_2^2) \frac{|z|}{2!} \sum\nolimits_2 (R_1 R_2 z) K_1(z) + (R_1^4 + R_2^4) \frac{z^2}{4!} \sum\nolimits_4 (R_1 R_2 z) K_2(z) +$$

$$+ (R_1^6 + R_2^6) \frac{|z|^3}{6!} \sum\nolimits_6 (R_1 R_2 z) K_3(z) + \dots ,$$

where: $K_i(z)$ = the Bessel function of the second kind of the ith order and with argument z, and

$$\sum\nolimits_j (R_1 R_2 z) = \frac{j!}{R_1^j} \left(\frac{R_1}{R_2 z}\right)^{\frac{j}{2}} \quad I_j \left(2\sqrt{v_1 v_2 z}\right) \text{ in which}$$

$I_j(\cdot)$ is the Bessel function of the first kind of the jth order.

When $v_1 = v_2 = \rho = 0$, the p.d.f. of $Z = X_1 X_2 / \sigma_1 \sigma_2$ is the simple Bessel function result expressed by (4-7).

Craig's results have unfortunately proved to be of little use computationally for it may be shown that for large v_1 and v_2 the series expansion (4-9) converges very slowly; in fact for v_1 and v_2 as small as 2, the expansion is unwieldy. Yet after this publication, many investigators consider the problem to be "solved" although the convergence problem exists. Others have been critical of this "solution", even

Craig himself who in 1942, [21], stated that even though his result is a mathematical solution to the problem, it falls far short of what is required for numerical computation.

Using Craig's formula for the cumulant generating function, J.B.S. Haldane in 1942 presented the moments (about the mean and the origin) and the cumulants of various products and powers of normally distributed random variables. He considered the correlated and uncorrelated cases of these statistics: a) the cube of a normal variable, b) an arbitrary power of a normal variate, and c) the product of n normal variates. These expressions are rather cumbersome however and their application in an applied problem requiring an extensive amount of numerical computation would be severely limited. This fact combined with the realization that Craig's series is subject to restrictive conditions prompted others to examine this product more from the viewpoint of establishing its analytical properties. For instance, B. Epstein [31] and I. Kotlarski [68] independently showed that the Mellin transform is a very useful analytical tool in examining the integral equation

$$\varphi(y_1) = \int_{\beta} f_{X_1}(y_1/y_2)\, g_{X_2}(y_2)\, dy_2 .$$

Their investigation are closely related to problems of the nature of those discussed in Chapter VI.

In 1947, L. Aroian [2] took up the problem of convergence in Craig's series expansion. Using Craig's notation, he showed that as v_1 and $v_2 \longrightarrow \infty$, the p.d.f. of Y_1 approaches the normal p.d.f. In addition, he demonstrated that the Type III function and the Gram-Charlier type A series afford excellent approximations to the distribution of Y_1 when $\rho = 0$.

Using the properties of the moment generating function of the $\wp(y_1)$, it is possible to show that $E[y_1] = \bar{y}_1 = V_1 V_2 + \rho$ and the standard deviation is $\sigma_{y_1} = \sqrt{V_1^2 + V_2^2 + 2\rho V_1 V_2 + 1 + \rho^2}$.

Aroian proved the following statements in the form of theorems:

1) The p.d.f. of Y_1 approaches the normal p.d.f. with mean $\bar{y}_1$ and variance $\sigma_{y_1}^2$ as V_1 and $V_2 \longrightarrow \infty$ (or $-\infty$) in any manner whatsoever, provided $-1 + \epsilon < \rho < 1$, $\epsilon > 0$.

2) The p.d.f. of y_1 approaches the normal p.d.f. with mean $\bar{y}_1$ and variance $\sigma_{y_1}^2$ if $V_1 \longrightarrow \infty$, $V_2 \longrightarrow -\infty$, provided $-1 \leq \rho < 1 - \epsilon$, $\epsilon > 0$

and

3) The p.d.f. of y_1 approaches $N(\bar{y}_1, \sigma_{y_1}^2)$ if V_1 remains constant and $V_2 \longrightarrow \infty$, $-1 + \epsilon < R \leq 1$, $\epsilon > 0$; or if V_1 remains constant and $V_2 \longrightarrow -\infty$ for $-1 \leq R < 1 - \epsilon$, $\epsilon > 0$.

Aroian demonstrated the close approximations to $\wp(y_1)$ by the Type III function and the Gram-Charlier type A series by numerically integrating Craig's expression (4-8) for the special case: $V_1 = 0$, $V_2 = 10$, and $\rho = 0$. A brief tabular comparison is presented in his article.

In Table X , Appendix A, the numerical integration of Craig's formula (4-8) for a few special cases is presented[11]. These are: (1) $V_1 = V_2 = \rho = 0$, (2) $V_1 = V_2 = ½$, $\rho = 0$ and (3) $V_1 = 1(0)$, $V_2 = 0(1)$, $\rho = 0$.

4.2.2 <u>The Quotient of Normally Distributed Random Variables</u>: The first investigations of the properties of the p.d.f. of a quotient of two normally distributed random variables were directed toward characterizing the quotient's properties in terms of the properties of the component variables.

K. Pearson's study [97] in 1910 of an opsonic index formed by the quotient of two normally distributed random variables represents the first published investigation of this problem. He succeeded in obtaining the first four moments of $Y_1 = X_1/X_2$ in terms of the moments of X_1 and X_2. Unfortunately, he found that they were "practically unworkable if X_1 and X_2 are correlated as we should have to find the third and fourth order product moments".

Later, C. C. Craig [20] , (1929), developed this approach by

11) A number of other cases are to be included in a forthcoming research report to be released by the Applied Mathematics Research Laboratory, Aerospace Research Laboratories.

deriving the moments of Y_1 in terms of the semi-invariants of the components. Craig demonstrated the advantages of using the semi-invariants (as opposed to the moments) by constructing the moments of Y_1 in the case of correlated X_1 and X_2 with relatively little difficulty. The resulting expressions are cumbersome by present day standards and are difficult to apply.

In this time period, the hypothesis that Y_1 must be near-normally distributed given that the components X_1 and X_2 are distributed normally was strongly supported. Investigations of Merrill [86] in 1928 and by Geary [38] , 1930, were instrumental in disproving this conjecture. Merrill's investigation, by graphical approximation, showed that when the correlation between X_1 and X_2 is high and the coefficients of variation $\delta_{x_i} = \dfrac{\sigma_{x_i}}{\mu_{x_i}}$, $i = 1, 2$, are large, there is a considerable deviation from the normal p.d.f. Geary established this result on a more rigorous foundation by formulating what is now a widely known approximation. He considered the problem as formulated in this manner:

Let X_1 and X_2 be two jointly distributed normal random variables with the p.d.f.,

$$g(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right)\right\}, \tag{4-11}$$

where: $\mu_{x_1} = \mu_1$, $\mu_{x_2} = \mu_2$, $\sigma_{x_1} = \sigma_1$, $\sigma_{x_2} = \sigma_2$ and ρ is the coefficient of correlation.

When the expected values $\mu_1 = \mu_2 = 0$, the $g(x_1, x_2)$ reduces to

$$g(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x_1^2}{\sigma_1^2} - \frac{2\rho x_1 x_2}{\sigma_1 \sigma_2} + \frac{x_2}{\sigma_2}\right)^2\right\}. \quad (4\text{-}12)$$

Consider the random variable Y_1, where

$$y_1 = \frac{b + x_2}{a + x_1} \quad (4\text{-}13)$$

and a and b are constants; then the function

$$t = \frac{a\, y_1 - b}{\sqrt{\sigma_1^2\, y_1^2 - 2\rho\,\sigma_1\,\sigma_2\, y_1 + \sigma_2^2}} \quad (4\text{-}14)$$

is approximately normally distributed with expected value $\bar{t} = 0$ and standard deviation $\sigma_t = 1$, provided that $(a + x_2)$ is unlikely to assume negative values. Geary shows that this latter condition is satisfied if $(a \geq 3\,\sigma_1)$, i.e., the coefficient of variation of $(a + x_1) \leq \frac{1}{3}$.

Geary's and Merrill's results gave definite proof that the hypothesis of the "normally distributed index" was incorrect in many cases and have lead to the problem termed "characterizing properties of a p.d.f." which is discussed in Chapter VI.

It is advantageous at this point to demonstrate that the integral methods discussed in Chapter III may be used to obtain an expression for the p.d.f. of Y_1 in order to lay further groundwork for discussing other known results.

In the independent case, the joint p.d.f. of X_1 and X_2 is of the form

$$g(x_1, x_2) = f_1(x_1)\, f_2(x_2). \tag{4-15}$$

When considering the random variable $Y_1 = X_1/X_2$, it is necessary to use the methods described in Section 3.4.2 in order to define a one-to-one transformation. By specially defining $\Pr(X_2 = 0) = 0$, the inverse functions $x_1 = y_1/y_2$ and $x_2 = y_2$ map the set

$\mathcal{Y} = \{-\infty < x_1 < \infty;\ -\infty < x_2 < 0,\ 0 < x_2 < \infty\}$ into the set

$\mathcal{B} = \{-\infty < y_1 < \infty;\ -\infty < y_2 < 0,\ 0 < y_2 < \infty\}$ where the joint p.d.f. $g(y_1, y_2) > 0$. The $|J| = |y_2|$ and the marginal p.d.f. of Y_1 is, by equation (2-27),

$$g(y_1) = \lim_{C \to 0^-} \int_{-\infty}^{C} \frac{-y_2}{\sqrt{2\pi}\,\sigma_1} \exp\left\{-\frac{1}{2\sigma_1^2}\, y_1^2\, y_2^2\right\} \frac{1}{\sqrt{2\pi}\,\sigma_2} \exp\left\{\frac{-y_2^2}{2\sigma_2^2}\right\} dy_2 + \tag{4-16}$$

$$+ \lim_{C \to 0^+} \int_{C}^{\infty} \frac{y_2}{\sqrt{2\pi}\,\sigma_1} \exp\left\{-\frac{1}{2\sigma_1^2}\, y_1^2\, y_2^2\right\} \frac{1}{\sqrt{2\pi}\,\sigma_2} \exp\left\{\frac{-y_2^2}{2\sigma_2^2}\right\} dy_2 .$$

Equation (4-16) is directly integrable, provided the limits exist, so that

$$g(y_1) = \frac{\sigma_1 \sigma_2}{\pi(y_1^2 \sigma_2^2 + \sigma_1^2)}, \quad y_1 \neq 0. \tag{4-17}$$

The more complicated result, $\mu_1 \neq \mu_2 \neq 0$ of $g(y_1)$ may also be obtained by applying the methods of Section 3.4.2. This result was first shown by Fieller [33] and Baker [4] in 1932.[12] It is

$$g(y_1) = \frac{1}{\pi} \frac{\sigma_1 \sigma_2 \sqrt{1-\rho^2}}{(\sigma_2^2 y_1^2 - 2\rho y_1 \sigma_1 \sigma_2 + \sigma_1^2)} \exp\left\{- \frac{1}{2(1-\rho^2)} \left(\frac{\mu_1^2}{\sigma_1^2} \frac{-2\rho\mu_1\mu_2}{\sigma_1 \sigma_2} + \frac{\mu_2^2}{\sigma_2^2}\right)\right\} +$$

$$+ \exp\left\{-\tfrac{1}{2} \left(\frac{(\mu_1 - y_1 \mu_2)^2}{\sigma_2^2 y_1^2 - 2\rho y_1 \sigma_1\sigma_2 + \sigma_1^2}\right)\right\} \cdot \frac{\sigma_1(\rho \mu_1 \sigma_2 - \mu_2 \sigma_1) + y_1 \sigma_2(\rho\mu_2 - \mu_1 \sigma_2)}{\pi(\sigma_2^2 y_1^2 - 2\rho y_1 \sigma_1 \sigma_2 + \sigma_1^2)^{3/2}}$$

$$\int_0^{\frac{\sigma_1(\rho \mu_1 \sigma_2 - \mu_2 \sigma_1) + y_1 \sigma_2(\rho \mu_2 \sigma_1 - \mu_1 \sigma_2)}{\sigma_1 \sigma_2 (1-\rho^2)(\sigma_2^2 y_1^2 - 2\rho y_1 \sigma_1 \sigma_2 + \sigma_1^2)^{1/2}}} e^{-\frac{1}{2} u^2} \, du \qquad (4\text{-}18)$$

We see that if $\mu_1 = \mu_2 = \rho = 0$, $g(y_1)$ reduces to Equation (4-17); and under the conditions $\mu_1 = \mu_2 = 0$, $\rho \neq 0$, $g(y_1)$ reduces to

$$g(y_1) = \frac{\sqrt{1-\rho^2} \quad \sigma_1 \sigma_2}{\pi(\sigma_2^2 y_1^2 - 2 \rho \sigma_1 \sigma_2 y_1 + \sigma_1^2)}, \quad y_1 \neq 0. \qquad (4\text{-}19)$$

The calculation of values of the c.d.f. of y_1 is accomplished merely by

---

12) This result is generally attributed to E. C. Fieller.

the use of a set of tables of arc-tangent in these two cases.

Fieller's paper is mainly devoted to obtaining, by quadrature, an expression which will simplify the numerical calculations associated with equation (4-18). His rather complicated quadrature method, formulated in terms of existing tabulated functions of the bivariate normal density function, was derived under the following hypothesis.

Consider the ratio $y_1 = x_1/x_2$ where random variables X_1 and X_2 are described by the joint p.d.f. of equation (4-11). The points (x_1, x_2) corresponding to a given value of y_1 lie on the line

$$x_1 = y_1 x_2 . \tag{4-20}$$

Thus, the probability that an arbitrary element of $g(x_1,x_2)$ will have an index $\{v_1 < y_1 < v_2\}$ is equal to the volume of the portion of $g(x_1, x_2)$ which lies above the area swept out in the (x_1, x_2) plane by the line $x_1 = y_1 x_2$ as it revolves from

$$x_1 = v_1 x_2$$

to

$$x_1 = v_2 x_2 .$$

Taking $v_1 = -\infty$, this probability is the probability that $y_1 \leq v_2$ and is calculated from

$$V = \int_0^{\infty}\int_{-\infty}^{v_2x_2} g(x_1,x_2)\ dx_1\ dx_2 + \int_{-\infty}^{0}\int_{v_2x_2}^{\infty} g(x_1,x_2)\ dx_1\ dx_2. \tag{4-21}$$

When $g(x_1, x_2)$ is of the form (4-11) and each variable is expressed in terms of deviations from its respective mean, V may be written as[13]

$$V = \iint\limits_{a+b} \frac{1}{2\pi\sigma_1 \sigma_2\sqrt{1-r^2}} \exp\left\{-\frac{1}{2(1-r^2)}\left[\frac{x_2^2}{\sigma_2^2} - \frac{2rx_1x_2}{\sigma_1\sigma_2} + \frac{x_1^2}{\sigma_1^2}\right]\right\} dx_1 dx_2 \tag{4-22}$$

where a and b are two portions of the (x_1, x_2) plane indicated in Figure 3. The boundaries of the a and b portions are the lines

$$x_2 + \bar{x}_2 = 0,$$
$$x_1 + \bar{x}_1 = y_1(x_2 + \bar{x}_2).$$

If the transformations $x_2 = \alpha$ and $x_1 - y_1 x_2 = \eta$ are made, the portions γ and β of the (α, η) plane which correspond to a and b in Figure 3 are bounded by the lines

$$\alpha + \bar{x}_2 = 0,$$
$$\eta + \bar{x}_1 - y_1 \bar{x}_2 = 0.$$

This area is shown in Figure 4.

Therefore if this change of variable is performed, V becomes

$$V = \iint\limits_{\gamma+\beta} \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} e^{-\frac{1}{2}\chi^2} \, d\alpha \, d\eta \tag{4-23}$$

where $\chi^2 = \frac{1}{1-r^2}\left\{\frac{\alpha^2}{\sigma_2^2} - 2r\frac{(\alpha)(\eta + y_1\alpha)}{\sigma_1\sigma_2} + \frac{(\eta + y_1\alpha)^2}{\sigma_1^2}\right\}.$

---

[13] In this derivation r denotes the coefficient of correlation.

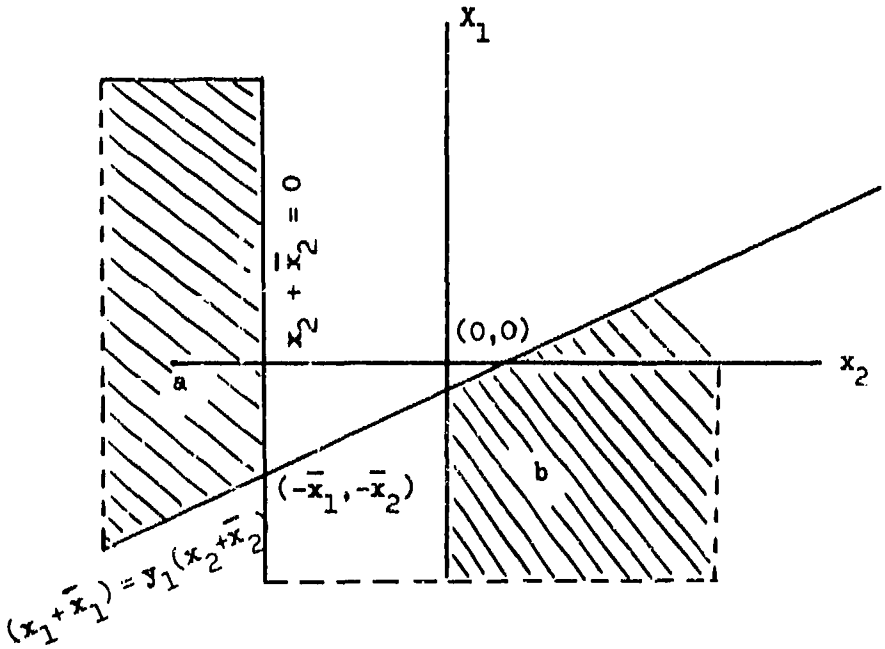


Figure 3. Geometrical Presentation of Bounded Areas (Non-Transformed Case).

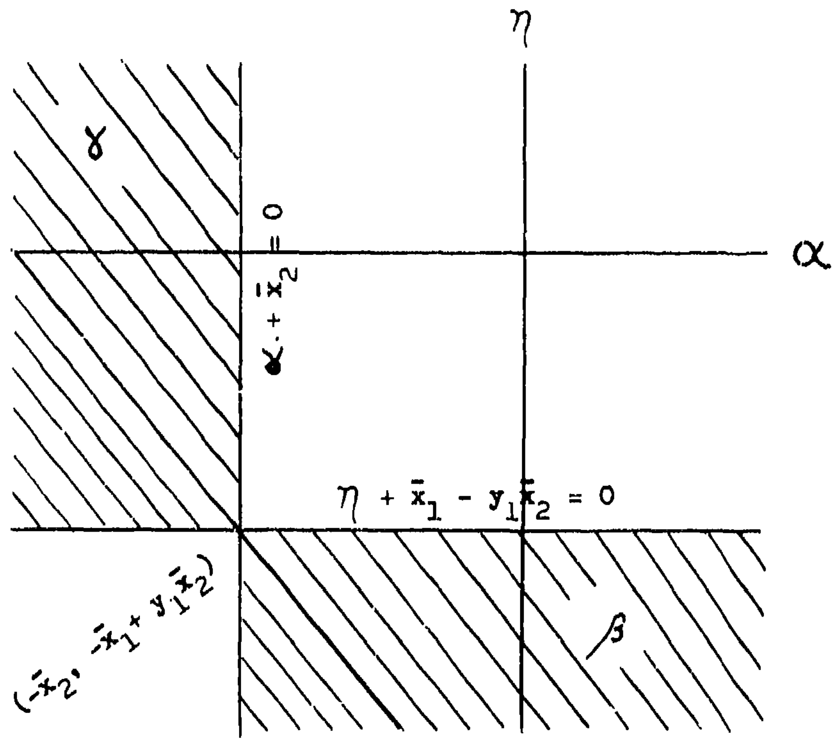


Figure 4. Geometrical Presentation of The Bounded Area in the Transformed Case.

which is identically $\frac{1}{1-\rho^2} \left\{ \frac{\alpha^2}{\sigma_\alpha^2} - 2\rho \frac{\alpha}{\sigma_\alpha} \frac{\eta}{\sigma_\eta} + \frac{\eta^2}{\sigma_\eta^2} \right\}$.

