AN ESTIMATE OF CORRELATION CORRECTED FOR ATTENUATION AND ITS DISTRIBUTION

By Edney Webb Stacy

University of North Carolina

This research was supported by the Office of Naval Research under Contract No. Nonr-855(06) for research in probability and statistics at Chapel Hill. Reproduction in whole or in part for any purpose of the United States Government is permitted.

Institute of Statistics
Mimeograph Series No. 233
August 1959

ACKNOWLEDGEMENTS

It is a pleasure to acknowledge indebtedness to Professor Harold Hotelling for suggesting the research presented here and for illuminating its course with further suggestion and assistance at every stage. The opportunity of working under his distinguished direction has been a privilege.

I wish to express my deep appreciation to Professors R. C. Bose, W. Hoeffding, and S. N. Roy for their advice, criticism, and assistance in preparation of the paper.

The Office of Naval Research supported initial stages of this research; and the Operations Analysis Office of the Air Force has made possible my recent stay at the University. To both of these organizations, I acknowledge my gratitude. Reproduction of this work for any purpose of the United States Government is freely permitted.

TABLE OF CONTENTS

INTRODUCTION*

Measurement of mental characteristics is subject to considerable error as a rule. Test scores are regarded as having two components--a true score plus a variable error. In the fields of psychology and education, statisticians are concerned both with the reliability and, more importantly, with the correlation between true measures of different characteristics. With regard to the latter problem, if perfectly reliable tests were available, correlation between two characteristics could be estimated directly from a sample product-moment correlation. The existence of test errors means that correlation between test scores may not be an adequate indication of correlation between characteristics which the tests purport to measure.

C. Spearman, G. U. Yule, and others noted this phenomenon around the turn of the century. In 1904,

---

*This research, begun in 1948, was supported until its interruption in 1950, by the United States Department of the Navy, Office of Naval Research, under Project NR 042031, Contract No. N 7 onr-284(02). In September 1958, research was renewed in conjunction with the United States Air Force policy regarding administration of those portions of Public Law 57, 84th Congress, which pertain to contract training and as a part of the Air Force Academic Training Program for Operations Analysts, established by order of the Chief of Staff via letter to the Commander-in-Chief, Continental Air Defense Command and dated December 3, 1956. Further assistance of several kinds, including typing, was provided through the Office of Naval Research Project NR 042031, now under Contract No. Nonr-855(06) for research in probability and statistics at Chapel Hill. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Spearman[1] [17] proposed a method for correcting the correlation between test scores to obtain a better estimate of correlation between characteristics. Errors of measurement tend to diminish or "attenuate" the correlation estimates. Hence Spearman's method is called a correction for attenuation. The correction (to be specified later) involves reliability coefficients of the tests. Though the meaning and method of estimating reliability has been the subject of much discussion in the literature, Spearman's concept was that of the "self-correlation" of a test.

Suppose, for example, that an arithmetic test is administered twice to the same group of students and that memory of the test items plays no part in determining the second set of scores. Discrepancies between the two sets of test scores are expected, and they evidence test error in measuring the mental characteristics involved. The sample correlation betweeen the test scores is a measure of test reliability in the Spearman sense. Current methods for estimating reliability are summarized and evaluated by Kuder and Richardson [10] and Loevinger [11].

Statistical textbooks such as [7], [8], and [12] include reliability and correlations corrected for attenuation as topics.

The Spearman correlation corrected for attenuation is $r/\sqrt{r_1 r_2}$, where r denotes the sample correlation

---

[1]Numbers in square brackets refer to bibliography listed at the end.

between two sets of test scores and r_1, r_2 are measures of self correlation for the respective tests. Two anomalies are apparent. The modulus of the statistic may exceed unity and the statistic may have an imaginary value.

A central difficulty in determining reliability is an ambiguity of definition. For example, the split-half method [10] requires the test items to be divided into two groups, each group containing half of the items. A test containing more than 100 items, say, may be "split" in a multitude of ways, each providing an estimate of reliability. This difficulty has received moderate attention in the literature of psychology, e.g. [10] and [12]; but no generally accepted methods have been found to escape it.