From this identity, these results are apparent:

$$\sigma_\alpha^2 (1-\rho^2) = \sigma_1^2 \sigma_2^2 (1-r^2) \Big/ (\sigma_1^2 - 2 r y_1 \sigma_1 \sigma_2 + y_1^2 \sigma_2^2)$$

$$\sigma_\eta^2 (1-\rho^2) = \sigma_1^2 (1-r^2) \qquad (4\text{-}24)$$

$$\frac{1}{\sigma_\alpha \sigma_\eta} \cdot \frac{\rho}{1-\rho^2} = (r \sigma_1 - y_1 \sigma_2) \Big/ \sigma_2 \sigma_1^2 (1-r^2) .$$

By squaring the last equation of(4-24) and multiplying the result by the first two equations, we obtain these results:

$$\rho^2 = (r \sigma_1 - y_1 \sigma_2)^2 \Big/ (\sigma_1^2 - 2 r y_1 \sigma_1 \sigma_2 + y_1^2 \sigma_2^2), \qquad (4\text{-}25)$$

so that

$$(1-\rho^2) = (1-r^2) \sigma_1^2 \Big/ (\sigma_1^2 - 2 r y_1 \sigma_1 \sigma_2 + y_1^2 \sigma_2^2), \qquad (4\text{-}26)$$

$$\sigma_\alpha = \sigma_2 , \qquad (4\text{-}27)$$

$$\sigma_\eta = (\sigma_1^2 - 2 r y_1 \sigma_1 \sigma_2 + y_1^2 \sigma_2^2)^{1/2}, \qquad (4\text{-}28)$$

and

$$\sigma_\alpha \sigma_\eta \sqrt{1-\rho^2} = \sigma_1 \sigma_2 \sqrt{1-r^2} . \qquad (4\text{-}29)$$

If we write $X = \frac{\alpha}{\sigma_\alpha} = \frac{x_2}{\sigma_2}$, and $Y = \frac{\eta}{\sigma_\eta} = \frac{x_1 - y_1 x_2}{(\sigma_1^2 - 2 r y_1 \sigma_1 \sigma_2 + y_1^2 \sigma_2^2)^{1/2}}$,

the quadrants A and B of the (X, Y) plane that correspond to the portions a and b of the (x_1, x_2) plane have as a common corner the point (-h, -k) where

$$h = \frac{\bar{x}_2}{\sigma_2},$$

$$k = \frac{\bar{x}_1 - y_1 \bar{x}_2}{(\sigma_1^2 - 2 r y_1 \sigma_1 \sigma_2 + y_1^2 \sigma_2^2)^{1/2}}. \tag{4-30}$$

From equation (4-23), V becomes

$$V = \iint_{A+B} \frac{\sigma_\alpha \sigma_\eta}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}} \exp\left\{-\frac{1}{2(1-\rho^2)} \quad (X^2 - 2\rho XY + Y^2) \quad dX\, dY\right\}. \tag{4-31}$$

Thus the probability of obtaining an index for which $y_1 \geq V$ is given by

$$C = 1 - V = \int_h^\infty \int_k^\infty + \int_{-h}^\infty \int_{-k}^\infty \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)} \left(X^2 - 2\rho XY + Y^2\right)\right\} dX\, dY. \tag{4-32}$$

Equation (4-25) provides two values of ρ. The appropriate value is chosen by noting that as $y_1 \longrightarrow \infty$, the point

$$(h\ ,k) \longrightarrow \left[\frac{x_2}{\sigma_2}\ , \frac{-\ x_2}{\sigma_2} \right]$$

so that the positive square root makes $(1-V) \longrightarrow 0$ as is desired. Extensive tables are found in [140] for which the value of the integral,

$$\int_h \int_k Z(\rho)\ dXdY = \int_h \int_k \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(- \frac{1}{2(1-\rho^2)} \left[X^2 - 2\rho XY + Y^2 \right] \right) dXdY \tag{4-33}$$

is calculated for small increments of ρ, $-1 < \rho < 1$. For non-positive h and k, the following relations hold:

$$(1) \quad \int_{-h} \int_k Z(\rho)\ dXdY = \int_k \frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2} Y^2}\ dY - \int_h \int_k Z(-\rho)\ dXdY$$

$$(2) \quad \int_h \int_{-k} Z(\rho)\ dXdY = \int_h \frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2} X^2}\ dX - \int_h \int_k Z(-\rho)\ dXdY$$

and

$$(3) \quad \int_{-h} \int_{-k} Z(\rho)\ dXdY = 1 - \int_h \frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2}X^2} dX - \int_k \frac{1}{\sqrt{2\pi}}\ e^{-\frac{1}{2}Y^2} dY + \int_h \int_k Z(\rho)\, dXdY$$

Probabilities of interest may be calculated with the aid of these relations and the appropriate values of ρ, h, k, X, and Y.

With the publication of Fieller's quadrature method, many agreed that the problem of the quotient of two normally distributed random variables was essentially solved. In later papers, emphasis was turned to the study of the general mathematical properties of the random variable quotient. Two such notable papers appearing in the late 1940's were published by Curtiss [24] and Gurland [48]. Curtiss approached the study of the properties of the quotient from an analysis of the application of the Lebesque-Stieltjes integral.

Gurland formulated the c.d.f. of the ratio of two linear combinations of correlated normal random variables. He presents two theorems for this distribution:

Theorem 1: (for positive denominator). Let $X_1, X_2, \ldots, X_n$ have a joint c.d.f., $F(x_1, x_2, \ldots, x_n)$ with the corresponding characteristic function $\phi(t_1, t_2, \ldots, t_n)$. Let G(x) be the c.d.f. of the linear combination, $(a_1 X_1 + a_2 X_2 + \ldots a_n X_n) / (b_1 X_1 + \ldots b_n X_n)$, where $a_1, a_2, \ldots, a_n, b_1, \ldots, b_n$ are real numbers.

If $P\left\{\sum_{j=1}^{n} b_j X_j \leq 0\right\} = 0$ then

$$G(x) + G(x-0) = 1 - \frac{1}{\pi i} \int \frac{\phi\{t(a_1-b_1 x_1),\ldots,t(a_n-b_n x_n)\}}{t} dt. \quad (4\text{-}34)$$

Theorem 2: (for positive or negative demoninator). Let G(x) be the c.d.f. of the ratio of linear combinations (of Theorem 1). If

$P\left\{\sum_{i=1}^{n} b_i X_i = 0\right\} = 0$ then

$$G(x) + G(x-0) = 1 - \frac{1}{\pi i} \oint \frac{\phi^{+}\left\{t(a_1 - b_1 x_1), \ldots, t(a_n - b_n t_n)\right\} + \ldots}{t} \qquad (4\text{-}35)$$

$$\ldots \frac{\phi\left\{t(a_1 - b_1 x_1), \ldots, t(a_n - b_n x_n)\right\}}{t} \, dt$$

where: $\phi^{+}(t_1, t_2, \ldots, t_n) = \underset{\sum b_k X_k > 0}{\int\int \cdots \int} e^{i(t_1 x_1, \ldots, t_n x_n)} \, d\,F(x_1, x_2, \ldots, x_n)$

and

$$\phi^{-}(t_1, t_2, \ldots, t_n) = \underset{\sum b_n X_k < 0}{\int\int \cdots \int} e^{i(t_1 x_1 \ldots t_n x_n)} \, d\,F(x_1, x_2, \ldots, x_n).$$

4.3 The Rectangular Density Function

A random variable, X, defined on a finite interval $\gamma = \{a \leq x \leq b\}$, (a, b, finite real numbers), is said to obey the uniform or rectangular p.d.f. if over the finite interval γ, the probability of A, a sub-interval of γ, is given by

$$P[A] = \begin{cases} \dfrac{\text{length of } A}{\text{length of } \gamma}, & A \text{ a subset of } \gamma, \\ 0, & \text{otherwise.} \end{cases} \qquad (4\text{-}36)$$

Thus the c.d.f. of the rectangular probability law, by this definition, is

$$F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & a \leq x \leq b, \\ 1, & b < x. \end{cases} \tag{4-37}$$

If we differentiate (4-37), we find the p.d.f. of the rectangular error function to be

$$\begin{aligned} f(x) &= \frac{1}{b-a}, & a \leq x \leq b, \\ &= 0, & \text{elsewhere.} \end{aligned} \tag{4-38}$$

4.3.1. The Product of Two Identically Distributed Rectangular Variables: The product of two identically distributed rectangular variables, say $y_1 = R_1R_2$, is easily found by quadrature utilizing the methods of sections 3.4.1. and 3.4.2. Four distinct cases arise in solving for the p.d.f. of $y_1 = R_1R_2$, where R is a rectangular variable defined on $\gamma = \{a \leq x \leq b\}$. The distinct cases may be enumerated as: Case (1) $y_1 = R_1R_2$, R defined on interval γ so that $0 < a < b$, Case (2) R defined on interval γ so that $a < 0 < b$, $|b| > |a|$, Case (3) R defined on interval γ so that $a < 0 < b$, $|a| > |b|$, and Case (4) R defined on interval γ so that $a < b < 0$. The p.d.f.'s of y_1 for each of these four cases are presented in Table II, Appendix A. The distribution of the product of any number of identically distributed rectangular variables can easily be obtained from Kendall's [146] derivation of the p.d.f. of the geometric mean of samples from a rectangular distribution.

4.3.2. The Quotient of Two Identically Distributed Rectangular Variables: The quotient of two identically distributed rectangular variables is also easily found using the methods of Sections 3.4.1 and 3.4.2. The same distinct cases depending upon the interval of definition as in 4.3.1 arise in the problem. The p.d.f.'s of these four cases are presented in Table III, Appendix A.

4.4 The Triangular Density Function.

The sum or the difference of two identically distributed rectangular r.v.'s defined on $\mathscr{X} = \{ a \leq x \leq b \}$ is triangularly distributed. The p.d.f. of T is

$$f(x) = \begin{cases} \dfrac{x-2a}{(b-a)^2}, & 2a < x \leq b+a , \\ \dfrac{2b-x}{(b-a)^2}, & b+a < x \leq 2b , \\ 0 , & \text{elsewhere.} \end{cases} \tag{4-39}$$

4.4.1 The Product of Two Identically Distributed Triangular Variables: Slightly more cumbersome expressions result from applying the methods of Chapter III to finding the p.d.f. of $Y_1 = T_1 T_2$ and $Y_1 = T_1/T_2$. This difficulty is due primarily to the fact that the p.d.f. of the triangular variable is defined by two branches as in (4-39). In deriving the p.d.f. of $T_1 T_2$, one must determine which combination of branches of T_1 and T_2 must be considered for each of the various mutually exclusive partitions of the set defined under the transformation. The condition is best illustrated by considering the problem of deriving this p.d.f. when T is defined on an arbitrary positive interval. Under the transformation $y_1 = T_1 T_2 = x_1 x_2$, $y_2 = x_2$, the set

$\gamma = \{2a - x_1 - 2b,\ \ 2a - x_2 - 2b\}$ is mapped into the set graphically illustrated in Figure 5. Here, the notation g, g' and h , h' denote the respective branches of $f(x_1)$ and $f(x_2)$ which apply to the partitions of β . The p.d.f. of Y_1 is derived by integrating over y_2 the product combinations as indicated in Figure 5 over the partitions A, B, C, D, and E of β . Two important special cases have been derived. (Table IV, Appendix A). These are the cases in which T is defined over a positive interval $[0, 2]$ and the interval $[-2, 2]$.

4.4.2 <u>The Quotient of Two Identically Distributed Triangular Variables</u>: The p.d.f.'s of two special cases for T defined on $[0, 2]$ and $[-2, 2]$ have been derived and are presented in Table V, Appendix A.

4.5 The Product and Quotient of "Mixed Components".

The product and quotient forms of the three measurement error variables under consideration which have not yet been treated are RN , TN , R/N and T/N. There is no requirement to investigate the reciprocal forms of these quotients since Cramér [22] has insured that the p.d.f. of the reciprocal of a quotient is immediately evident upon derivation of the p.d.f. of the quotient.

The four forms above lead to rather cumbersome expressions in respect to numerical calculations. A more convenient method of analysis in the cases of the quotient is demonstrated.

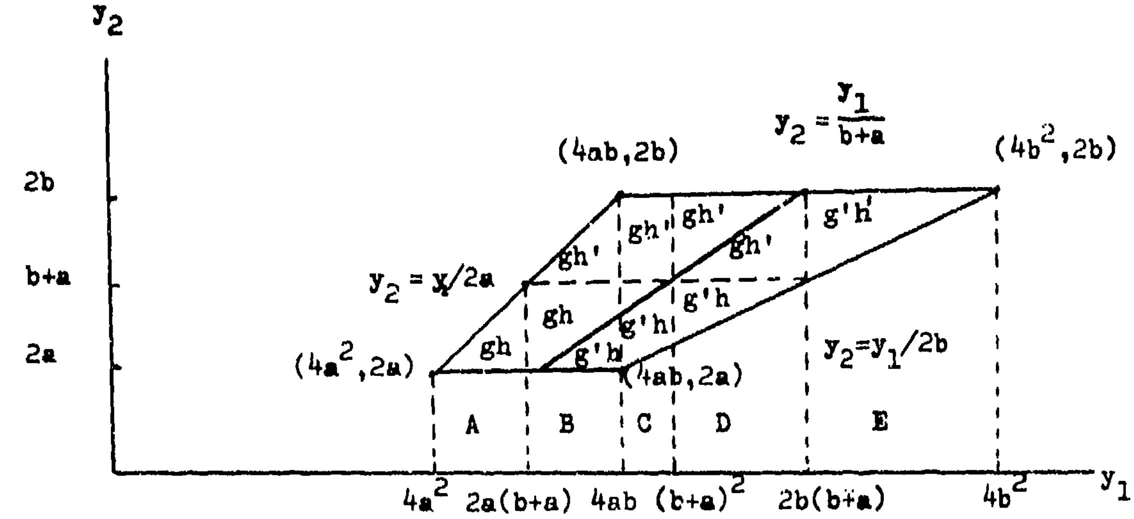


Figure 5. Geometric Representation of the Region β.

4.5.1.1 The Product of a Rectangular and a Normally Distributed Variable: Certain cases of this random variable product may be formulated in terms of the tabulated exponential integral. As an example, consider $Y_1 = RN = X_1 X_2$ where X_1 is rectangularly distributed on $\{0 \leq x_1 \leq 1\}$ and X_2 is normally distributed with parameters N(0, 1) and defined such that $Pr(X_2 = 0) = 0$. Defining the transformation, $y_1 = x_1 x_2$, $y_2 = x_2$, the p.d.f. of Y_1 is by Section 3.4.2, Chapter III,

$$\varphi(y_1) = \begin{cases} \displaystyle\int_{-\infty}^{y_1} -\frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{1}{2} y_2^2}}{y_2} \, dy_2 , & -\infty \leq y_1 \leq 0, \\[2ex] \displaystyle\int_{y_1}^{\infty} \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{1}{2} y_2^2}}{y_2} \, dy_2 , & 0 < y_1 < \infty. \end{cases}$$

Thus, with the substitution $\sigma = y_2^2$, $\varphi(y_1)$ satisfies

$$\varphi(y_1) = \frac{1}{2\sqrt{2\pi}} \int_{y_1^2}^{\infty} \frac{e^{-\frac{1}{2}\sigma}}{\sigma} \, d\sigma , \qquad y_1 \neq 0;$$

which, in turn, may be expressed as a function of the tabulated exponential integral after the substitution

$$v = \tfrac{1}{2}\sigma .$$

Thus,

$$\varphi(y_1) = \frac{1}{2\sqrt{2\pi}} \int_{\frac{1}{2} y_1^2}^{\infty} \frac{e^{-v}}{v}\, dv, \qquad y_1 \neq 0. \qquad (4\text{-}40)$$

All product forms of $Y_1 = RN$, where N is taken as N(0, 1), may be expressed as a function of the tabulated exponential integral. An extensive table of this integral is presented in reference [143].

Other product forms in which N is $N(\mu \neq 0, \sigma^2 \neq 0)$ are not expressible as simple functions of this integral.

4.5.1.2 <u>The Product of a Triangular and a Normally Distributed Variable</u>: As was the case in the preceding section, only the product cases of $Y_1 = TN$ in which N is taken N(0, 1) are easily attainable by the quadrature methods. One of the less complicated examples is the case of $Y_1 = TN = X_1 X_2$ where T is defined on $\{0 \le x_1 \le 2\}$ and N is N(0, 1). The p.d.f. of X_1 is,

$$f(x_1) = \begin{cases} g(x_1) = x_1, & 0 \le x_1 \le 1, \\ g'(x_1) = 2 - x_1, & 1 \le x_1 \le 2. \end{cases}$$

The p.d.f. of X_2 is defined such that

$$f(x_2) = \begin{cases} h(x_2) = \dfrac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2} x_2^2}, & -\infty < x_2 < 0,\ 0 < x_2 < \infty, \\ = 0, & x_2 = 0. \end{cases}$$

The regions of definition under the transformation $y_1 = x_1 x_2$, $y_2 = x_2$ are shown in Figure 6.

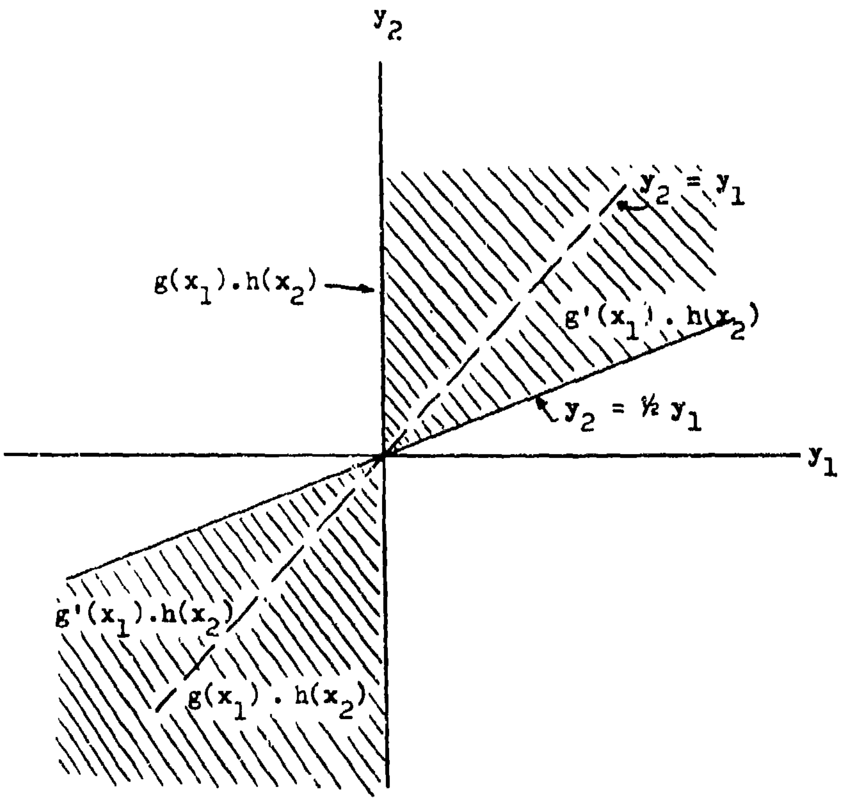


Figure 6. Geometric Interpretation of the Area of Definition under the Transformation $y_1 = x_1 x_2$, $y_2 = x_2$.

The p.d.f. of Y_1 is expressed as

$$\varphi(y_1) = \begin{cases} \dfrac{1}{\sqrt{2\pi}}\left[\displaystyle\int_{-\infty}^{y_1} -\frac{y_1}{y_2^{\,2}}\, e^{-\frac{1}{2}y_2^{\,2}}\, dy_2 + \int_{y_1}^{y_1/2} \left(2-\frac{y_1}{y_2}\right) e^{-\frac{1}{2}y_2^{\,2}}(-y_2)\,dy_2\right], & \begin{cases} -\infty < y_2 < 0, \\ -\infty < y_1 < 0, \end{cases} \\[4ex] \dfrac{1}{\sqrt{2\pi}}\left[\displaystyle\int_{y_1/2}^{y_1} \left(2-\frac{y_1}{y_2}\right) e^{-\frac{1}{2}y_2^{\,2}}\, y_2\, dy_2 + \int_{y_1}^{\infty} \frac{y_1}{y_2^{\,2}}\, e^{-\frac{1}{2}y_2^{\,2}}\, dy_2\right], & \begin{cases} 0 < y_2 < \infty, \\ 0 < y_1 < \infty, \end{cases} \end{cases}$$

In the above, $\varphi(y_1)$ reduces to an expression involving functions of the error functions and of the exponential integral so that

$$\varphi(y_1) = \frac{2}{\sqrt{2\pi}}\left[e^{-\frac{1}{2}y_1^{\,2}} - e^{-\frac{1}{8}y_1^{\,2}}\right] + \frac{y_1}{\sqrt{2\pi}}\left[\int_{y_1/2}^{y_1} e^{-\frac{1}{2}y_2^{\,2}}\, dy_2 - \int_{y_1}^{\infty} e^{-\frac{1}{2}y_2^{\,2}}\right] \tag{4-41}$$

$$+ \frac{1}{\sqrt{2\pi}} \int_{\frac{y_1^{\,2}}{8}}^{y_1^{\,2}} \frac{e^{-v}}{v}\, dv, \quad -\infty < y_1 < \infty \ .$$

All other cases of Y_1 in which N(0, 1) may be reduced to combinations of these special functions. Tables of the exponential integral and the error functions are presented in references [143], [144], and [145].

4.5.2.1 The Quotient of a Rectangular and a Normally Distributed Variable: A rectangular variable, say X_o, defined on $\{a \leq x_o \leq b\}$ may be redefined on the interval $[0, 1]$ through the use of the transformation,

$$X_o = X(b-a) + a$$

where X is rectangularly distributed on $[0, 1]$. The original rectangular variable may also be redefined on the interval $[-1, 1]$ by the transformation

$$X_o = X\beta + \eta,$$

where β = half-range of the interval $[a, b]$ and η = mean of the interval $[a, b]$.

The distribution of $Y_1 = R/N$ is greatly simplified by using one of the above transformations. In the special case of $Y_1 = R/N = X_1/X_2$ where R is defined on $\{0 \leq x_1 \leq 1\}$ or $\{-1 \leq x_1 \leq 1\}$ and N is taken to be N(0, 1) where $Pr(X_2 = 0) = 0$, the p.d.f. of y_1 is simply,

$$\phi(y_1) = \frac{1}{\sqrt{2\pi}} \left[1-e^{-\frac{1}{2y_1^2}} \right], \qquad -\infty < y_1 < \infty. \quad (4\text{-}40)$$

In the more general cases, R may be distributed as any of the four cases enumerated in Section 4.3, and N taken to be $N(\mu, \sigma^2)$. The respective p.d.f.'s of y_1 in these four cases are shown in Table VII, Appendix A. It is apparent that these expressions are rather awkward and difficult for numerical computations.

Broadbent [15] has contributed a general method with which to investigate a ratio in which the numerator is a rectangularly distributed random variable and the denominator is unspecified but possesses certain general properties. The c.d.f. of $Y_1 = R/N$ may be obtained in terms of tabulated integrals by treating this special case by Broadbent's method of analysis.

Any arbitrary random variable quotient of independent components whose numerator is rectangularly distributed may be written in the following "standardized" form:

$$y_1' = (\gamma + \beta\eta)/(\mu + \sigma\xi), \tag{4-43}$$

where:

η is rectangularly distributed on $[-1, 1]$,

γ is the expected value of the original R variable,

β is the half-range of the original R variable,

ξ is normally distributed with parameters (0, 1),

and μ and σ are the parameters of the original N variable.

The hypothesis is only slightly restricted by requiring that

$$\gamma > 0,\ \beta/\gamma = \alpha \text{ where } 0 < \alpha < 1,\ \mu < 0,\ \frac{\sigma}{\mu} = \psi \text{ where } 0 < \psi < \infty .$$

Since $y_1' = \left(\frac{\gamma}{\mu}\right)(1 + \alpha\eta)/(1 + \psi\xi),$ (4-44)

the standardized quotient y_1 may be considered, where

$$y_1 = \left(\frac{\mu}{\gamma}\right) y_1' = \frac{x_1}{x_2} . \tag{4-45}$$

In (4-45), X_1 is rectangularly distributed on $(1 - \alpha, 1 + \alpha)$ and X_2 is normally distributed with mean 1 and variance ψ^2. In this special case, X_2 is described by a probability density function $f(x_2) = \frac{d}{dx} F(x_2)$ existing almost everywhere and a c.d.f. $F(x_2) = z$ which has an inverse, say $G(z)$, which is defined, non-zero, and finite for almost all z. This condition allows the preclusion of probability measure at points $(x_2 = 0)$ and $(x_2 = \infty)$.

The function z is independent of X_1 and is rectangularly distributed on (0,1), a property of all c.d.f.'s. The joint p.d.f. of z and X_1 is by theorem 2, Chapter III,

$$\varphi(z, x_1) = \frac{1}{2\alpha} dz\, dx_1.$$

In order to obtain the joint p.d.f. of z and x_1, note that for any z, $dx_1 = |G(z)| dy_1$ and y_1 is bounded by

$$((1 - \alpha)/G(z),\ (1 + \alpha)/G(z)), \quad \text{if } z > \epsilon,$$

$$((1 + \alpha)/G(z),\ (1 - \alpha)/G(z)), \quad \text{if } z < \epsilon,$$

where $F(0) = \epsilon$, ϵ not necessarily small.

Therefore, z and y_1 have the joint p.d.f.,

$$\varphi(z, y_1) = \frac{1}{2\alpha} |G(z)|\ dz\ dy,$$

over the region shown in Figure 7.

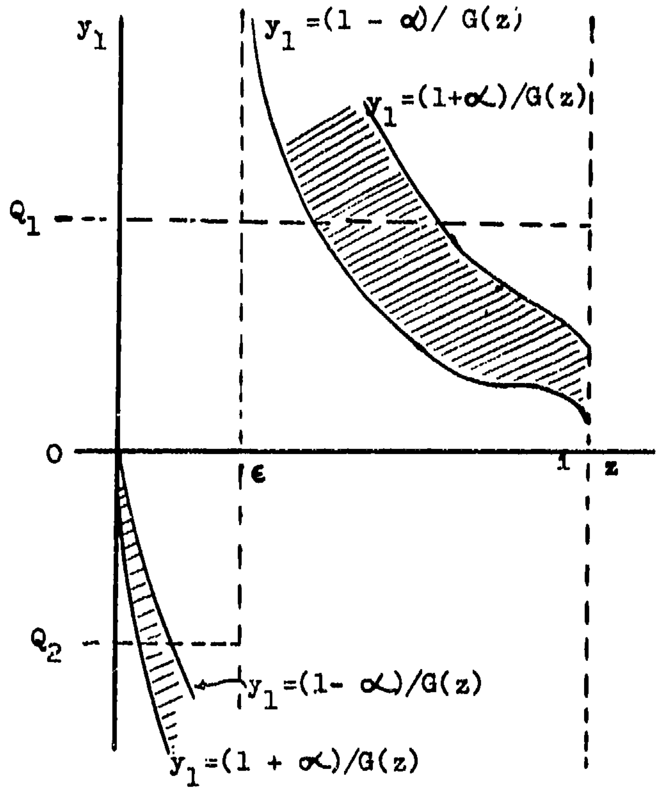


Figure 7. Geometric Interpretation of the Region of Definition of $\varphi(z, y_1)$.

The $\Pr(0 < y_1 \leq Q_1)$, for $(Q_1 > 0)$, is given by the integral of this function over the region A bounded by $y_1 = Q_1$, $z = 1$ and $y_1 = (1 - \alpha)/G(z)$ <u>minus</u> the integral over the region S by $y_1 = Q_1$, $z = 1$ and $y_1 = (1 + \alpha)/G(z)$, so that

$$\Pr(0 < y_1 < Q_1) = \int_A\int \frac{1}{2\alpha} \left|G(z)\right| \, dz \, dy_1 - \int_S\int \frac{1}{2\alpha} \left|G(z)\right| \, dx \, dy_1 . \quad (4\text{-}46)$$

After integration of this repeated integral in respect to y_1, the first integral may be expressed as

$$\frac{1}{2\alpha} \int_{z = F\left[(1-\alpha)/Q_1\right]}^{1} \left[Q_1 \, G(z) - (1 - \alpha)\right] dz. \quad (4\text{-}47)$$

If we change the variable to ξ and write

$$g(\xi) = \psi f(1 + \psi \xi),$$

the result is the p.d.f. of the standardized form of X_2. Now let

$$X = (1 - \alpha - Q_1)/(\psi Q_1),$$

and

$$Y = (1 + \alpha - Q_1)/(\psi Q_1).$$

The integral (4-47) becomes

$$\frac{\psi Q_1}{2\alpha} \, J_1 \, g(X) = \frac{\psi Q_1}{2\alpha} \int_{\xi = X}^{\infty} (\xi - X) \, g(\xi) \, d\xi .$$

In this special case where random variable X_2 is normally distributed,

$J_n\, g(X) = I_n(x)$ is the Hermitian probability integral of order n tabulated in the British Association tables, 2nd ed. (1946).[12]

The second integral of (4-46) is obtained in the same manner and is denoted by $J_1\, g(Y)$. Since $\Pr(y_1 \leq 0) = \Pr(x_2 < 0) + \Pr(x_2 = \infty) = \epsilon$ as previously noted, the c.d.f. of y_1 is, for $Q_1 > 0$,

$$F(y_1) = \Pr(y_1 \leq Q_1) = \epsilon + \frac{\psi Q_1}{2\alpha} \left\{ J_1\, g(X) - J_1\, g(Y) \right\} . \qquad (4\text{-}48)$$

It follows that from (4-48), the

$$\Pr(y_1 > Q_1) = \frac{\psi Q_1}{2\alpha} \left\{ \left[Y + J_1\, g(Y) \right] - \left[X + J_1\, g(X) \right] \right\} - \epsilon . \qquad (4\text{-}49)$$

The forms (4-48) and (4-49) are useful when ϵ is small for the calculation of percentage points at the tails of the distribution of y_1 since as $Q_1 \longrightarrow 0$ and $X, Y \longrightarrow \infty$, both $J_1\, g(X)$ and $J_1\, g(Y) \longrightarrow 0$ — the former more slowly than the latter. For small Q_1, $J_1\, g(Y)$ is small in comparison with $J_1\, g(X)$ in the case of normally distributed x_2. In this case the lower percentage points of y_1' are $(\frac{\alpha}{\mu})$ times those of y_1.

When ϵ is not small, the $\Pr(y_1 \leq Q_2)$, $(Q_2 < 0)$, must be considered. The integrals over the region in Figure 7 bounded by $y_1 = Q_2$, $z = \epsilon$, $y_1 = (1 + \alpha)/G(z)$ and $y_1 = Q_2$, $z = \epsilon$ and $y_1 = (1 - \alpha)/G(z)$ may be

12) A detailed account of these integrals is also given by Fisher (B. A. Tables, 1931).

expressed as

$$\Pr(y_1 \leq Q_2) = \frac{\psi Q_2}{2\alpha}\left\{\int_{\xi=Y}^{-1/\psi} (\xi - Y)\, g(\xi)\, d\xi - \int_{\xi=X}^{-1/\psi} (\xi - X) g(\xi)\, d\xi\right\}$$

where:

$$Y = (1 + \alpha - Q_2)/(\psi Q_2), \quad X = (1 - \alpha - Q_2)/(\psi Q_2).$$

4.5.2.2. The Quotient of a Triangular and a Normal Variable: As shown in Section 4.2, the sum (or difference) of two independent rectangular variables distributed on the same interval is triangularly distributed. Considerable difficulty is encountered in deriving the respective p.d.f.'s of $Y_1 = T/N$. This difficulty is attributable to the fact that the p.d.f. of T is defined in two branches as Equation (4-39). The resulting p.d.f.'s are functions of the error function. In the simplest case, i.e., T is triangularly distributed on [0, 2] and N is taken to be N(0, 1), the p.d.f. of Y_1 is expressible as the difference of two tabulated error functions. So that

$$\varphi(y_1) = \frac{y_1}{2}\left[\int_{2/y_1}^{1/y_1} e^{-\frac{1}{2} y_2^2}\, dy_2 - \int_{1/y_1}^{0} e^{-\frac{1}{2} y_2^2}\, dy_2\right], \quad y_1 \neq 0. \qquad (4\text{-}50)$$

Using Broadbent's method of analysis, an expression convenient for numerical computation may be derived.

Consider the sum (or difference) of two independently distributed rectangular random variables defined on $[a, b]$ with the same half-range β. Allow the sum (or difference) of their means to be denoted as γ. The ratio T/N may be written in a similar form to that of (4-43) or

$$y_1' = (\gamma + \beta\eta)/(\mu + \sigma\xi) .$$

(Conditions being such that $\gamma > 0$, $\frac{\beta}{\gamma} = \alpha$ where $0 < 2\alpha < 1$, $\mu > 0$, $\frac{\sigma}{\mu} = \psi$ where $0 < \psi < \infty$). The variable η is distributed on $[-2, 2]$ and ξ is independent of η and satisfies those conditions outlined in the preceding section.

The standardized quotient is

$$y_1 = \left(\frac{\mu}{\gamma}\right) y_1' = \frac{x_1}{x_2} .$$

In this case, χ_1 has the p.d.f.

$$\varphi(x_1) = \begin{cases} \dfrac{x_1 - 1 + 2\alpha}{4\alpha^2} , & \text{for } (1 - 2\alpha \leq x_1 \leq 1) , \\ \dfrac{1 + 2\alpha - x_1}{4\alpha^2} , & \text{for } (1 < x_1 \leq 1 + 2\alpha) . \end{cases}$$

The joint p.d.f. of $z = F(x_2)$ and y_1 is

$$\begin{aligned} &G(Z) \left\{ y_1 G(Z) - 1 + 2\alpha \right\} dz\, dy_1/4\alpha^2, \ ((1 - 2\alpha)/G(z)) \leq y_1 \leq 1/G(z)) \\ &G(Z) \left\{ 1 + 2\alpha - y_1 G(Z) \right\} dz\, dy_1/4\alpha^2, \ ((1/G(z) \leq y_1 \leq (1 + 2\alpha)/G(z)), \end{aligned} \qquad (4\text{-}51)$$

when $z > \epsilon$ and corresponding expressions when $z < \epsilon$ over the region shown in Figure 8.

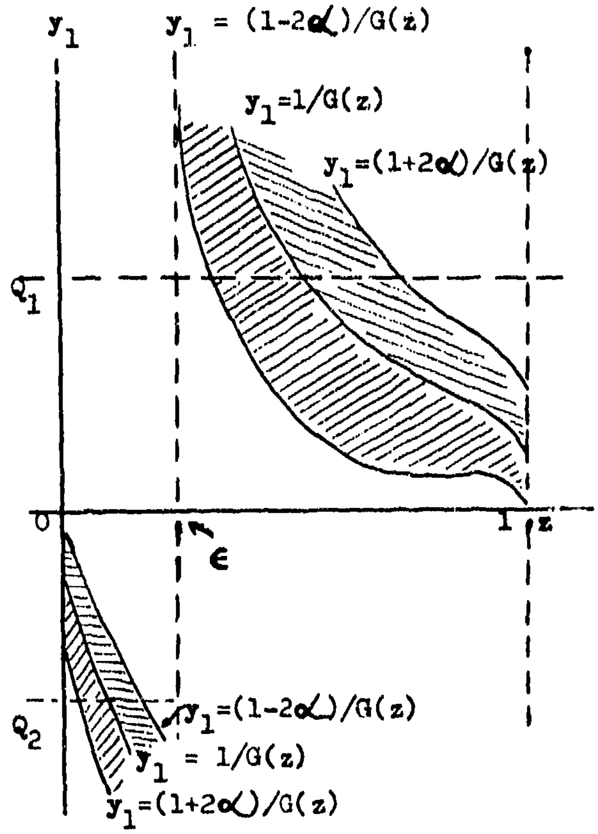


Figure 8. Geometrical Interpretation of $f(z, y_1)$.

The $\Pr(0 < y_1 \leq Q_1)$, $(Q_1 > 0)$, is the integral of the first element over the region A bounded by $y_1 = Q_1$, $z = 1$ and $y_1 = (1-2\alpha)/G(z)$ minus the integral of the difference of the two elements bounded by $y_1 = Q_1$, $z = 1$ and $y_1 = 1/G(z)$, minus the integral of the second element over the region bounded by $y_1 = Q_1$, $z = 1$ and $y_1 = (1+2\alpha)/G(z)$.

The first of these integrals is

$$\frac{1}{4\alpha^2} \int_{z = F\left[(1-2\alpha)/Q_1\right]}^{1} ½ \left\{ Q_1^2 \left[G(z)\right]^2 - 2Q_1\, G(z)\, (1-2\alpha) + (1-2\alpha)^2 \right\} dz . \quad (4\text{-}52)$$

The integral is simplified by changing the variable of integration to ξ and writing

$$U = (1 - 2\alpha - Q_1)/(\psi Q_1).$$

The integral (4-52) becomes

$$\frac{\psi^2\, Q_1^2}{4\alpha^2}\ ½ \int_{\xi = U}^{\infty} (\xi - U)^2\, g(\xi)\, d(\xi).$$

Thus, if

$$J_2\, g(U) = \frac{1}{2!} \int_{\xi = U}^{\infty} (\xi - U)^2\, g(\xi)\, d(\xi),$$

$V = (1-Q_1)/(\psi Q_1)$, and $W = (1+2\alpha - Q_1)/(\psi Q_1)$, the full expression for $\Pr(y_1 < Q)$ is

$$\Pr(y_1 \leq Q) = \epsilon + \frac{\varphi^2 Q^2}{(2\alpha)^2} \left\{ J_2\, g(U) - 2\, J_2\, g(V) + J_2 g(W) \right\} . \quad (4\text{-}53)$$

It may be concluded also that

$$\Pr(y_1 > Q) = \frac{\varphi^2 Q^2}{(2\alpha)^2} \left\{ \left[\frac{U^2 + 1}{2} - J_2\, g(U) \right] - 2 \left[\frac{V^2 + 1}{2} - J_2\, g(V) \right] + \right.$$

$$\left. + \left[\frac{W^2 + 1}{2} - J_2\, g(W) \right] \right\} - \epsilon . \quad (4\text{-}54)$$

Equations (4-53) and (4-54) have similar computational advantages as noted for (4-48) and (4-49).

V. LIMITING DISTRIBUTIONS, APPROXIMATIONS AND ASYMPTOTIC RESULTS

5.1 Introduction.

A frequently encountered problem in the missile and space industry may be described as follows:

> A complex mechanism (say, a missile) contains many individual and, to some extent, independently operating subsystems. Such structures include airborne subsystems such as payload, guidance, electrical, hydraulic, flight control, pressurization and airframe as well as ground support subsystems, i.e., a ground power, facilities and propellant subsystem.

Frequently, component data are available pertaining to failure rates of each of these subsystems. These data are used to estimate the overall system failure rate. Of interest then is the statistical statement which may be made concerning overall system failure probability.

This classical problem in reliability analysis has received considerable attention; however, most attempts, in practice, end in one of two methods of analysis. There may be a decision to use the exponential failure probability density function to estimate the failure probability of each subsystem. As an alternative, the use of a digital computer is made to simulate operation of the overall system and thus gain, through repetition, a confidence statement of the overall system failure probability. The use of the exponential density function features the highly important property of forming a complex product simply by summing the exponents of the variables representing component failure rates.

In general, problems concerning the establishment of a reliability estimate for various equipment or events may be regarded as an investigation of the properties of a random variable formed from the product of several components. Problems of this type may be investigated by studying the distribution of Y_1 which may be expressed in the general form

$$Y_1 = \prod_{i=1}^{n} X_i \Bigg/ \prod_{j=n+1}^{s} X_j \, . \tag{5-1}$$

As seen in Chapter IV, Geary [38] produced an approximation for $Y_1 = N_1/N_2$ where N represents the normal density function. Fieller [32] shows exact percentages of N_1/N_2. Craig [20] and Aroian [2] have studied and tabulated results of $N_1 N_2$. In addition, Broadbent [14], [15] studied measures of efficiency which are of the form R/N and T/N where R and T represent the Rectangular and Triangular distributions. Certain linear combinations of random variables in quotient form have been examined such as Gurland's [48] investigation of

$$K = (a_1 X_1 + a_2 X_2 + \ldots + a_n X_n)/(b_1 X_1 + b_2 X_2 + \ldots + b_n X_n)$$

where the joint density function $f(x_1, x_2, \ldots, x_n)$ is known. All of these results are of a specific nature and are not generally applicable to investigating distributions of the type (5-1).

Several approximations to distributions of random variables of this type (5-1) have been investigated. One of these will now be described.

5.2 Lognormal Approximations.

Consider a measure Y_1 of the form (5-1) in which each component is subject to rectangular or normally distributed measurement error with known half-range or coefficient of variation. In many applications, the distributions of the sampling or measurement errors are known to be one of these two forms, and their standard deviations or half-ranges can usually be found by simple investigation. In other applications, the distribution of these errors can only be estimated and then perhaps the roughest approximations should be used only as bounds for experimental error.

It is impossible to list exhaustively all possible combinations of errors to consider in (5-1). In the event of a large number of errors being combined, the asymptotic distribution may be used with confidence [Cramér [22]]. It is in the intermediate cases that approximations must be critically considered. The choice of a suitable family of approximating distributions, however, will always be rather subjective. In this respect, the consideration of an approximation is directed towards finding a "working, first-order approximation". Several authors, i.e., Shellard [115] , Finney [35], Gaddum, and Johnson, have investigated the use of lognormal approximations to (5-1) as an alternative to attempting to find its exact distribution using the methods of Section 3.4. They have shown that the p.d.f. of

Y_1 (5-1) tends to the lognormal as $s \longrightarrow \infty$ under very general conditions.