In Chapters I through V, we shall consider two random variables ξ and η which may be interpreted as mental traits. The variable ξ, for example, might be the arithmetic ability of college freshmen; and η might be the reading comprehension ability of those students. We shall suppose that each trait may be measured (inaccurately) in more than one way, e.g., by more than one test. If we denote two measures of ξ by X_1 and X_2, for example, we write

$$X_1 = \xi + \varepsilon_1$$
$$X_2 = \xi + \varepsilon_2$$

where ε_1 and ε_2 are random errors which are independent of each other and of ξ. The error variables ε_1 and ε_2 are assumed to have zero means and equal standard deviations. The scores X_1 and X_2 may stem from a variety of situations. However, we shall avoid ambiguity if we always think of X_1 and X_2 as having been produced by the test-and-retest situation previously illustrated. Then the correlation between X_1 and X_2 may be called a reliability coefficient. (Similar remarks pertain to the mental trait which we denote by η.) Extension of these concepts to cases in which more than two measures are available for each trait is straightforward. The correlation between ξ and η is denoted by $\rho_{\xi\eta}$ and is called the correlation corrected for attenuation.

Study of correlation corrected for attenuation involves two sets of variates: one set measuring ξ and one set measuring η. Hotelling [5] has developed the theory of relations between two sets of variates and has shown that, referring to features which depend only upon correlations and to non-singular internal linear transformations within the sets, canonical correlations and functions of them are the only invariants of the two sets. Texts by Anderson [1] and Roy [15] include many subsequent developments in the theory. The present study gives consideration to canonical correlations for the special cases, i.e., models, which are discussed.

The theory of simple correlation is widely known and is included in modern texts in statistics. A historical account of its development is given by Hotelling along with his research in [6].

The primary subject to be discussed herein is the sample correlation between variates whose measurements are subject to error. Thus, in the sense of the preceding paragraphs, the study concerns correlations corrected for attenuation. A statistic, derived from maximum likelihood considerations is proposed, and its exact probability distribution is given.

The new statistic cannot result in an imaginary number, as is possible with Spearman's correction. However, its modulus may exceed unity, just as Spearman's may. This illustrates a class of problems which has been drawn to the author's attention by Professor Hotelling and which apparently has not been treated in the literature of statistics. Suppose, for example, that independent random variables X and Y are observable only in the form X and $X + Y$. Suppose further that information concerning X and $X + Y$ is obtained separately from independent samples. The variances of X and $X + Y$ may be estimated directly from the samples. But how should such information be used to estimate the variance of Y? The variance of $X + Y$ is the sum of the variance of X and the variance of Y. Yet, simple subtraction of variance estimates for $X + Y$ and X (based upon independent samples) may result in a negative value. Obviously, some modification is required if such a result is to become the basis for estimating the variance of Y. Apparently, similar considerations are required in applications of reliability coefficients and correlations corrected for attenuation.

A mathematical model is specified in Chapter I. The present study of reliability and correction for attenuation postulates two sets of variates. Each set contains $p \geq 2$ variates, and each variate includes a random error as one of its two components. The two-component concept, with attendant assumptions, is called the structure of the variate.

Both the case $p = 2$ and its generalization, $p \geq 2$ are treated. The covariance matrix, as determined by the assumed structures, is established for the variates in each case. The concepts are illustrated at the end of Chapter I.

Density functions for the study are specified in Chapter II. It is assumed that the joint distribution of the variates is multivariate normal with zero means and specified positive definite covariance matrix. The symbol ρ_1 (or ρ_2) always refers to correlation between variates in the same set; and ρ to that between variates of different sets. Canonical correlations stemming from the model are discussed for $\rho_1 \neq \rho_2$ and for $\rho_1 = \rho_2$. Under the assumptions of Chapter II, an exact test is provided for the hypothesis $\rho_1 = \rho_2$ versus the hypothesis $\rho_1 \neq \rho_2$.