The most widely used lognormal approximation to (5-1) is a variate z such that $\log(z-\xi)$ is normally distributed with mean μ and variance σ^2. Here, for simplicity, only the methods of selecting μ and σ^2 are considered, ξ being considered as equal to zero. The notation $\alpha_1, \alpha_2, \ldots, \alpha_n$ denotes the coefficient of variation or the quotient of the half-range and the mean of a rectangular variable for each of the variates $x_1, x_2, \ldots, x_n$. Thus $Y_1 = R/N$, $(\alpha_1 = 0.33, \alpha_2 = 0.10)$, denotes the quotient of a rectangular variate whose mean is three times its half-range and an independent normal variate whose coefficient of variation is 10%. The variate z is necessarily positive although Y_1 may have negative values. When $\alpha_1, \alpha_2, \ldots$, are small positive values, the approximation by this lognormal distribution may be satisfactory, since the probability attached to the possibility of negative values is exceedingly small.

There are two methods of fitting the lognormal approximation. The first method of choosing μ and σ^2 is to calculate the moments of $\log y_1$ and to set μ and σ^2 equal to the first and second corrected moments. This method, termed the <u>method of fitting by moments to log</u> y_1, was first detailed by Finney [34] (1941). An alternate method is to choose the lognormal approximation whose mean and variance are equal to the mean and variance of y_1. This method generally termed, <u>the method of fitting by moments to</u> y_1, differs from that of the first method but the difference is seen to disappear as n increases. The possibility of two methods raises the question of which method should be used.

Let a lognormal p.d.f. have mean m and variance v^2 and let the p.d.f. of the logarithm of the variate from the lognormal have mean μ and variance σ^2. It is easily shown that

$$
\begin{aligned}
m &= \exp\,(\mu + \sigma^2/2), \\
v^2 &= \exp\,\{(2\mu + \sigma^2)\}\;\{\exp(\sigma^2) - 1\}, \qquad (5\text{-}2)\\
\mu &= \log m - \tfrac{1}{2}\log\,(1 + v^2/m^2), \\
\sigma^2 &= \log\,(1 + v^2/m^2).
\end{aligned}
$$

The fitting to $\log y_1$ requires that we select the mean and variance of $\log y_1$ and set μ and σ^2 equal to them. The k% point of this fit is $\exp(\mu + p_t\,\sigma)$, where the probability that a standardized normal variable is less than p_t is t/100.

The fit to y_1 requires that the mean and variance of y_1 be found and set equal to m and v^2. The k% point of this fit is

$$\left(m \exp\left[p_t\,\{\log\,(1 + t^2)\}^{1/2}\right]\right)/\,(1 + t^2)^{1/2}, \qquad (5\text{-}3)$$

where: t = s/m, the coefficient of variation of y_1.

The k% points of (5-3) are tabulated in Table VIII, Appendix A. The appropriate points are found by: (1) determining the expected value and coefficient of variation of y_1, (2) entering the table with the appropriate t, and (3) multiplying the value in the table by m.

In order to find the moments of y_1 or of $\log y_1$, expressions are derived for the moments of various positive and negative powers of the normally and rectangularly distributed variables. These variables may be expressed as $x = \mu(1 + \alpha w)$ where α and w are as previously defined.

In the case of normally distributed w, the moments of $\log(1 + \alpha w)$ are more easily derived from the truncated distribution,

$$\exp\ (-\tfrac{1}{2} w^2)\ dw / \left\{(1-\epsilon)/\sqrt{2\pi}\right\}, \qquad y > -\frac{1}{\alpha}, \quad (5\text{-}4)$$

which has the property,

$$\Pr(1 + \alpha w < 0) = \epsilon\ , \ (\epsilon < 10^{-10} \text{ when } 0 < \alpha \leq 0.15).$$

For all practical purposes, (5-4) is indistinguishable from the normal p.d.f. with the above restrictions.

The moment-generating function of $\log(1 + \alpha w)$ is $E\left[(1 + \alpha w)^{it}\right]$ and may be written as

$$\frac{1}{\sqrt{2\pi}(1-\epsilon)} \int_{-1/\alpha}^{1/\alpha} (1 + it\ \alpha w + \tfrac{1}{2}\, it(it-1)\alpha^2\ w^2 + \ldots)\exp(-\tfrac{1}{2}\ w^2)dw + K \quad (5\text{-}5)$$

where $|K| \leq \epsilon\ /(1-\epsilon)$.

Broadbent [14] has discussed this integral and its convergence and has shown that it leads to the cumulants of $\log\ (1 + \alpha w)$. He has also shown that the cumulants of $\log\ (1 + \alpha w)$ when w is rectangularly distributed may be obtained and that a similar analysis leads to the expected values of various powers of $(1 + \alpha w)$ for both cases of w.[13] These expressions for $E\left[(1 + \alpha w)^r\right]$ and $E\left[\log(1 + \alpha w)^r\right]$ for $r = \pm 1/2, \pm 1, \pm 2$, and ± 4 are shown in Table IX, Appendix A.

Before considering a complicated case for the lognormal approximation, it is worthwhile to consider the very simple independent case of $Y_1 = \frac{N_1}{N_2} = \frac{(1 + \alpha_1 w)}{(1 + \alpha_2 w)}$ where N_i represent the normal p.d.f. The coefficients of variation, α_1 and α_2 of the numerator and denominator are taken as $100\,\alpha_1$, and $100\,\alpha_2$ respectively.

Using Table IX, these results are obtained:

$$E\left[y_1\right] = E\left[N_1\right] E\left[\frac{1}{N_2}\right] = m = (1)\,(1 + \alpha_2^{\,2} + 3\,\alpha_2^{\,4} + \ldots),$$

$$E\left[Y_1^{\,2}\right] = E\left[N_1^{\,2}\right] E\left[\frac{1}{N_2^{\,2}}\right] = (1+ \alpha_1^{\,2})(1+3\,\alpha_2^{\,2} + 15\,\alpha_2^{\,4} + \ldots),$$

$$V\left[y_1\right] = E\left[y_1^{\,2}\right] - E\left[y_1\right]^2 = v^2 = \alpha_1^{\,2} + \alpha_2^{\,2} + 3\,\alpha_1^{\,2}\,\alpha_2^{\,2} + 8\,\alpha_2^{\,4} + \ldots).$$

Similarly for the method of fitting to $\log y_1$, we obtain these results:

$$E\left[\log y_1\right] = \mu = -(\tfrac{1}{2}\,\alpha_1^{\,2} + \tfrac{3}{4}\,\alpha_1^{\,4} + \ldots) + (\tfrac{1}{2}\,\alpha_2^{\,2} + \tfrac{3}{4}\,\alpha_2^{\,4} + \ldots),$$

$$V\left[\log y_1\right] = v^2 = (\alpha_1^{\,2} + \tfrac{5}{2}\,\alpha_1^{\,4} + \ldots) + (\alpha_2^{\,2} + \tfrac{5}{2}\,\alpha_2^{\,4} + \ldots).$$

---

13) J. Haldane [51] has also derived these moments.

The first two moments of the lognormal fitted to log y_1, using these values of μ and σ^2, are:

$$\text{Mean} = 1 + \alpha_2^2 + \tfrac{1}{2}\alpha_1^4 + \tfrac{5}{2}\alpha_2^4 + \dots ,$$

$$\text{Variance} = \alpha_1^2 + \alpha_2^2 + 3\alpha_1^4 + 3\alpha_1^2\alpha_2^2 + 5\alpha_2^4 + \dots$$

These estimates agree with m and v^2 to $O(\alpha_1^2, \alpha_2^2)$, and to higher orders when $\alpha_1 = \alpha_2$, thus the difference in the two fits is small when α_1 and α_2 are small. In fact, for this case, there is complete agreement in the first three significant digits at the 1%, 5% and the 95% points between the two methods of fitting and the exact points computed by Craig's methods (4-8). The differance between exact points and the two methods of lognormal approximation may be shown to be $\leq \pm .002$ at the 1%, 5%, 95% and 99% points in the cases of R_1/R_2, RN, and N_1/N_2.

We now return to the problem of setting limits to

$$y_1 = (x_1 x_2 \dots x_j)/(x_{j+1} \dots x_n).$$

If the number of component variates is larger than two, the lognormal approximation will give satisfactory results when the coefficients of variation of the components are small and do not vary greatly. It is, therefore, a suitable approximation for general use.

To calculate the first two moments of y_1, it is necessary to know the coefficient of variation or quotient of half-range and mean, of each component, and then to combine the values given in Table IX. The percentage points are then given in Table VIII. Finally, these points are to be multiplied by

$$(\mu_1 \mu_2 \cdot \cdot \cdot \mu_j)/(\mu_{j+1} \cdot \cdot \cdot \mu_n),$$

where $\mu_i = \bar{x}_i$.

As an example, consider

$$y_1 = RN_1N_2/N_3N_4$$

and $\alpha_1 = .01$, $\alpha_2 = .02$, $\alpha_3 = .005$, $\alpha_4 = .015$, $\alpha_5 = 0.005$ and $\alpha_6 = 0.005$.

Using Tables VIII and IX, we find the mean and coefficient of variation of Y_1 to be 1.00026 and 2.57%. The 1% and 99% points of Y_1 are 0.937 and 1.063.

VI. CHARACTERIZING PROPERTIES OF STATISTICAL DISTRIBUTIONS

6.1 Introduction.

If X_1 and X_2 are independent and $Y_1 = X_1 + X_2$ is known to obey a normal p.d.f., it is well known (Cramér [22]) that X_1 and X_2 must each be normally distributed. This suggests the general problem of determining when the p.d.f. of an algebraic combination of random variables uniquely determines the respective p.d.f.'s of the component random variables. Investigations of this type have led to a class of statistical problems formally termed "characterizing properties of statistical distributions."

An important phase of this problem is that of determining the properties of observations from their estimated p.d.f. when it is known that the samples are in reality either a product or a quotient of random variables.

6.2 Three Important Problems.

By Theorem 3, ChapterIII, the p.d.f. of a random variable quotient $Y_1 = X_1/X_2$, where X_1 and X_2 are arbitrary independent random variables, is given in its most general form as

$$\varphi(y_1) = \int_{-\infty}^{\infty} f_{X_1}(y_1 y_2)\, g_{X_2}(y_2) \quad |y_2| \; dy_2 , \quad (6\text{-}1)$$

where the transformation $y_1 = x_1/x_2$, $y_2 = x_2$ is implied and $f_{X_1}(\cdot)$ and $g_{X_2}(\cdot)$ are the respective independent p.d.f's of X_1 and X_2.

Several important problems may be formulated regarding (6-1); three of which will now be briefly considered.

The first and most straight-forward is that of determining $\varphi(y_1)$ when the functional forms of $f_{X_1}(\cdot)$ and $g_{X_2}(\cdot)$ are known. This problem simply reduces to the integration of (6-1).

The second problem which many consider as being the "classical" characterization problem may be formulated as follows:

Let X_1 and X_2 be two independent and identically distributed random variables having a common p.d.f., f(x). Let the quotient $Y_1 = X_1/X_2$ be distributed insome known p.d.f., $\varphi(y_1)$. Is it possible to obtain a characterization of f(x) by the properties of the quotient p.d.f?

In this case, Equation (6-1) may be written as

$$\varphi(y_1) = \int_{-\infty}^{\infty} f(y_1\, y_2)\; f(y_2)\; |y_2|\; dy_2 \tag{6-2}$$

where only the form of $\varphi(y_1)$ is known. Among the questions which naturally arise are: (1) how many independent solutions of f($\cdot$) satisfy the linear integral equation (6-2) for a given function of $\varphi(y_1)$?, and (2) if more than one, is it possible to deduce the general properties of the class of all f(x) satisfying (6-2)?

Several investigators have dealt with a Cauchy form for $\varphi(y_1)$. When X_1 and X_2 are independently and normally distributed random variables with parameters (0, 1) or (0, σ^2), it is well known that

$Y_1 = X_1/X_2$ is Cauchy distributed symmetrically about $y_1 = 0$. It has been conjectured that this is a unique property of the normal p.d.f.; a hypothesis which implies the existence of only one solution of $f(x)$, namely, $f(x) = N(0, \sigma^2)$, which satisfies the linear integral equation,

$$\frac{C}{1+y_1^2} = \int_{-\infty}^{\infty} f(y_2) f(y_1 y_2) \, |y_2| \, dy_2, \; (C = \text{a constant}). \quad (6\text{-}3)$$

Laha [75], [76], Steck [117] and Mauldon [85] all have subsequently proved this conjecture to be false by exhibiting several non-normal solutions. Goodspeed [40] has shown in fact that an infinity of solutions to (6-3) exists.

Laha [76] has shown that if X_1 and X_2 are independently and identically distributed random variables having a common c.d.f. $F(x)$ such that the quotient $Y_1=X_1/X_2$ follows the Cauchy Law distributed symmetrically about zero, then $F(x)$ has these general properties:

i) $F(x)$ is symmetric about the origin $x = 0$;

ii) $F(x)$ is absolutely continuous and has a continuous p.d.f. $f(x) = F'(x) > 0$;

iii) the random variable X has an unbounded range;

iv) the p.d.f. $f(x)$ satisfies the linear integral equation

$$\frac{C'}{1+ y_1^2} = \int_{0}^{\infty} f(y_1 \, y_2) \, f(y_2) \, y_2 \, dy_2 , \quad (6\text{-}4)$$

which holds for all y_1, where C' is a constant. (Here again the transformation, $y_1 = x_1/x_2$, $y_2 = x_2$ is implied).

It is now apparent that an answer to the question of the number of f(x)'s satisfying the equation (6-4) is the equivalent of completely enumerating all solutions to this integral equation. This problem is very difficult and still remains to be solved.

Mauldon [85] has studied similar problems showing in particular that: (1) the ratio of two independently and identically distributed chi-square variables is not the only quotient distributed according to the general F-distribution,

$$\varphi(y_1) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{y_1^{b-1}}{(1+y_1)^{b+a}}, \quad (0 < y_1 < \infty, \ a, \ b > 0) \qquad (6\text{-}5)$$

and (2) that we cannot assert that f(x) is normal even when a sample of x_i, i = 1, . . . , n observations is taken from the same p.d.f. distributed symmetrically about zero such that the sample statistic $t = \bar{x} \sqrt{n/s}$ is distributed as Student's t-distribution ($\bar{x}$, s denoting the usual sampling statistics.)

The third problem is a general case of the preceding problem and involves solving the integral equation

$$\varphi(y_1) = \int_{-\infty}^{\infty} f_{X_1}(y_1 y_2)\, g_{X_2}(y_2)\, |y_2| \, dy_2 \qquad (6\text{-}6)$$

when one of the two combinations of functions, ($\varphi(y_1)$, $f_{X_1}(\cdot)$) or ($\varphi(y_1)$, $g_{X_2}(\cdot)$), is known.

In certain sampling studies, it is conceivable that collected data y_i, $(i = 1, \ldots, n)$, will be of a ratio form, say $y_i = x_{1_i}/x_{2_i}$, and at the same time estimates of the respective p.d.f.'s of the ratio y_i and of one of the components, say X_1, can be made with a high degree of certainty. In these cases, it is reasonable to question whether or not the general properties of X_2 may be ascertained.

A simple method of obtaining an estimate of the p.d.f. of X_2 is to consider the ratio $X_2 = X_1/Y_1$ using the estimated p.d.f. of the respective components. However, the p.d.f.($\varphi(x_2)$) may represent only one of several possible solutions. In this case only from the complete enumeration of all solutions satisfying equation (6-6) will it be possible to deduce the general properties of $g_{X_2}(\cdot)$.

APPENDIX A

STATISTICAL TABLES

TABLE I

The Probability Density Function of $Y_1 = N_1 N_2$, the Product of Two Normally Distributed Random Variables, for Certain Dependent and Independent Cases.*

Case I $N_1 = N(0,1)$, $N_2 = N(0,1)$, correlation coefficient $\rho = 0$.	Case II $N_1 = N(0,1)$, $N_2 = N(0,1)$, correlation coefficient $\rho = k$.
$\wp(y_1) = \frac{1}{\pi} K_o(\lvert y_1 \rvert)$, $y_1 \neq 0$, $K_o(\lvert y_1 \rvert)$ = modified Bessel function of zero order of the second kind.	$\wp(y_1) = \frac{1}{\pi\sqrt{1-k^2}}\, e^{\frac{\boxed{k y_1}}{(1-k^2)}} K_o\left(\frac{\lvert y_1 \rvert}{(1-k^2)}\right)$, $y_1 \neq 0$.
Case III $N_1 = N(0, \sigma_1^2)$, $N_2 = (0, \sigma_2^2)$, correlation coefficient $\rho = 0$	Case IV $N_1 = N(0, \sigma_1^2)$, $N_2 = N(0, \sigma_2^2)$, correlation coefficient $\rho = k \neq 0$.
$\wp(y_1) = \frac{1}{\pi \sigma_1 \sigma_2} K_o\left(\frac{\lvert y_1 \rvert}{\sigma_1 \sigma_2}\right)$, $y_1 \neq 0$.	$\wp(y_1) = \frac{1}{\pi \sigma_1 \sigma_2 \sqrt{1-k^2}}\, e^{\frac{k y_1}{(1-k^2)\sigma_1\sigma_2}} K_o\left(\frac{\lvert y_1 \rvert}{(1-k^2)\sigma_1\sigma_2}\right)$, $y_1 \neq 0$.

* see section 4.2.1, chapter IV

TABLE II

The Probability Density Functions of $Y_1 = R_1R_2$ for Identically Distributed Rectangular Variables Defined on the Interval $\gamma = \{a \le x \le b\}$. (see section 4.3.1, chapter IV)

Case I	Case II
R defined on $\gamma = \{a \le x \le b\}$, $0 < a < b$.	R defined on $\gamma = \{a \le x \le b\}$, $a < 0 < b$, $\lvert b\rvert > \lvert a\rvert$.
$\varphi(y_1) = \begin{cases} \dfrac{\log y_1 - \log a_2^2}{(b-a)^2}, & a^2 \le y_1 \le ab \\ \dfrac{\log (b_1^2) - \log (y_1)}{(b-a)^2}, & ab < y_1 < b^2 \end{cases}$	$\varphi(y_1) = \begin{cases} \dfrac{\log (a^2b^2) - 2\log \lvert y_1\rvert}{(b-a)^2}, & ab \le y_1 < 0, \\ 0, & y_1 = 0, \\ \dfrac{\log (a^2b^2) - 2\log \lvert y_1\rvert}{(b-a)^2}, & 0 < y_1 \le a^2, \\ \dfrac{\log (b^2) - \log \lvert y_1\rvert}{(b-a)^2}, & a^2 < y_1 \le b^2. \end{cases}$
Case III	**Case IV**
R defined on $\gamma = \{a < x < b\}$, $a < 0 < b$, $\lvert b\rvert < \lvert a\rvert$.	R defined on $\gamma = \{a < x < b\}$, $a < b < 0$.
$\varphi(y_1) = \begin{cases} \dfrac{\log(a^2b^2) - 2\log\lvert y_1\rvert}{(b-a)^2}, & ab \le y_1 \le 0, \\ 0, & y_1 = 0, \\ \dfrac{\log(a^2b^2) - 2\log\lvert y_1\rvert}{(b-a)^2}, & 0 \le y_1 \le b^2, \\ \dfrac{\log (a^2) - \log \lvert y_1\rvert}{(b-a)^2}, & b^2 < y_1 \le a^2. \end{cases}$	$\varphi(y_1) = \begin{cases} \dfrac{\log a^2 - \log \lvert y_1\rvert}{(b-a)^2}, & a^2 < y_1 \le ab, \\ \dfrac{\log(b^2) - \log\lvert y_1\rvert}{(b-a)^2}, & ab < b^2. \end{cases}$

TABLE III

The Probability Density Functions of $Y_1=R_1/R_2$ for Identically Distributed Rectangular Variables Defined on the Interval $\gamma = \{a \le x \le b\}$. (see section 4.3.2, chapter IV)

Case I R defined on $\gamma = \{a \le x \le b\}$, $0 < a < b$.	Case II R defined on $\gamma = \{a \le x \le b,\ a < 0 < b,\ \lvert b \rvert > \lvert a \rvert\}$.
$\varphi(y_1)=\begin{cases} \dfrac{y_1^2 b^2 - a^2}{2 y_1^2 (b-a)^2}, & \dfrac{a}{b} < y_1 < 1, \\ \dfrac{b^2 - a^2 y_1^2}{2 y_1^2 (b-a)^2}, & 1 < y_1 < \dfrac{b}{a}. \end{cases}$	$\varphi(y_1)=\begin{cases} \dfrac{b^2 + a^2}{2 y_1^2 (b-a)^2}, & -\infty < y_1 \le \dfrac{b}{a}, \\ \dfrac{a^2 (y_1^2 + 1)}{2 y_1^2 (b-a)^2}, & \dfrac{b}{a} < y_1 \le \dfrac{a}{b}, \\ \dfrac{a^2 + b^2}{2(b-a)^2}, & \dfrac{a}{b} < y_1 \le 1, \\ \dfrac{a^2 + b^2}{2 y_1^2 (b-a)^2}, & 1 < y_1 < \infty. \end{cases}$
Case III R defined on $\gamma = \{a \le x \le b\}$, $a < 0 < b$, $\lvert b \rvert < \lvert a \rvert$.	Case IV R defined on $\gamma = \{a \le x \le b\}$, $a < b < 0$.
$\varphi(y_1)=\begin{cases} \dfrac{b^2 + a^2}{2 y_1^2 (b-a)^2}, & -\infty < y_1 \le \dfrac{a}{b}, \\ \dfrac{b^2 (1 + y_1^2)}{2 y_1^2 (b-a)^2}, & \dfrac{a}{b} < y_1 \le \dfrac{b}{a}, \\ \dfrac{a^2 + b^2}{2(b-a)^2}, & \dfrac{b}{a} < y_1 \le 1, \\ \dfrac{a^2 + b^2}{2 y_1^2 (b-a)^2}, & 1 < y_1 < \infty. \end{cases}$	$\varphi(y_1)=\begin{cases} \dfrac{y_1^2 b^2 - a^2}{2 y_1^2 (b-a)^2}, & \dfrac{a}{b} < y_1 \le 1, \\ \dfrac{b^2 - a^2 y_1^2}{2 y_1^2 (b-a)^2}, & 1 < y_1 < \dfrac{b}{a}. \end{cases}$

TABLE IV

The Probability Density Functions of $Y_1 = T_1 \cdot T_2$ for Identically Distributed Triangular Variables Defined on the Interval $\gamma = \{ 2a \leq x \leq 2b \}$.*

Case I T defined on $\gamma = \{-2 \leq x \leq 2\}$.	Case II T defined on $\gamma = \{0 \leq x \leq 2\}$.
$\varphi(y_1) = \begin{cases} \frac{1}{8}(y_1 + 4) \log \left\lvert \frac{y_1}{4} \right\rvert - (y_1 + 4), & -4 \leq y_1 < 0, \\ 0, & y_1 = 0, \\ \frac{1}{4}(y_1 - 4) - \frac{1}{8}(y_1 + 4) \log \left\lvert \frac{y_1}{4} \right\rvert, & 0 < y_1 \leq 4 \end{cases}$	$\varphi(y_1) = \begin{cases} 2 y_1 [1 - \log \lvert y_1 \rvert] + y_1 (\log \lvert \tfrac{1}{2} y_1 \rvert), & 0 < y_1 < 1, \\ 2(4 - 3y_1) + (4 + 2y_1) \log \lvert y_1 \rvert + y_1, & 1 < y_1 < 2, \\ (4 + y_1) \log 2 - (2 + y_1), & 2 < y_1 < 4. \end{cases}$

*see section 4.4.1, chapter IV

TABLE V

The Probability Density Function of $Y_1 = T_1/T_2$ for Identically Distributed Triangular Variables on the Interval $\gamma = \{2a \le x \le 2b\}$.*

Case I T defined on $\gamma = \{0 \le x \le 2\}$.	Case II T defined on $\gamma = \{-2 \le x \le 2\}$.
$\varphi(y_1) = \begin{cases} \frac{1}{6} y_1, & 0 \le y_1 \le \frac{1}{2}, \\ \frac{8}{3} - \frac{3}{2} y_1 - \frac{2}{3 y_1^2} + \frac{1}{6 y_1^3}, & \frac{1}{2} < y_1 \le 1, \\ \frac{5}{6} y_1 + \frac{2}{y_1^2} - \frac{5}{6 y_1^3} - \frac{4}{3}, & 1 < y_1 \le 2, \\ \frac{7}{6 y_1^3}, & 2 < y_1 < \infty. \end{cases}$	$\varphi(y_1) = \begin{cases} \frac{1}{3}\left(\frac{1}{2 y_1^3} + \frac{1}{y_1^2}\right), & -\infty < y_1 \le 1, \\ \frac{1}{3} + \frac{1}{6} y_1, & -1 < y_1 \le 0, \\ \frac{1}{3} - \frac{1}{6} y_1, & 0 < y_1 \le 1, \\ \frac{1}{3}\left(\frac{1}{y_1^2} - \frac{1}{2 y_1^3}\right), & 1 < y_1 < \infty. \end{cases}$

* see section 4.4.2, chapter IV

TABLE VI

The Distribution of the Product $Y_1 = R \cdot N$ in which R is defined on $\gamma = \{a < x_1 < b\}$ and N is N(0, 1). (see section 4.5.1.1, chapter IV)

Case I. R is defined on $\gamma = \{a < x_1 < b\}$, $0 < a < b$.	Case II. R is defined on $\gamma = \{a < x_1 < b\}$, $\lvert a \rvert < \lvert b \rvert$.
$\varphi(y_1) = \dfrac{2}{\sqrt{2\pi}\,(b-a)} \displaystyle\int_{\frac{y_1^2}{2b^2}}^{\frac{y_1^2}{2a^2}} \frac{e^{-v}}{v}\,dv, \; -\infty < y_1 < \infty .$	$\varphi(y_1) = \dfrac{2}{\sqrt{2\pi}\,(b-a)} \left[\displaystyle\int_{\frac{y_1^2}{2a^2}}^{\infty} \frac{e^{-v}}{v}\,dv + \int_{\frac{y_1^2}{2b^2}}^{\infty} \frac{e^{-v}}{v}\,dv \right],$ $-\infty < y_1 < \infty .$
Case III. R is defined on $\gamma = \{a < x_1 < b\}$, $\lvert a \rvert > \lvert b \rvert$	Case IV. R is defined on $\gamma = \{a < x_1 < b\}$, $a < b < 0$
Same as Case II.	$\varphi(y_1) = \dfrac{2}{\sqrt{2\pi}\,(b-a)} \displaystyle\int_{\frac{y_1^2}{2a^2}}^{\frac{y_1^2}{2b^2}} \frac{e^{-v}}{v}\,dv,$ $-\infty < y_1 < \infty .$

TABLE VII

The Probability Density Functions of $Y_1 = R/N$ for $N(\mu, \sigma^2)$ and R is Rectangularly Distributed on the Interval $\gamma = \{a < x < b\}$ as Indicated Below. (see section 4.5.2.1, chapter IV)

Case I R Defined on $\gamma = \{a < x < b\}$, $0 < a < b$	Case II R defined on $\gamma = \{a < x < b\}$ $a < 0 < b$ $\lvert a\rvert < \lvert b\rvert$
$\varphi(y_1) = \dfrac{\sigma^2}{(b-a)\sqrt{2\pi}}\left[e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2} - e^{-\frac{1}{2}\left(\frac{z'-\mu}{\sigma}\right)^2}\right] - \dfrac{\mu}{(b-a)\sqrt{2\pi}}\displaystyle\int_{\frac{z'-\mu}{\sigma}}^{\frac{z-\mu}{\sigma}} e^{-\frac{1}{2}u^2}\,du,\ -\infty < y_1 < \infty.$ for $\left(z = \dfrac{a}{y_1},\ z' = \dfrac{b}{y_1}\right)$	$\varphi(y_1) = \dfrac{\sigma^2}{(b-a)\sqrt{2\pi}}\left[2e^{-\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2} - e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2} - e^{-\frac{1}{2}\left(\frac{z'-\mu}{\sigma}\right)^2}\right] - \dfrac{\sigma^2}{(b-a)\sqrt{2\pi}}\left[\displaystyle\int_{\frac{z'-\mu}{\sigma}}^{-\frac{\mu}{\sigma}} e^{-\frac{1}{2}\mu^2}\,du - \int_{\frac{\mu}{\sigma}}^{\frac{z-\mu}{\sigma}} e^{-\frac{1}{2}u^2}\,du\right], -\infty < y_1 < 0,$ $= \dfrac{\sigma^2}{(b-a)\sqrt{2\pi}}\left[2e^{-\frac{1}{2}\left(\frac{\mu}{\sigma}\right)^2} - e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2} - e^{-\frac{1}{2}\left(\frac{z'-\mu}{\sigma}\right)^2}\right] - \dfrac{\mu}{(b-a)\sqrt{2\pi}}\left[\displaystyle\int_{\frac{z-\mu}{\sigma}}^{-\frac{\mu}{\sigma}} e^{-\frac{1}{2}\mu^2}\,du - \int_{-\frac{\mu}{\sigma}}^{\frac{z'-\mu}{\sigma}} e^{-\frac{1}{2}\mu^2}\,du\right],$ $0 < y_1 < \infty.$
Case III R Defined on $\gamma = \{a < x < b\}$ $a < 0 < b$ $\lvert a\rvert > \lvert b\rvert$	**Case IV** R defined on $\gamma = \{a < x < b\}$ $a < b < 0$
$\varphi(y_1)$ same as in Case II	$\varphi(y_1) = \dfrac{\sigma^2}{(b-a)\sqrt{2\pi}}\left[e^{-\frac{1}{2}\left(\frac{z'-\mu}{\sigma}\right)^2} - e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2}\right] - \dfrac{\mu}{(b-a)\sqrt{2\pi}}\displaystyle\int_{\frac{z-\mu}{\sigma}}^{\frac{z'-\mu}{\sigma}} e^{-\frac{1}{2}\mu^2}\,du,\ -\infty < y_1 < \infty.$

TABLE VIII

Standardized Lognormal Percentage Points

Given the mean m and standard deviation , let t = 100 s/m. The percentage point of the lognormal distribution with this mean and standard deviation is the entry in the table, multiplied by m.*

Lower 1% points (k = 1)

t	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Δ
0	1.0000	9977	9954	9930	9907	9884	9861	9838	9815	9792	-23
1	0.9770	9747	9724	9701	9679	9656	9633	9611	9588	9566	-23
2	0.9544	9521	9499	9477	9454	9432	9410	9388	9366	9344	-22
3	0.9322	9300	9278	9256	9233	9213	9169	9148	9148	9126	-22
4	0.9105	9083	9061	9040	9019	8997	8976	8955	8933	8912	-21
5	0.8892	8870	8849	8828	8807	8787	8766	8745	8724	8702	-21
6	0.8633	9662	8641	8621	8600	8580	8559	8539	8519	8498	-20
7	0.8478	8458	8438	8418	8398	8378	8357	8338	8318	8298	-20
8	0.8278	8358	8238	8219	8199	8179	8160	8140	8121	8101	-20.
9	0.8082	8062	8043	8024	8004	7985	7966	7947	7928	7909	-19
10	0.7890	7871	7852	7833	7814	7795	7776	7758	7739	7720	-19
11	0.7702	7683	7665	7646	7628	7609	7591	7572	7554	7536	-18
12	0.7518	7500	7481	7462	7445	7427	7409	7391	7374	7356	-18
13	0.7338	7320	7302	7285	7267	7249	7232	7214	7197	7179	-17
14	0.7162	7144	7127	7110	7093	7075	7058	7041	7024	7007	-17
15	C.6900	-	-	-	-	-	-	-	-	-	-

Lower 5% points (k = 5)

t	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Δ
0	1.0000	9983	9967	9951	9934	9918	9902	9885	9860	9853	-16
1	0.9836	9820	9804	9788	9771	9755	9739	9723	9707	9691	-16
2	0.9675	9658	9642	9626	9610	9594	9578	9562	9546	9530	-16
3	0.9514	9498	9482	9467	9451	9435	9419	9403	9388	9372	-16
4	0.9356	9340	0324	9309	9293	9278	9262	9246	9231	9215	-16
5	0.9200	9184	9168	9153	9137	9122	9106	9091	9075	9060	-15
6	0.9045	9029	9014	8999	8983	8968	8953	8937	8922	8907	-15
7	0.8892	8877	8862	8846	8831	8816	8801	8786	9771	8756	-15
8	0.8741	8726	8711	8696	8681	8666	8651	8636	8622	8607	-15
9	0.9592	8577	8562	8548	8533	8518	8503	8489	8474	8460	-15
10	0.8445	8430	8415	8401	8386	8372	8357	8343	8328	8314	-15
11	0.8299	8235	8271	8256	8242	8228	8213	8199	8185	8170	-14
12	0.8156	8142	8128	8113	8099	8085	8071	8057	8043	8029	-14
13	0.8015	8001	7987	7973	7959	7045	7931	7917	7903	7889	-14
14	0.7875	7861	7848	7834	7820	7806	7793	7779	7765	7751	-14
15	0.7738	-	-	-	-	-	-	-	-	-	-

*see chapter V

Upper 1% points(k = 99)											
t	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Δ
0	1.0000	0023	0047	0070	0094	0117	0141	0164	0188	0211	+23
1	1.0235	0259	0282	0306	0330	0354	0378	0402	0426	0450	+24
2	1.0474	0498	0522	0547	0571	0595	0620	0644	0669	0693	+24
3	1.0718	0743	0767	0792	0817	0841	0866	0891	0916	0941	+25
4	1.0966	0991	1016	1041	1067	1092	1117	1142	1168	1193	+25
5	1.1219	1244	1279	1296	1321	1347	1372	1398	1424	1450	+26
6	1.1476	1502	1528	1554	1580	1606	1632	1659	1685	1711	+26
7	1.1738	1764	1790	1817	1843	1870	1896	1923	1950	1977	+27
8	1.2004	2030	2057	2084	2111	2138	2165	2193	2220	2247	+27
9	1.2274	2302	2329	2356	2384	2411	2439	2466	2494	2522	+28
10	1.2549	2577	2605	2633	2661	2689	2716	2745	2773	2801	+28
11	1.2829	2857	2885	2914	2942	2970	2999	3027	3056	3084	+28
12	1.3113	3141	3170	3199	3228	3257	3285	3314	3343	3372	+29
13	1.3402	3431	3460	3489	3518	3547	3577	3606	3636	3665	+29
14	1.3695	3724	3753	3785	3813	3843	3872	3902	3932	3962	+30
15	1.3992	-	-	-	-	-	-	-	-	-	-

Upper 5% points (k = 95)											
t	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Δ
0	1.0000	0016	0033	0049	0066	0083	0099	0115	0132	0149	+17
1	1.0165	0182	0198	0215	0232	0249	2065	0282	0299	0316	+17
2	1.0332	0349	0366	0383	0400	0417	0433	0450	0467	0484	+17
3	1.0501	0518	0535	0552	0569	0586	0603	0620	0637	0654	+17
4	1.0671	0688	0705	0723	0740	0757	0774	0791	0809	0826	+17
5	1.0843	0860	0878	0895	0912	0930	0947	0964	0982	0999	+17
6	1.1017	1034	1051	1069	1086	1104	1121	1139	1156	1174	+17
7	1.1191	1209	1226	1244	1262	1279	1297	1315	1332	1350	+18
8	1.1368	1385	1403	1421	1439	1456	1474	1492	1510	1528	+18
9	1.1545	1563	1581	1599	1617	1635	1653	1671	1689	1707	+18
10	1.1725	1742	1760	1779	1797	1815	1833	1851	1869	1887	+18
11	1.1905	1923	1941	1959	1978	1996	2014	2032	2050	2069	+18
12	1.2087	2104	2123	2142	2160	2178	2196	2215	2233	2252	+18
13	1.2270	2288	2307	2325	2343	2362	2380	2399	2417	2436	+18
14	1.2454	2472	2491	2510	2528	2547	2565	2584	2602	2621	+19
15	1.2639	-	-	-	-	-	-	-	-	-	-

TABLE IX

Expected Values of $(1+\alpha w)^r$ For Various r.
(See chapter 5.)

	r	w is distributed normally with mean zero and variance one	w is distributed rectangularly in the interval (-1, 1)
$E(1+\alpha w)^r$	½	$1-\alpha^2/8-15\alpha^4/128-,\ldots\ldots$	$1-\alpha^2/24 - \alpha^4/128 - \ldots\ldots$
	1	1	1
	2	$1+\alpha^2$	$1+\alpha^2/3$
	4	$1+6\alpha^2+3\alpha^4$	$1+2\alpha^2+\alpha^4/5$
	-½	$1+3\alpha^2/8 + 105\alpha^4/128 + .$	$1+\alpha^2/8 + 7\alpha^4/128 + \ldots.$
	-1	$1+\alpha^2 + 3\alpha^4+ \ldots..$	$1+\alpha^2/3 + \alpha^4/5 + \ldots..$
	-2	$1+ 3\alpha^2+ 15\alpha^4 + \ldots..$	$1/(1-\alpha^2)$
	-4	$1+10\alpha^2+ 105\alpha^4+ \ldots..$	$1+10\alpha^2/3 + 7\alpha^4+ \ldots.$
$E\ \log(1+\alpha w)^r$		$-r(\alpha^2/2 + 3\alpha^4/4 + \ldots)$	$-r(\alpha^2/6 + \alpha^4/20 + \ldots)$
$V\ \log(1+\alpha W)^r$		$r^2(\alpha^2+ 5\alpha^4/2 + \ldots.)$	$r^2(\alpha^2/3 + 7\alpha^4/45 + \ldots)$

TABLE X

1. Tables of the cumulative distribution funct on of the random variable $Z = \frac{X_1}{\sigma_1}\frac{X_2}{\sigma_2}$, where X_i are normally distributed.

The p.d.f. of Z is $f(z) = I_1(z) - I_2(z)$ in which

$$I_1(z) = \frac{e^{\frac{-v_1^2 - 2\rho v_1 x_2 + v_2}{2(1-\rho^2)} + \frac{\rho z}{1-\rho^2}}}{2\pi\sqrt{1-\rho^2}} \cdot$$

$$\cdot \int_0^{\infty} \exp\left\{-\frac{1}{2(1-\rho^2)}\left(x^2 + \frac{z^2}{x^2}\right) + \frac{1}{1-\rho^2}\left[(v_1-\rho v_2)\,x + (v_2-\rho v_1)\frac{z}{x}\right]\right\}\frac{dx}{x}$$

and $I_2(z)$ is the same integral defined on the interval $(-\infty, 0)$.*

2. Values of $f(z)$, the probability density function, and $F(z)$, the cumulative distbibution function of z for various values of $v_1 = \frac{\mu_{x_1}}{\sigma_{x_1}}$, $v_2 = \frac{\mu_{x_2}}{\sigma_{x_2}}$ and the coefficient of correlation ρ.

* see section 4.2, chapter IV

TABLE X.a

Case 1: $\rho = 0$, $v_1 = 0$, $v_2 = 0$. The p.d.f. of f(z) is symmetrically distributed about the singularity point $z = 0$ with $E(z) = 0$, $\sigma_z = 1$. In this case $F(z) = 1 - F(-z)$.

Z	f(z)	F(z)
-10.0	.00001	.00001
- 9.0	.00002	.00002
- 8.0	.00005	.00004
- 7.2	.00011	.00019
- 6.4	.00026	.00021
- 5.6	.00061	.00056
- 4.8	.00146	.00139
- 4.0	.00355	.00323
- 3.6	.00557	.00509
- 3.2	.00878	.00796
- 2.8	.01395	.01231
- 2.4	.02235	.01943
- 2.0	.03625	.03090
- 1.8	.04645	.03914
- 1.6	.05983	.04969
- 1.4	.07756	.06336
- 1.2	.10138	.08114
- 1.0	.13402	.10450
- 0.9	.15493	.11890
- 0.8	.17996	.13144
- 0.7	.21025	.15508
- 0.6	.24749	.17790
- 0.5	.29425	.20491
- 0.4	.35477	.23721
- 0.3	.43887	.27656
- 0.2	.55790	.32585
- 0.1	.77256	.45242
0	$+\infty$	.50000

TABLE X b

Case 2: $\rho = 0$, $v_1 = (1.0)(0)$, $v_2 = 0(1.0)$. The p.d.f., f(z) is symmetrically distributed about the singularity point z = 0 with $E(z) = 0$ and $\sigma_z = \sqrt{2}$. In this case $F(z) = 1-F(-z)$.

z	f(z)	F(z)
-12.0	.00001	.00001
-11.0	.00002	.00002
-10.0	.00005	.00004
- 9.0	.00012	.00013
- 8.0	.00029	.00034
- 7.2	.00058	.00069
- 6.4	.00118	.00139
- 5.6	.00239	.00282
- 4.8	.00485	.00576
- 4.0	.00981	.01166
- 3.6	.01397	.01642
- 3.2	.01992	.02319
- 2.8	.02846	.03287
- 2.4	.04078	.04672
- 2.0	.05873	.06662
- 1.8	.07069	.07949
- 1.6	.08533	.09509
- 1.4	.10340	.11397
- 1.2	.12595	.13690
- 1.0	.15460	.16496
- 0.9	.17195	.18129
- 0.8	.19193	.19948
- 0.7	.21519	.21984
- 0.6	.24270	.24273
- 0.5	.27593	.26866
- 0.4	.31736	.29833
- 0.3	.37159	.33278
- 0.2	.44891	.37380
- 0.1	.58215	.42535
0	$+\infty$	.50000

TABLE X.c

Case 3: $\rho = 0$, $v_1 = v_2 = \frac{1}{2}$. The p.d.f., $f(z)$, has parameters $E(z) = 0.25$ and $\sigma_z = \frac{\sqrt{6}}{2}$ and possesses a singularity at $z = 0$.

z	f(z)	F(z)
-9.6	.00001	.00001
-8.8	.00002	.00002
-8.0	.00004	.00005
-7.2	.00010	.00010
-6.4	.00023	.00026
-5.6	.00054	.00061
-4.8	.00128	.00112
-4.0	.00311	.00288
-3.6	.00488	.00448
-3.2	.00769	.00699
-2.8	.01221	.01097
-2.4	.01954	.01732
-2.0	.03165	.02756
-1.6	.05213	.05560
-1.2	.08809	.10436
-0.8	.15568	.16520
-0.4	.30423	.24302
-0.2	.47388	.32084
-0.1	.64994	.37703
-0.0	∞	.50000
0.1	.68106	.50855
0.2	.51947	.56858
0.4	.36322	.65685
0.8	.21768	.77030
1.2	.14230	.85791
1.6	.09621	.89273
2.0	.04589	.92520
2.4	.06614	.94760
2.8	.04589	.96318
3.2	.02241	.97407
3.6	.01571	.98177
4.0	.01103	.98712
4.8	.00545	.99371
5.6	.00269	.99697
6.4	.00133	.99858

APPENDIX B - ANNOTATED BIBLIOGRAPHY

The components of a quotient or product of a random variable assume many different representations. Not only is one interested in ratios and products of random variables described by a specific frequency function, but also of these functional forms of variances, ranges, proportions, various forms of Mills' ratio, correlation coefficients, etc.

The following bibliography is believed to be very comprehensive in respect to articles pertaining to the distribution theory of products and quotients of random variables. However only a small sample of articles pertaining to other subjects as those mentioned above are given.

An index is presented in which listings appear by headings classifying many of the various combinational forms of quotients and products.

Index

1. Aroian, L. A., "A Study of R. A. Fisher's Z Distribution and the Related F Distribution," Annals of Math. Stat. Vol. 12, 1941.

The purpose of this paper is to discuss the semi-invariants of the Z distribution and to find useful approximations for them. The distribution of $Z = \frac{1}{2} \log \frac{s_1^2}{s_2^2}$, where s_i^2, $(i = 1, 2)$, are sample estimates of variances, is well known;

$$\rho(z) = \frac{2n_1^{\frac{1}{2}n_1} \; n_2^{\frac{1}{2}n_2}}{\beta\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \quad \frac{e^{n_1 z}}{\left(n_1 e^{2z} + n_2\right)^{\frac{1}{2}(n_1+n_2)}} dz.$$

The author shows that as n_1 and n_2 approach infinity in any manner the distribution of Z approaches normality.

2. Aroian, L. A., "The Probability Function of a Product of Two Normally Distribution Variables," Annals of Math. Stat. Vol. 18, 1947.

Let X_1 and X_2 follow a normal bivariate probability function with means $\bar{X}_1$ and $\bar{X}_2$, standard deviations σ_1 and σ_2, and correlation coefficient r_{x_1,x_2}. Also, let $\rho_1 = \bar{X}_1/\sigma_1$ and $\sigma_2 = \bar{X}_2/\sigma_2$. Craig [21] found the probability function of $Z = X_1 X_2 / \sigma_1 \sigma_2$ in closed form as the difference of two integrals. Craig, for the purposes of numerical computation, expanded this result in an infinite series involving powers of Z, ρ_1, ρ_2 and Bessel functions of a certain type. Difficulty arises, for large ρ_1 and ρ_2, in the convergence of this expansion. Aroian demonstrates that Z may be approximated by the Gram-Charlier series and the Type III function and presents the percentage points of F(z) for the special case $V_1 = 1$, $V_2 = 10$ and $r = 0$.

3. Aroian, L.A., Tables and Percentage Points of the Distribution Function of a Product, Hughes Aircraft Company, California, 1957.

Let X_1 and X_2 be normal uncorrelated variables with means m_1 and m_2 and variances σ_1 and σ_2, respectively, and let $\delta_1 = m_1/\sigma_1$. The author finds: (1) The 100 α percentage points of the distribution of $Z = X_1X_2/\sigma_1\sigma_2$ are tabulated for different values of δ_i ($i = 1,2$); (2) The cumulative distribution of Z is also tabulated.

4. Baker, G. A., "Distribution of the Means Divided by the Standard Deviations in Samples From Nonhomogeneous Populations", Annals of Math. Stat. Vol. 3, 1932.

It is the purpose of this paper to discuss the distribution of the means of samples(of size two) measured as the mean of the population which have been divided by the standard deviations of the samples. Experimental results for samples of four from normal populations are presented.

5. Baker, G. A., "Distribution of the Ratio of Sample Range to Sample Standard Deviation for Normal and Combinations of Normal Distributions," Annals of Math. Stat. Vol. 17, 1946.

This note summarizes, in tabular form, some of the numerical results obtained in previous studies of the distribution of sample ranges in terms of the standard deviation of the sampled population for homogeneous populations.

6. Banerjee, S, K., "A Lower Bound to the Probability of Student's Ratio", Sankhya, Vol. 18, 1957.

Samples of size n are drawn from a population having mean m, and 2nd and 4th cumulants k_2 and k_4 respectively. Starting with Tchebycheff's lemma that if a variate X has mean m and variance σ^2, then

$$P(X \leq m-k^2) \leq \frac{k^4}{\sigma^2 + k^4}$$

whatever k^2 may be, the behavior of Student's ratio is considered from the viewpoint of estimating the confidence interval. In particular, a lower bound to the probability of the event $\bar{x} + \frac{ts}{\sqrt{n}} \geq m \geq \bar{x} - \frac{ts}{\sqrt{n}}$ is calculated to be $\dfrac{1}{\dfrac{B_2-3}{n} + 1 + \dfrac{2}{(t^2-1)^2}\left\{\dfrac{t^4}{n-1} + 1\right\}}$, where $B_2 = \dfrac{k_4}{k_2^{\,2}} + 3$.

A table is given for $t = 3$ and several values of B_2 and n.

7. Barrow, D. T., and Cohen, A. C. Jr., "On Some Functions Involving Mills' Ratio," Annals of Math. Stat., Vol. 25, 1954.

In this note, it is proved that for all finite values of h,

$$\psi(h) = \frac{m_2}{m_1^{\,2}} = \frac{1-h(Z-h)}{(Z-h)^2}$$

is monotonic increasing so that $2\,m_1^{\,2} - m_2 > 0$ and $1 < \psi(h) < 2$. The function Z is the reciprocal of Mills' ratio and m_1 and m_2 are the first and second moments of a singly truncated normal population about a point of truncation. The function $\psi(h)$ frequently arises in connection with maximum likelihood estimates of population parameters from singly truncated normal samples.

8. Bartlett, M., "The Distribution of Second Order Moment Statistics in the Normal System," Proc. Camb. Phil. Soc., Vol. 28, 1932.

Let x be a normally distributed variable and the distribution is written as

$$dp = (2\pi k_2)^{-\frac{1}{2}}\, e^{-\frac{1}{2}x^2/k_2}\, dx$$

where k_2 is the semi-invariant of order 2. The moment generating function of the distribution of x^2 is developed. From this the generating function of the semi-invariants (k_r) of x^2 is developed.

9. Barton, D. E., David, F. N., and O'Neill, A., "Some Properties of the Logarithm of the Non-Central F", Biometrika, Vol. 47, 1960.

Given two independent variables: δ^2 distributed as a non-central chi square variable with f_1 degrees of freedom and non-central parameter and χ^2 distributed as a central chi square variable with f_2 degrees of freedom, the non-central F is defined as

$$F = \frac{(\delta^2/f_1)}{(\chi^2/f_2)}$$

The log transformation of F and the possibility of a normal approxamation to the transformation are discussed here.

10. Basu, D. "On the Dependence of Linear Functions of Independent Chance Variables," Bull. Inter. Stat. Inst., Vol. 33, 1953.

Relying on the assumption that the random variables $X_1, X_2, \ldots, X_n$ are independently distributed and possess finite moments of all order, the author proves the following: If there exist two linear functions $\zeta = \sum_{j=1}^{n} a_j x_j$, and $\xi = \sum_{j=1}^{n} b_j x_j$ with $a_j b_j \neq 0$ and ζ and ξ are stochastically independent, then each x_j is normally distributed.

11. Bliss, C. I., "The Calculation of Dosage-Mortality Data", Annals of Applied Biology, Vol. 22, 1935.

Very detailed applications of experimental data to the problem of estimating the quotient of two quantities subject to experimental error are given.

12. Bose, S., "On the Distribution of the Ratio of Variances of Two Samples Drawn From a Given Normal Bivariate Correlated Population," Sankhya, Vol. 2, 1935.

A formula for the p.d.f. of the bivariate chi distribution is given This formula is corrected by Krishnaiah, et. al. [70]. The

distribution of $Z = U/V$, the ratio of two correlated chi variates is given.

13. Boyd, A. V., "Bounds for Mills' Ratio for the Type III Distribution," <u>Annals of Math. Stat.</u>, Vol. 29, 1958.

Des Raj [101] and Cohen have shown that in estimating the parameters of the truncated Type III populations, it is necessary to calculate for several values of x, the Mills' ratio of the ordinate of the standardized Type III curve at x to the area under the curve from x to ∞. For large values of x, existing tables are inadequate. The object of this note is to establish lower and upper bounds for this ratio.

14. Broadbent, S. R., "Lognormal Approximations to Products and Quotients," <u>Biometrika</u>, Vol. 43, 1956.

The problem considered is to obtain bounds of limits upon combinations of random variables (random in the sense of measurement error) in a product or quotient form. The general form is considered to be

$$q = (x_1 \cdot x_2 \ldots x_n) / (x_{n+1} \cdot x_m \ldots x_r), \; 1 \leq n \leq r .$$

A summary of the known exact results is first given. The lognormal distribution which is asymptotically exact is shown to give useful approximations when fitted by moments to the combination above. Tables are given which make its application relatively simple.

15. Broadbent, S. R., "The Quotient of a Rectangular or Triangular and a General Variate," <u>Biometrika</u>, Vol. 41, 1954.

See Chapter IV of this report.

16. Buehler, R. J., "Confidence Intervals for the Product of Two Binomial Parameters," Journal of the American Stat. Assoc., Vol. 52, 1957.

The author generalizes the important practical problem of obtaining confidence limits on products of binomial parameters. The analysis is specialized to small probabilities of failure and moderate sample sizes. The results are restricted to the two-parameter case and no efforts are reported on the general case of the product of n parameters. The use of inequalities forms the basis of this analysis.

17. Chung, Kai-Lai, "The Approximate Distribution of Student's Distribution," Annals of Math. Stat., Vol. 17, 1946.

The asymptotic expansion of the distribution of certain statistics in a series of powers $n^{-1/2}$ with a remainder term gives an estimate of the accuracy of the normal approximation to that distribution. H. Cramér first obtained the asymptotic expansion of the mean and P. L. Hsu obtained that of the variance. In this paper, the methods of Cramér-Hsu are applied to Student's statistic. The important results are that if certain conditions are met by the population distribution, an asymptotic expansion may be obtained for this statistic.

18. Cohen, A. C. See Barrow [7]

19. Craig, C. C., "The Frequency Function of X/Y," Annals of Math., Vol. 30, 1929.

At this time (1929) only Pearson [97] had studied y/x for known distributions. Pearson investigated the problem by attempting to find expressions for the moments of y/x in terms of the known characteristics of the joint distribution function F(x, y), such as the coefficients of variation, etc. Craig extends his work by obtaining expressions for the semi-invariants. Very tedious and cumbersome mathematical expressions arise which hinder application. (See Chapter IV)

20. Craig, C. C., "On the Frequency Function of XY," Annals of Math. Stat., Vol. 7, 1936.

The p.d.f. of a product of a pair of normally distributed random variables is discussed under the following conditions: Let X_1 and X_2 follow a joint normal bivariate p d.f. with mean $\bar{x}_1$, $\bar{x}_2$, standard deviations σ_1, σ_2 respectively, and correlation coefficient ρ. Let $V_1 = \bar{x}_1/\sigma_1$ and $V_2 = \frac{\bar{x}_2}{\sigma_2}$. Craig finds the p.d.f. of $z = x_1x_2/\sigma_1\sigma_2$ in closed form as the difference of two integrals. (See Section 4.2)

21. Craig, C. C., "On the Frequency Distribution of the Quotient and the Product of Two Statistical Variables," American Math. Monthly, Vol. 49, 1942.

This work is divided into three sections: 1) the author discusses the relatively simple methods of finding the quotient y/x in which both x and y are distributed in a bivariate normal distribution, 2) a treatment of the distribution XY is given and, 3) a discussion of the quotient $\alpha_3 = m_3/m_2^{3/2}$ where m_2 and m_3 are the second and third central moments calculated from a sample of N observations is given.

22. Cramér, H., Mathematical Methods of Statistics, Princeton University Press, 1957.

23. Creasy, M. A., "Limits of the Ratio of Means," Journal of the Royal Stat. Soc., (Series B), Vol. 10, 1954.

Limits for the ratio of the means of two normal p.d.f.'s are obtained by a method which is analogous to that used in finding fiducial limits for the difference of two means with possibly unequal variances, (the Fisher-Behrens problem).

24. Curtiss, J. H., "On the Distribution of the Quotient of Two Chance Variables," Annals of Math. Stat., Vol. 12, 1941.

A rigorous treatment of methods of finding the distribution of the ratio $Z = X_1/X_2$ for known distributions of X_1 and X_2 is given. Particular attention is paid to existence proofs in the presentation of the variable transformation method of finding the distribution of Z in terms of the joint density function of (x_1, x_2). Four important formulae concerning the transformation of variables method of finding F(Z) are given.

25. Darmois, G., "Sur Une Propriété Caractéristique de la Loi de Probabilité de Laplace," C. R. Acad. Science, Paris, Vol. 239, 1951.

A remarkable characterization of the normal p.d.f. is that if x_1 and x_2 are two independent chance variables such that two linear functions $ax_1 + bx_2$ $(ab \neq 0)$ and $cx_1 + dx_2$ $(cd \neq 0)$ are distributed independently, then both X_1 and X_2 are normally distributed. This theorem has been proved without any assumption about the existence of moments by the author using earlier results of Gnedenko and Kac [63].

26. David, H. A., Hartley, H.O. and Pearson, K., "The Distribution of the Ratio, in a Single Normal Sample, of Range to Standard Deviation," Biometrika, Vol. 41, 1954.

The author presents certain percentage points of the ratio of samples of n observations from an underlying normal population. The ratio concerned is a) range, w, of a sample of n observations to b) the standard deviation, s, where both w and s are calculated from the same sample of n observations. The method of analysis uses the moments of the distribution of w/s.

27. David, F. N., "Reciprocal Bernoulli and Poisson Variables," Anais de Faculdade de Ciencios do Porto, Vol. 37, 1953.

Tables of the p.d.f. of $\sigma E\left(\frac{1}{x}\right)$ where x is a non-zero random variable described by the positive Bernoulli and Poisson frequency functions are presented.

28. David, F. N. See Barton [9]

29. DeGroot, M. H., "Some Aspects of the Use of 'Sequential' Probability Ratio Tests," Journal of the American Stat. Assoc., Vol 53, 1958.

This paper investigates the use of Wald's sequential probability ratio test as to the circumstances of applying the test when there is no known population variance. The behavior of the test when an erroneous value is taken for the value of the variance is investigated in two applications. Additionally, the use of the test is discussed for applications in which the variance can be restricted to a given finite interval.

30. Doob, J. L., "Limiting Distributions of Certain Statistics," Annals Of Math Stat., Vol. 6, 1935.

A rigorous treatment of some of the fundamental concepts of frequency distributions is first presented. This presentation provides a foundation for the treatment of the derivation of a limiting distribution. Many fundamental lemmas dealing with the concepts of a limiting distribution are derived.

31. Epstein, Benjamin, "Some Applications of the Mellin Transformation in Statistics," Annals of Math. Stat., Vol. 19, 1948.

It is well known that the Fourier transform is a useful analytic tool for studying the distribution of the sum of independent random variables. It is the purpose of this paper to study the Mellin Transform in relation to products of independent random variables of the form $\eta = \eta_1 \eta_2$, where η_i is defined on positive intervals. In this paper Epstein uses the Mellin transform to obtain $\eta = N_1 N_2$ as a Bessel function.

32. Fieller, E. C., "The Biological Standardization of Insulin," Journal of the Royal Stat. Soc., (Series B), Vol. 7, 1940.

Results involving quotients of random variables and the p.d.f. of an index are used in an application of a biological assay of insulin.

33. Fieller, E. C., "The Distribution of an Index of a Normal Bivariate Population," Biometrika, Vol. 20, 1932.

This is a notable paper concerning the ratio of two normally distributed random variables. By quadrature, Fieller presents a solution to the p.d.f. of $Y_1 = N_1/N_2$ for the general case. The normal case is expressible in terms of the bivariate normal p.d.f.

34. Finney, D.J., "The Distribution of the Ratio of Estimates of Two Variances in a Sample From a Normal Bivariate Population," Biometrika, Vol. 30, 1938.

Finney shows the p.d.f. of the ratio $Z = U/V$ of two correlated chi variates. From this result, he obtains the c.d.f. of Z and discusses certain pertinent applications.

35. Finney, D. J., "On the Distribution of a Variate Whose Logarithm is Normally Distributed," Journal of the Royal Stat, Soc., (B), Vol. 7, 1940.

The problem considered here is to determine efficient estimates of both the mean and the variance of a given population from the sufficient

statistics for the normal distribution of the transformed data where the logarithm of the observed values is taken as the transformation (the distribution of the logs is known to be normal).

The estimates for the mean and variance are given in terms of an infinite series (actually a Bessel function) with slow convergence except for small values of parameters involved, and these are modified for better arithmetic computation.

The efficiency of estimations without considering the above type transformation is also discussed.

36. Fisher, R. A., and Tippett, L., "Limiting Forms of the Frequency Distribution of the Smallest and Largest Number in a Sample," Proceedings of the Camb. Phil. Soc., Vol 24, 1927.

Constants for obtaining the first four moments of the distribution of the largest member of a sample from a normal population for samples up to 1000 are given. Possible limiting forms of such distributions in general, are discussed. Many tabular and graphic illustrations are presented.

37. Furstenberg, H., and Kesten, H., "Products of Random Matrices," Annals of Math. Stat., Vol. 31, 1960.

R. Bellman, "Limit Theorems for Non-Commutative Operations," Duke Math. Journal, 1954, considers this problem: Let X^1, X^2, X^3, ..., form a stationary stochastic process with values in the set of k x k matrices. He investigated the asymptotic behavior of ${}^nY^1 = X^n X^{n-1} \ldots X^1$ and showed that if the X^n are independent, then the $\lim_{n \to \infty} n^{-1} E \{\log({}^nY^1)i,j\}$ exists, (i, j refers to the i j^th^ entry of the matrix.) Bellman conjectured that $n^{-1/2}(\log({}^nY^1)i, j - E\{\log({}^nY^1)\ i, j\})$ is asymptotically normally distributed. This paper adds more evidence that Bellman's conjecture is correct.

38. Geary, R. C., "The Frequency Distribution of the Quotient of Two Normal Variables," Journal of the Royal Stat. Soc., Vol. 93, 1930.

Geary supplies an approximation of $q = x/y$ which is widely used. He shows that for X and Y normally distributed with respective means $\bar{X}$ and $\bar{Y}$ and Q defined as x/y, the statistic

$$Z = (q-Q)/(\alpha_1^2 q^2 + \alpha_2^2 Q^2)^{1/2}$$

is approximately normally distributed where α_i $(i = 1, 2)$ represents $\frac{\bar{x}}{\sigma_x}$ and $\frac{\bar{y}}{\sigma_y}$.

39. Geisser, S., "The Distribution of the Ratio of Certain Quadratic Forms in Time Series," Annals of Math. Stat., Vol. 28, 1957.

In testing hypotheses that successive members of a series of observations are serially correlated, a number of statistics have been proposed. Durbin and Watson gave exact distributions for some of these when slightly modified. This paper extends this work for a non-null case of two of these modified statistics and gives a simple expression for the moments of another.

40. Goodspeed, F. M., "The Relation Between Functions Satisfying A Certain Integral Equation and the General Watson Transform" Canadian Journal of Math., Vol. 2, 1950.

Goodspeed investigates the integral equation

$$\int_0^\infty F(x)\ F(u,x)dx = \frac{1}{1+u}$$

discussed in Chapter VI. In particular, a class of functions F(x) satisfying the equation is sought.

41. Gordon, R. D., "The Estimation of a Quotient When the Denominator Is Normally Distributed," Annals of Math. Stat., Vol. 12, 1941.

This investigation is devoted to obtaining an estimate of the means of components of ratio values obtained in a time series. The time series is of the form x/y. Of corresponding interest in the time series are mean values x and y, (true value, b/a). The search is for an estimate of (b/a) which has the property that it may be numerically integrated. Gordon defines a function $\gamma(x)$ so that

$$E\left\{y \cdot \gamma(x)\right\} = E(y) \cdot E\left\{\gamma(x)\right\} = (b)\left(\frac{1}{a}\right) = b/a.$$

The function derived by the author is a function of Mills' ratio.

42. Grab, E. L., "Tables of the Expected Value of 1/X for Positive Bernoulli and Poisson Variables," Journal of the American Stat. Assoc., Vol. 49, 1954.

Tables of the probability function of σ E $\left(\frac{1}{X}\right)$, where X is a non-zero random variable described by the positive Bernoulli and Poisson frequency functions, are presented. Methods of solutions are presented which are more efficient than the factorial series method first presented by Stephen [118] in 1945.

43. Greenwood, J. A., and Greville, T., "On the Probability of Attaining a Given Standard Deviation Ratio in an Infinite Series of Trials," Annals of Math. Stat., Vol. 10, 1938.

Investigations are directed toward finding the probability that the ratio of the deviation from the expected number of successes in a Binomial experiment of n trials, $(x_n - np)$, to the standard deviation $\sqrt{npq}$ (recomputed after each trial) will exceed some positive number k. The authors prove that if $t_n = (x_n - np)/\sqrt{npq}$, then for some n the probability $t_n \geq k$ is unity.

44. Greville, T. See Greenwood [43] .

45. Gumbel, E. J. See Keeney [64] .

46. Gumbel, E. J., "The Geometric Range for Distributions of Cauchy's Type," Annals of Math. Stat., Vol. 21, 1950.

The author considers the distribution of the reciprocal of the geometric range of large samples drawn from population characterized by the Cauchy (or nearly Cauchy) distribution. Brief tables and graphs of the probability functions are presented.

47. Gumbel, E. J., and Herbach, L. H., "The Exact Distribution of the Extremal Quotient," Annals of Math. Stat., Vol. 22, 1951.

A method of obtaining the exact distribution of the quotient $\wp$, of the extreme values found in a sample of $n \geq 2$ observations taken from the same distribution is presented. This is an extension of the author's first published work (1950) dealing with the asymptotic distribution of the extremal quotient. The authors, in this paper, consider the Laplace, Exponential, Gamma, Normal and Cauchy distributions as components of the quotient and present brief graphs of the probability function associated with the sample size n.

48. Gurland, J., "Inversion Formulae for the Distribution of Ratios," Annals of Math. Stat., Vol. 19, 1948.

Gurland gives theorems and proofs that the distribution function $G(x)$ of $(a_1 x_1 + a_2 x_2 + \cdots + a_n x_n)/(b_1 x_1 + b_2 x_2 + \cdots + b_n x_n)$ is obtainable if the characteristic function of the joint density function $(x_1, x_2, \ldots, x_n)$ is known. He also presents inversion formulae for $G(x)$ and shows the density function of $G(x)$. Certain other ratios of quadratic forms are investigated.

49. Hack, H. R. B., "An Empirical Investigation into the Distribution of the F-Ratio in Samples from Two Non-Normal Populations," Biometrika, Vol. 45, 1958.

Experimental data obtained from an agricultural experiment are used to estimate the distribution of the F-ratio selected from two highly skewed populations.

50. Hagis, P. See Krishnaiah [70].

51. Haldane, J. B. S., "Moments of the Distribution of Powers and Products of Normal Variates," Biometrika, Vol. 32, 1942.

Various distributions are found, the derivations of which are greatly detailed. They are: (1) the distribution of a cube of a normal (0, 1) variable, (2) the general case of the distribution of any power of a normal variable, (3) the distribution of the product of n independent normal variables, (4) the special case of the product of 3 normal independent random variables, (5) the product of two correlated normal variables, and (6) the Galton-Macalister distribution.

52. Hart, B. I., and VonNeumann, J., "Tabulation of the Probability of the Ratio of the Mean Square Successive Difference to the Variance," Annals of Math. Stat., Vol. 13, 1942.

VonNeumann determined the distribution of δ^2 / s^2, the ratio of the mean square successive difference to the variance estimate, for odd values of sample size n. In this paper, the probability function developed is evaluated for other values (specific) of n. The evaluation methods are dependent upon the Incomplete Beta functions.

53. Hartley, H. O., and Ross, A., "Unbiased Ratio Estimators," Nature, Vol. 174, 1954.

Ratio estimation used in sample surveys to estimate the population mean $\overline{Y}$ of a variate y with the help of the known population mean $\overline{X}$ of

some correlated variable X suffer from the defect that they are biased estimators and by amounts for which only approximate formulae are offered.

Confined here to simple random samples of n pairs y_i, x_i from a population of N pairs, various ratio estimators of $\bar{Y}$ can be formed such as $\tilde{y} = \bar{X}\ \frac{\bar{y}}{\bar{x}}$, $\hat{y} = \bar{X}\bar{r}$ where $\bar{y}$, $\bar{x}$, $\bar{r}$ are arithmetic means of samples of y_i, x_i and $\frac{y_i}{x_i}$, respectively. To get exact formulas for the biases one obtains $E(\hat{y}) = \bar{Y} - \mathrm{Cov}\left(\frac{y}{x}, x\right)$, $E(\tilde{y}) = \bar{Y} - \mathrm{cov}\left(\frac{\bar{y}}{\bar{x}}, x\right)$ which give bounds on the biases: $|\text{bias in } \hat{y}| \leq \sigma y/x\ \sigma_x$, $|\text{bias in } \tilde{y}| \leq \sigma \bar{y}/\bar{x} - \sigma_y$. These have previously been attained for large sample approximation.

54. Hartley, H. O. See David, H. A. [26].

55. Haviland, E. K., "On the Distribution Function of the Reciprocal of a Function and of a Function Reduced Mod. 1," American Journal of Math., Vol. 63, 1941.

Proofs are presented which show that under suitable conditions, the functions $1/x(t)$ and $x(t)$ possess asymptotic distribution functions if $x(t)$ does. $x(t)$ is a measur[illegible] real function defined for every t. This asymptotic distribution function is expressed in terms of the distribution function of $x(t)$ for certain specified cases.

56. Herbach, L. H., See Gumbel, [47].

57. Hess, J., See Kish [67].

58. Hirshfeld, H. O., "The Distribution of the Ratio of Covariance Estimates in Two Samples Drawn from Normal Divariate Populations," Biometrika, Vol. 29, 1937.

The question of testing whether there is a difference between correlation in two normal populations is considered. The distribution of the ratio of covariance estimates is found and a detailed discussion of the appropriateness of its use in this test is presented. The development of the distribution requires frequent use of the incomplete Beta function. Some agricultural experiments are detailed.

59. Hogg, R. V., "On Ratios of Certain Algebraic Forms," Annals of Math. Stat., Vol. 22, 1951.

Necessary and sufficient conditions for the stochastic independence of the ratio of two random variables and its denominator are given and this result is applied to special linear forms. More specifically let x, y be one dimensional random variables with joint density function g(x, y). Let $P(y \leq 0) = 0$ and assume the moment generating function $M(u, t) = E\left[\exp(ux + ty)\right]$ exists for $-T < u,\ t < T,\ T > 0$; then in order that y and $r = \frac{x}{y}$ be stochastically independent, it is necessary and sufficient that

$$\frac{\partial^k M(0,t)}{\partial u^k} = \frac{\dfrac{\partial^k M(0,0)}{\partial u^k}}{\dfrac{\partial^k M(0,0)}{\partial t^k}} \cdot \frac{\partial^k M(0,t)}{\partial t^k}, \quad k=0,1,2,\ldots$$

60. Huntington, E. V., "Frequency Distributions of Products and Quotients," Annals of Math. Stat., Vol. 10, 1939.

Proofs of four well known theorems are presented. These theorems result in a mathematical formulation for obtaining the distribution of a sum, product, quotient and difference of two random variables. As an example, a geometric proof is given to show that the distribution of the quotient $Z = x/y$ is given by

$$Q(z) = \int_{-\infty}^{\infty} \int_{0}^{\infty} f(z, y)\, g(y)\, y\, dy$$

where x and y, $(f(y) > 0)$, are independently distributed random variables.

61. James, G. S., "On the Accuracy of Weighted Means and Ratios," Biometrika, Vol. 43, 1956.

$x_1, \ldots, x_k$ are k quantities derived from observations which are independent and normally distributed about the same mean μ but with different variances $\lambda_1 \sigma_1^2, \lambda_2 \sigma_2^2, \ldots, \lambda_i \sigma_k^2$. The λ_i are known but estimates s_i^2 are used. A confidence limit for μ is required. Let w_i denote $1/\lambda_i \sigma_i^2$, the reciprocal -variance or weight of x_i, and let z_i denote $1/(\lambda_i \sigma_i^2)$. Allow $w = \sum w_i$ and $z = \sum z_i$. It is reasonable, in the absence of firm knowledge of the true weights to estimate $\hat{x} = \sum w_i x_i / w$. The distribution of $\mu = w^{1/2}(\hat{x} - \mu)$ depends on the ratio of w_i. The desired result is then to find a function $\mu\left(\frac{w_1}{w}, \frac{w_2}{w}, \ldots, \frac{w_k}{w}\right)$ such that

$$\Pr\left[|\mu| \leq \mu(w)\right] = P.$$

This function is found and tabulated for the case $k = 2$.

62. Johnson, N. L., "Some Notes on the Investigation of Heterogeometry in Interactions," Trabajos De Estadistica, Vol. 13, 1962.

In this paper certain important distribution theory of use in the Analysis-of-Variance model for analysis of interactions is developed.

63. Kac, M., "On a Characterization of the Normal Distribution," American Journal of Math., Vol. 61, 1939.

A simple characterization of the normal distribution is presented. The characterization is based on an invariant principle and admits a physical interpretation.

64. Keeney, R. D., and Gumbel, E. J., "The Extremal Quotient," Annals of Math. Stat., Vol. 21, 1950.

After defining the extremal quotient as the ratio of the largest to the absolute value of the smallest observation in a sample, the authors investigate the analytical properties of symmetrical, continuous and unlimited distributions through a ratio defined as a ratio of two non-negative variates with identical distributions. Among the important results is the fact that the logarithm of the extremal quotient is symmetrically distributed. Asymptotic distributions of the extremal quotients from the Cauchy and Exponential distributions are exhibited.

65. Kesten, H. See Furstenberg [37].

66. King, E. P. See Lukacs [82].

67. Kish, L., and Hess, I., "On Variances of Ratios and Their Differences in Multi-Stage Sampling," Journal of the American Stat. Assoc., Vol. 54, 1959.

The aim of this presentation is to discuss the computation for the variances for the estimators $r = x/y$ and $(r - r')$ where the random variables y and x are sample totals for two variables obtained from a multi-stage design. The variate x often represents the sample size which represents the simplest case of r. This and several other cases are considered. Several useful computational forms are presented for $var(r-r') = var\ (r) + var\ (r') - 2\ Cov\ (r,r')$.

68. Kotlarski, I., "On Random Variables Whose Quotient Follows the Cauchy Law," Colloquium Mathematicum, Vol. 7, 1960.

The question of describing a class $\mathfrak{X}$ of random variables possessing the property that the ratio of any two independent random variables from the class having the same frequency function follows the Cauchy distribution is considered. A method is given for constructing an arbitrary number of random variables belonging to $\mathfrak{X}$ along with the necessary and sufficient conditions for belonging to $\mathfrak{X}$. Mellin transforms are used throughout the analysis. It is shown that the random variables whose frequency functions are determined by the Mellin transforms which are solutions to the functional equation $h(z)\, h(-z) = \dfrac{1}{\cos \frac{z}{2} \pi}$, $(-1 < R_e(z) < 1)$, in the class of Mellin transforms of symmetrical random variables belong to the class $\mathfrak{X}$.

69. Krishnaiah, P. R., P. Hagis, and L. Steinberg, The Bivariate Chi Distribution, Technical Report #3, Applied Math. Dept. Remington Rand Univac, Philadelphia, 1961.

70. Krishnaiah, P. R., Hagis, P., and Steinberg, L, "A Note on the Bivariate Chi Distribution," SIAM Review, Vol. 5, 1963.

Several properties of the bivariate chi p.d.f. are discussed with some mention of possible applications. Moments of the joint p.d.f. are presented and the distributions of the sum of U and V and the ratio $Z = U/V$ are shown.

71. Kruskal, W., "The Monotonicity of the Ratio of Two Non-Central t Density Functions," Annals of Math. Stat., Vol. 25, 1954.

The ratio of two different non-central t density functions with the same number of degrees of freedom is shown to be strictly monotone with sense depending on the relative values of the two non-central constants. The author also cites several statistical applications in which the ratio of two non-central t density functions arise.

72. Kuh, E., and Mayer, J., "Correlation and Regression Estimates When the Data are Ratios," Econometrica, Vol. 23, 1955.

The authors examine the influence of ratio transformations on correlation and regression estimates. After a discussion of the "spurious" ratio correlation problem, necessary and sufficient conditions are deduced for the correlation between two series with a common denominator to equal the partial correlation between numerator series with the deflating variable's influence held constant. It is shown that conditions must be fulfilled to obtain the best linear unbiased least squares estimates when the data are in ratio form.

73. Kullback, S., "The Distribution Laws of the Difference and Quotient of Random Variables Independently Distributed in Pearson Type III Laws," Annals of Math. Stat., Vol. 7, 1936.

The distribution of the difference u = x-y where

$$f(x) = \frac{e^{-x}x^{p-1}}{\Gamma(p)} \; ; \quad f(y) = \frac{e^{-y}y^{q-1}}{\Gamma(q)} \quad \text{for} \quad \begin{cases} 0 \leq x \leq \infty \\ 0 \leq y \leq \infty \end{cases}$$

is first found. The distribution function of the quotient $w = x/y$ is investigated through the utilization of the transformation $w = e^u$ where $u = \log x - \log y$. The distribution function P(u) is found and with appropriate substitution the solution of the distribution function of the quotient is found.

74. Laderman, J., "The Distribution of Student's Ratio for Samples of Two Items Drawn from Non-Normal Universes," Annals of Math. Stat., Vol. 10, 1939.

The formal expression for the distribution of student's t is derived from samples of two drawn from any population having a continuous frequency function. A geometric method similar to that used by Rider is used.

75. Laha, R. G., "An Example of a Non-Normal Distribution where the Quotient Follows the Cauchy Law," Proceedings of the National Academy of Science, Vol. 44, 1958.

For x, y, independently and identically distributed random variables, (the density function of x being $f(x) = \frac{\sqrt{2}}{\pi} \cdot \frac{1}{1+X^4}, -\infty < x < \infty,$), it is shown that the ratio x/y follows a Cauchy distribution thus proving untrue a previously well known conjecture that this was a unique property of normal variates.

76. Laha, R. G., "On the Laws of Cauchy and Gauss," Annals of Math. Stat., Vol. 30, 1959.

It is well known that x/y follows the Cauchy distribution when x and y are normal $(0, \sigma^2)$. The question arises, "is it possible to obtain a characterization of the normal distribution by this property of the quotient?". Laha constructs a class of functions which have the property that their ratios are Cauchy distributed but the components are non-normal. He also presents and proves several important lemmas concerning characterizing distributions.

77. Laha, R. G., "On a Class of Distribution Functions where the Quotient Follows the Cauchy Law," Transactions of the American Math.Soc., Vol. 93, 1959.

The author again concerns himself with the distribution of the ratio

of identically distributed random variables with a common distribution function $F(x)$. If the ratio follows the Cauchy law, is $F(x)$ necessarily normal? Laha constructs a different class of functions in this article.

78. Loève, M., "Fundamental Limit Theorems on Probability Theory," Annals of Math. Stat., Vol. 21, 1950.

The fundamental limit theorems on probability theory are classified into two groups. One group deals with the problems of limit laws of sequences of sums of random variables, the other deals with the problem of limits of random variables in the sense of almost sure convergence. This article is a detailed development of the significant results in each class as they were chronologically achieved.

79. Link, R. F., "The Sampling Distribution of the Ratio of Two Ranges From Independent Samples," Annals of Math. Stat.; Vol. 21, 1950.

Investigation of the joint sampling distribution of the ratio of ranges, w_1/w_2, where $w_i = (x_n - x_1)/\sigma$, are made for certain specified distributions. Tables of values of Prob $(w_1/w_2 \leq R)$ are presented. The underlying distributions considered are 1) Rectangular Distribution, 2) e^{-x}, and 3) Normal Distribution.

80. Link, R. F., "A Correction To: The Sampling Distribution of the Ratio of Two Ranges From Independent Samples," Annals of Math. Stat., Vol. 23, 1952.

A correction showing that the probabilities computed in the reference above are correct only when $R \leq 1$ is shown. P. R. Rider ("The Distribution of the Quotients of Ranges in Samples from a Rectangular Population," Journal. American Stat. Assoc., Vol. 45, 1951) gives the correct density function for $R \geq 1$.

81. Lukacs, E., and Szász, O., "Certain Fourier Transforms of Distributions," Canadian Journal of Mathematics, Vol. 3, 1951.

The question of when a given function can be the characteristic function of a probability distribution is considered for a restricted class of functions; namely, functions which are reciprocals of polynomials, with the view in mind of deriving conditions which are easy to apply. The following necessary conditions are derived:

If the reciprocal of a polynomial without multiple roots is a characteristic function, then the following two conditions are satisfied:

(1) The polynomial has no real roots. Its roots are either all on the imaginary axis or in pairs $\pm b + ia$ symmetric with regard to this axis.

(2) If $b + ia$ (a, b real; $a \neq 0$, $b \neq 0$) is a root of the polynomial then it has at least one root $i\alpha$ such that $\operatorname{sgn} \alpha = \operatorname{sgn} a$ and $|\alpha| \leq |a|$.

The assumption concerning multiple roots is used in deriving (2).

82. Lukacs, E., and King, E. P., "A Property of the Normal Distribution," Annals of Math. Stat., Vol. 25, 1954.

The authors prove the theorem under Basu [10] imposing slightly less restricted conditions. Their conditions require only that the random variables $x_1, x_2, \ldots, x_n$ may or may not be identically distributed and possess finite moments of order n.

83. Mann, H. B., and Whitney, D. R., "On a Test of Whether One Variable is Larger than the Other," Annals of Math. Stat., Vol. 18, 1947.

Let x and y be two random variables with continuous cumulative distribution functions F and G. A statistic u depending upon the relative ranks of the x's and y's is proposed in this paper for testing the hypothesis $F = G$. The conditions under which $f(u)$ is developed are stated and uses of the test of hypothesis are given.

84. Mayer, J. See Kuh [72].

85. Mauldon, J. G., "Characterizing Properties of Statistical Distributions," Journal of Math, Oxford, Vol. 7, 1956.

Three problems dealing with the general theme of when does the distribution of a sampling statistic determine that of the population are discussed. A representative problem is described as follows:

Consider X and Y to be independently distributed variates such that $r = X/Y$ has the general F distribution,

$$dF = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \cdot \frac{v^{b-1}}{(1+v)^{a+b}}\, dv, \quad 0 < v < \infty;\ a, b > 0 .$$

Is it still true that X and Y need not be chi-square variates?

86. Merrill, A. S., "Frequency Distribution of an Index when Both Components Follow the Normal Law," Biometrika, Vol. 20, 1928.

One of the earliest papers to investigate the deviation of error in the normal distribution approximation for the ratio of two random variables. In this paper, the author investigates the ratio of the form $(y + \bar{y})/(x + \bar{x})$ and, through graphic means, demonstrates that when correlation is high between x and y and the coefficients of variation large, there is a considerable deviation from normality.

87. Mickey, M. R., "Some Finite Population Unbiased Ratio and Regression Estimators," Journal of the American Stat. Assoc., Vol. 54, 1959.

A class of ratio and regression type estimators is given such that the estimators are unbiased for random sampling, without replacement, from a finite population. Non-negative, unbiased estimators of population variance are provided for a subclass of these. Similar results are given for the case of the generalized procedures of sampling without replacement. Efficiency is compared with comparable sample selection and estimation methods for this case.

88. Mills, J. P., "Table of the Ratio: Area of Bounding Ordinate for any Portion of the Normal Curve," Biometrika, Vol. 18, 1926.

The tabular values for $.00 \leq x \leq 10.00$ (in increments of .01) of

$$R_x = e^{\frac{1}{2} x^2} \int_x^{\infty} e^{-\frac{1}{2} x^2} dx$$

and a discussion of its derivation and many possible uses are presented.

89. Moore, P. G., "Transformations to Normality Using Fractional Powers of the Variable," Journal of the American Stat. Assoc., Vol.52, 1951.

An examination of the properties of a certain class of transformations is made under the assumption that they are designed to transform the variate X into some form of the normal distribution. The class of transformations is that of

$x^r, 0 < r < 1$ and non-negative X.

Some consideration is given to a Type III variate X, in respect to the use of $(X - \alpha)^r$ where α is also a random variable and is distributed uniformly on the interval $[0, 1]$.

90. Murty, V. N., "The Distribution of the Quotient of Maximum Values in Samples From a Rectangular Distribution," Journal of the American Stat. Assoc., Vol. 50, 1955.

When there is a priori knowledge that two samples have been drawn from rectangular populations with the same lower bounds, the hypothesis that both samples have been drawn from the same population can be tested by means of the quotient of maximum values. The distribution of this statistic is derived, and its properties studied. Explicit expressions for the power function of the test are given, and a table of 5% values of the quotient is presented for sample sizes up to ten.

91. Nair, K. R., "The Distribution of the Extreme Deviate from a Sample Mean and Its Standardized Form," Biometrika, Vol.35, 1948.

Denote $x_1, \ldots, x_n$ as a random sample of n observations drawn from any statistical universe so that the observations are ranked in ascending order of magnitude. The author first summarizes the efforts of other studies of the distribution of x_r. In this paper attention is concentrated on McKay's method of solution [Biometrika, Vol.27, p. 466]. It is shown that the distribution of $u = (x_n - x_1/\sigma)$, (termed the McKay statistic), can be found in a more direct way than that utilized by McKay. The distribution of u is reduced to certain integrals termed G-functions. Tabular values of f(u) are presented.

92. Newman, D., "The Distribution of the Range in Samples From a Normal Population in Terms of an Independent Estimate of the Standard Deviation," Biometrika, Vol. 31, 1939.

The author makes use of the results obtained by K. Pearson in computing probability levels for w/σ, where w = the sample range and σ = population variance, to determine appropriate corresponding levels of w/s where s is the independent estimate of σ based on the sample.

The distribution of q = w/s is obtained and factors to convert E(q) to $E(w/\sigma)$ are presented.

93. Nicholson, C., "A Geometric Analysis of the Frequency Distribution of the Ratio Between Two Variables," Biometrika, Vol.32, 1941.

An interpretation of the geometry of the distribution of $(Y+\bar{Y})/(X+\bar{X})$ where X, Y are normal $(0, \sigma_x, \sigma_y)$ and independent is presented. The case of dependency is given some consideration.

94. Noether, G. E., "Two Confidence Intervals for the Ratio of Two Probabilities and Some Measures of Effectiveness," Journal of the American Stat. Assoc., Vol.52, 1957.

The problem of comparing a success rate of hypothetical experimental method with that of a standard method is attached using the statistic

$$p = \frac{P_2 - P_1}{1 - P_1}$$

where P_i, $i = 1, 2$, represent the success ratio of the old and new methods, respectively. This is a departure from the usually method of finding a confidence limit on the difference $\Delta = P_2 - P_1$. The relative advantages and disadvantages are discussed.

95. Olkin, I., "Multivariate Ratio Estimation for Finite Populations," Biometrika, Vol. 45, 1958.

This paper is concerned with the extension of ratio estimation to the case where multi-auxillary variables are used to increase precision. The following model is presented:

$$\begin{array}{llll} \text{Population} & Y_1, \ldots, Y_n; & \bar{Y} \text{ unknown} & \\ & X_{11}, \ldots, X_{1n}; & \bar{X}_1 \neq 0 \text{ known}, & R_1 = \dfrac{\bar{Y}}{\bar{X}_1} \\ & \vdots & \vdots & \\ & X_{p1}, \ldots, X_{pn}; & \bar{X}_p \neq 0 \text{ known}, & R_p = \dfrac{\bar{Y}}{\bar{X}_p} \end{array}$$

and the (p+1) x (p+1) covariance matrix is known.

The proposed estimator of $\bar{Y}$ is $\tilde{y} = w_1 r_1 \bar{X}_1 + \ldots + w_p r_p \bar{X}_p$ where $w = (w_1, \ldots, w_p)$, $\sum w_i = 1$ is a weighting function and $r_i = \frac{\bar{y}}{\bar{x}_i}$.

As is the result from the univariate case, $\tilde{y}$ is shown to be biased in general but consistent. The Hartly-Ross estimator is generalized to yield an unbiased estimator of $\bar{Y}$.

A large sample approximation is given for the mean, variance and mean square error (to $\Theta(n^{-2})$) and the optimum choice of the weighting function (so as to minimize the variance of $\tilde{y}$) is discussed.

96. O'Neill, A. See Barton [9] .

97. Pearson, K., "On the Constants of Index Distributions," Biometrika, Vol. 7, 1910.

The frequency distribution of $R = X/Y$ is investigated by calculating the distribution of $R = ZX$ when $Z = 1/Y$. The known properties of X and Y distributions are used to develop expressions for the moments of $R = X/Y$. Pearson applies this method to obtain a distribution of an Opsonic index. The coefficients of variation, means, and standard deviations of the raw data are known.

98. Pearson, E. S. See David, H. A. [26] .

99. Pillai, K. C. S., and Ramachandran, K. V., "On the Distribution of the Ratio of the ith Observation in an Ordered Sample from a Normal Population to an Independent Estimate of the Standard Deviation," Annals of Math. Stat., Vol. 25, 1954.

The distribution of an observation, x_i, in an ordered sample of size n from a normal population with zero mean and unit standard deviation is developed as a series of Gamma functions. This distribution, in turn, is utilized to find the distribution of $q_i = (x_i/s)$

where s is an independent estimate of the standard deviation with v degree of freedom. In a similar manner, the distribution of the "studentized" maximum modulus $u_n = |x_n/s|$ is obtained. Uses of these statistics are investigated.

100. Pothoff, R. F. See Roy [113] .

101. Des Raj, "On Mill's Ratio for Type III Populations," Annals of Math. Stat., Vol. 24, 1953.

Several authors have studied the Mills' ratio, i.e., the ratio of the area of the standardized normal curve from x to ∞ and the ordinate at x. The objective of this note is to establish the monotonic character of, and to obtain lower and upper bounds for, the ratio of the ordinate of the standardized Type III curve at x and the area of the curve from x to ∞.

102. Ramachandran, K. V. See Pillai, [99] .

103. Reiter, S., "Estimates of Bounded Relative Error for the Ratio of Variances of Normal Distributions," Journal of the American Stat. Assoc., Vol. 51, 1956.

A method of comparing the variability of two competing industrial processes is investigated. In this paper, the process whose variability is in question behaves like a normally distributed random variable with σ_1^2 representing the variance of the first process and σ_2^2 that of the second. The ratio σ_1^2/σ_2^2 is investigated not through the normal F distribution, but through methods of obtaining a bounded relative error, a minimax estimate as suggested by Girshick in Theory of Games and Statistical Decision, Wiley and Sons, 1954.

104. Rider, P. R., "The Distribution of the Quotient of Ranges in Samples From a Rectangular Distribution," Journal of the American Stat. Assoc., Vol. 22, 1951.

The distribution of the quotient of the ranges of two independent random samples from a continuous rectangular distribution is investigated. The distribution is shown to be independent of the population range. Uses of the distribution are discussed. The special case of equal sample sizes is investigated and four brief tables of probabilistic values are presented.

105. Rider, P. R., "The Distribution of the Product of Ranges in Samples From a Rectangular Population," Journal of the American Stat. Assoc., Vol. 48, 1953.

A general formula is given for the product of the ranges of two independent samples from a rectangular population. The formula does not apply to cases where the sample sizes are the same or if they differ by unity. Special consideration has been given to these two cases.

106. Rider, P. R., "The Distribution of the Product of Maximum Values in Samples from a Rectangular Distribution," Journal of the American Stat. Assoc., Vol. 50, 1955.

This note is to supplement the article by Murty [90] by deriving the distribution of the product of maximum values in random samples from a rectangular p.d.f. The case of the product of K maximum values of samples of equal sizes is also considered.

107. Rider, P. R., "Generalized Cauchy Distributions," Annals of Inst. Stat. Math., Vol. 9, 1958.

This paper compares the asymptotic variance of medians in samples of size n from any distribution (which is continuous and possesses a continuous first derivative in the vicinity of the median) with exact

variances from distributions of the following type:

$$f(x) = C\left(1 + \left[X-\Theta\right] k\right)^{-h}$$

where: $C = k \dfrac{\Gamma(h)}{2\Gamma\left(\frac{1}{K}\right)\Gamma\left(\frac{h-1}{k}\right)}$.

It is seen that for appropriate values of k and h, f(x) becomes the Cauchy and Student-Fisher distributions.

108. Rider, P. R., "Variance of the Median of Samples from A Cauchy Distribution," Journal of the American Stat. Assoc., Vol.55, 1960.

The exact values of the variances of the medians of small samples from a Cauchy distribution are given. The tabular values have been computed from the integral expression representing the frequency function of the median of a sample of $\left[2K + 1\right]$ observations. Values of k considered are integer values up to 15.

109. Rider, P. R., "Expected Values and Standard Deviations of the Reciprocal of a Variable from a Decapitated Negative Binomial Distribution," Journal of the American Stat. Assoc., Vol. 57, 1962.

Tables are given of values for the expected value of $1/x$, $1/x^2$ and the standard deviation of $1/x$ where x follows the negative binomial distribution whose zero class is deleted. A recurrence relation is established using the form

$$E_m^{(k)} = \sum_{x=1}^{\infty} x^{-k} f(x) .$$

A closed form of $E_m^{(1)}$ (the expected value of $1/x$) is obtained and results are computed numerically.

110\. Rietz, H. L., "On Certain Properties of Frequency Distributions Obtained by a Linear Fractional Transformation on the Variates of a Given Distribution," Annals of Math. Stat., Vol. 2, 1931.

Certain properties are examined of the distribution of the variates $u_i = (e\ x_i + f)/(g\ x_i + h)$ obtained by a linear fractional transformation of the x's where e, f, g, and h are real numbers so selected that

$$u = (e\ x + f)/(g\ x + h)$$

is continuous $-1 \leq x \leq 1$.

This investigation results in observing the properties of the differential equation $\frac{du}{dx} = \frac{he - fg}{(gx+h)^2}$.

111\. Rietz, H. L. "On the Frequency Distribution of Certain Ratios," Annals of Math. Stat., Vol. 7, 1936.

A general resumé of the published articles of the early investigations of ratios is presented. In addition, a geometrical description of the distribution of the ratio $t = y/x$ for several cases in which x, y are taken to be uniformly distributed over certain simple geometric shapes is presented. The author considers four cases, the simplest being that in which x and y are uniformly distributed over the offset-rectangular plane.

112\. Rietz, H. L., "On the Distribution of the Student Ratio for Small Samples From Certain Non-Normal Populations," Annals of Math. Stat., Vol. 10, 1939.

The author investigates the behavior of student's ratio when observations are taken from certain non-normal distributions. Measures of efficiency are given as the effects of departure from normality are described through regression techniques applied to the original data.

113. Ross, A. See Hartley [53] .

114. Roy, S. N., and Pothoff, R. F., "Confidence Bounds on Vector Analogues of the Ratio of Means and the Ratio of Variances for Two Correlated Normal Variates and Some Associated Tests," Annals of Math. Stat., Vol. 29, 1958.

In this paper confidence bounds are obtained on the ratio of variances of a (possibly) correlated bivariate normal population, and then by generalization, on a set of parametric functions of a correlated p + p variate normal population and on the ratio of means of these two populations.

115. Scheffé, H., "On the Ratio of the Variances of Two Normal Populations," Ann. of Math. Stat., Vol. 13, 1942.

This paper presents a concise and rigorous presentation of the methods of testing whether two variances may be considered equal when estimated from samples from a normal population. The presentation details the significant tests and confidence interval methods based on the F distribution and on Neyman's various criteria.

116. Shellard, G. D., "On Estimating the Product of Several Random Variables," Journal of the American Stat. Assoc., Vol. 47, 1952.

Through the applications of the central limit theorem and logarithmic transformations, the author finds a suitable approximation to the following problem: Let x_i, $i = 1, \ldots, n$, be random variables with mutually independent distributions and let $X = \pi x_i$. What is the probability that X lies between A and B, i.e., $Pr\{A \leq X \leq B\}$? An investigation of the error introduced in a few simple cases is investigated.

117. Steck, G. P., "A Uniqueness Property Not Enjoyed by the Normal Distribution," Annals of Math. Stat., Vol. 29, 1958.

Counter-examples are constructed showing that a ratio following a Cauchy distribution does not necessarily have normally distributed components. The author uses Fourier transforms of specially defined functions to construct his counter-examples.

118. Steinberg, L. See Krishnaiah [69], [70].

119. Stephan, F. F., "The Expected Value and Variance of the Reciprocal and Other Negative Powers of a Positive Bernoulli Variate," Annals of Math. Stat., Vol. 16, 1945.

Moments of the positive Bernoulli distribution,

$$p(x) = \binom{n}{x} p^x q^{n-x} / (1-q^n)$$

where x, n are integers, $1 \leq x \leq n$,

are found through the application of factorial series. Other distributions such as the positive Hypergeometric and Poisson are considered.

120. "Student", "The Probable Error of the Mean," Biometrika, Vol. 6, 19[illegible]

The classic paper is divided into nine sections, the most noteworthy are those showing:

I. The derivation of the frequency distribution of the standard deviations of samples drawn from a normal population, and

II. that the mean and standard deviation of a sample are independent.

121. Szász, O. See Lukacs [81].

122. Thompson, W. A., "The Ratio of Variances in a Variance Components Model," Annals of Math. Stat., Vol. 26, 1955.

This discussion concerns λ, the ratio of two variances which arise in "mixed" incomplete block designs. A class of invariant statistics for a test involving this ratio is developed as well as the joint distributions of this statistic. The test is used for the hypothesis $\lambda < \lambda_o$ vs. $\lambda > \lambda_1$.

123. Thompson, W. R., "On a Criterion for the Rejection of Observations and the Distributionof the Ratio of Deviation to Sample Standard Deviation, " Annals of Math. Stat., Vol. 6, 1935.

The distribution of $r = \frac{\delta}{s}$ where s is the sample standard deviation and δ is the deviation of an arbitrary observation of the sample from the sample mean is developed. This distribution is discussed in relation to its use in the criterion for rejecting certain elements from a sample.

124. Thompson, G. W., "Bounds for the Ratio of Range to Standard Deviation," Biometrika, Vol. 42, 1955.

This work supplements the work of David, Hartley, and Pearson on the distribution of the ratio of the range w to the standard deviation s. Bounds are shown to exist for w/s for all populations with non-zero variance and percentage points are given for samples of three from a normal population. It is also evidenced that the bounds are distribution free.

125. Tippett, L. See Fisher [36].

126. Tukey, J., and Wilks, S. S., "Approximation of the Distribution of the Product of Beta Variables by a Single Beta Variable," Annals of Math. Stat., Vol. 17, 1946.

Certain maximum likelihood ratio test criteria have been shown to be distributed as a product of k-1 independent Beta variables. The purpose of this note is to consider a method of finding a fractional power of this test criterion which is approximately distributed according to the incomplete Beta distribution function

$$dF(u) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(z)} u^{p-1}(1-u)^{q-1} du$$

and to find appropriate p and q.

127. VanUven, M. J., "Adjustment of a Ratio," Annals of Eugenics, Vol.9, 1939.

Given n pairs of observations $x_1, y_1, \ldots, x_n, y_n$ where x_k, y_k are assumed to have the same weight g_k and all observations are mutually independent the different ratios $m_k = \frac{x_k}{y_k}$ are adjusted to be "best" value $\bar{m}$. Taking x, y as rectangular coordinates this means to find the straight line through the origin that fits the points $P_K(x_k, y_k)$ best. In this paper the probability distributions of certain quantities involved in the adjustment are investigated assuming that the true errors involved in the observations are normally distributed.

128. Von Neumann, J., "Distribution of the Ratio of the Mean Square Successive Difference to the Variance," Annals of Math. Stat., Vol. 12, 1941.

Let $x_1, \ldots, x_n$ be variables representing n successive observations in a population which obeys the normal distribution law. Define the mean and standard deviation estimates in the usual way and let the

mean square successive difference δ^2 be

$$\delta^2 = \frac{1}{n-1} \sum_{\mu=1}^{n-1} (x_{\mu+1} - x_{\mu})^2.$$

The distribution of δ^2 and, in particular, its moments are studied here.

129. Von Neumann, J. See Hart [52].

130. Whitney, D. R. See Mann [83].

131. Wilks, S. S., "Certain Generalizations in the Analysis of Variance," Part 4 of "Moments and Distribution of the Ratio of Independent Generalized Variances," Biometrika, Vol. 24, 1932.

The moments and distribution of the ratio of independent generalized variances for samples from a multivariate normal population are determined. The generalized sample variance is defined as the determinant $|a_{ij}|$ where $a_{ij} = a_{ji} = \frac{1}{N} \sum_{j=1}^{N} (x_{i\alpha} - \bar{x}_i)(x_{j\alpha} - \bar{x}_j)$ $(i, j = 1, \ldots, n)$, when a sample of N items from a n-variate normal population is taken. $\bar{X}_i = \frac{1}{N} \sum_{\alpha=1}^{N} X_{i\alpha}$ is the sample mean of the i^{th} variate, $[X_{i\alpha}$ the value of the i^{th} variate X_i for the α^{th} individual$]$. The generalized population variance λ is defined to be $\lambda = |\sigma_i \sigma_j p_{ij}|$ where σ_i is the standard deviation in the population of the i^{th} variate and (p_{ij}) is the matrix of population correlations.

132. Wilks, S. S., See Tukey [125].

133. Williams, J. D., "Moments of the Ratio of the Mean Square Successive Difference to the Mean Square Difference in Samples from a Normal Universe," Annals of Math. Stat., Vol. 12, 1941.

Consider a sample of n items, $x_1, x_2, \ldots, x_n$, from a normal population with zero mean and variance σ^2, the variates arranged in temporal order. The moments of the ratio S^2 to δ^2 are derived when

$$(n-1)\, S^2 = \sum_{j=1}^{n-1} (x_j - x_{j+1})^2$$

and

$$n\, \delta^2 = \sum_{j=1}^{n} (x_j - \bar{x})^2 .$$

134. Wishart, J., and Bartlett, M., "The Distribution of Second Order Moment Statistics in the Normal System," Proc. Camb. Phil. Soc., Vol. 28, 1932.

Let x be a normally distributed variable and the distribution is written as

$$dp = (2\pi k_2)^{-1/2}\, e^{-1/2\, x^2/k_2}\, dx$$

where k_2 is the semi-invariant of order 2. The moment generating function of the distribution of x^2 is developed. From this the generating function of the semi-invariants (k_r) of x^2 is developed.

135. Yuan, P. T., "On the Logarithmic Frequency Distribution and the Semi-Logarithmic Frequency Surface," Annals of Math. Stat., Vol.4, 1933.

The logarithmic frequency function

$$\frac{1}{\sqrt{2\pi c(z-a)}}\, e^{\frac{1}{2c^2}(\log \frac{z-a}{b})^2}$$

is investigated.

Various methods of determining the parameters of this frequency function have been proposed by different authors. This paper utilizes the method of moments and presents tables facilitating the computation of constants by this method.

136. Tables of the Bivariate Normal Distribution Function and Related Functions, National Bureau of Standards, Vol. 50, Applied Math. Series, 1956.

137. Owen, D. B., "Tables for Computing Bivariate Normal Probabilities," Annals of Math. Stat., Vol. 27, 1956.

138. Pearson, K., Tables for Statisticians and Biometricians, Part II, Cambridge University Press, 1931.

139. Cadwell, J. H., "The Bivariate Normal Integral," Biometrika, Vol. 38, 1951.

140. Gupta, S. S., "Probability Integrals of Multivariate Normal and Multivariate t," Annals of Math. Stat., Vol. 34, 1963.

141. Watson, G. N., Theory of Bessel Functions, Cambridge University Press, London, 1922.

142. Luke, Y. L., Integrals of Bessel Functions, McGraw-Hill Book Co., Inc., New York, 1962.

143. Tables of the Exponential Integral, National Bureau of Standards, Vol. 14, Applied Math. Series, 1951.

144. The Probability Integral, British Assoc. for the Advancement of Science, Math. Tables, Vol. VII, 1939.

145. Tables of the Error Function and Its Derivative, National Bureau of Standards, Vol. 41, Applied Math. Series, 1954.

146. Kendall, M. G., The Advanced Theory of Statistics, Vol. 1, Hafner Publishing Comp., New York, 1958.