The covariance matrix for the variates is denoted by $\tilde{\Sigma}$ or, when $\rho_1 = \rho_2$, by Σ. Chapter III concerns the estimation of elements in Σ. It is assumed that the variates have a common variance $\sigma^2 > 0$ and that the correlation between any two variates from distinct sets is ρ. It is shown that, if the specified structures are adopted, $\rho_1 > 0$ and $|\rho/\rho_1| < 1$. These restrictions are waived to facilitate estimation of σ^2, ρ_1, and ρ. The maximum likelihood estimates of ρ_1 and ρ_2, denoted by $\hat{\rho}_1$ and $\hat{\rho}$, may or may not satisfy the inequalities $\hat{\rho}_1 > 0$ and $|\hat{\rho}/\hat{\rho}_1| < 1$, as is discussed at the end of Chapter III. Asymptotic distributions of the estimators are derived, and canonical correlations are considered. Chapter III includes an alternate derivation of maximum likelihood estimates and a discussion of their anomalous properties.

The next chapter, Chapter IV, provides the distribution of the statistic w. It is shown that cumulative probabilities for w can be expressed as linear functions of four types of integrals. These integrals are evaluated and give the cumulative distribution function, $F(w)$, as a terminating series which involves incomplete Beta-functions. $F(w)$ and the previously mentioned asymptotic distributions are functions of ρ_1 and $\rho_{\xi\eta}$. Methods of calculation are illustrated by numerical example.

Chapter V is a summary of assumptions, techniques and results. It includes a discussion of some possible extensions of the results and of some problems which are associated with the present research.

CHAPTER I

MODEL FOR TWO SETS OF p VARIATES

1.1 Basic structure

The introductory remarks suggest the study of variables whose structures are of the form

$$X = \xi + \varepsilon \quad ,$$

where X is a test score, say, ξ is the true component, and ε is a random error which is independent of ξ . We adopt this structure and assume throughout that the expectations of ξ and ε are zero.[1]

1.2 The 2-variate case

Suppose the random variables $X_1, X_2; X_3, X_4$ have structures

$$X_i = \begin{cases} \xi + \varepsilon_i & , \quad i = 1,2 \\ \eta + \varepsilon_i & , \quad i = 3,4 \end{cases} \quad ,$$

where

(a) $\varepsilon_1, \ldots, \varepsilon_4$ constitute an independent set of

---

[1] It is well known [6] that for complete samples of n + 1 individuals from a p-variate normal population, the distributions of functions of the covariances, computed by dividing the sums of the products of deviations by the number of degrees of freedom n , and with arbitrary population means, is the same as if sample means were not thus eliminated but the means were known to be zero. Accordingly we have assumed that the population means of all variables are zero.

variates with means zero and variances $\sigma^2_{\varepsilon_i}$, $i = 1,\ldots,4$, respectively, and

(b) ξ and η are independent of the error variables $\varepsilon_1,\ldots,\varepsilon_4$, have zero means and variances σ^2_ξ , $\sigma^2_\eta > 0$, respectively, and have covariance $\sigma_{\xi\eta}$. Denoting the variance of X_i by σ^2_i $(i = 1,\ldots,4)$ and the covariance of X_i with X_j by σ_{ij} ,we have

$$\sigma^2_i = \begin{cases} \sigma^2_\xi + \sigma^2_{\varepsilon_i} , & i = 1,2 ; \\ \sigma^2_\eta + \sigma^2_{\varepsilon_i} , & i = 3,4 ; \end{cases}$$

$$(1.1) \qquad \sigma_{12} = \sigma^2_\xi > 0$$

$$\sigma_{34} = \sigma^2_\eta > 0$$

$$\sigma_{ij} = \sigma_{\xi\eta} , \quad i = 1,2;\ j = 3,4 .$$

Using a notation similar to that above in connection with the letter ρ to indicate correlations between variates, we note that $\rho_{12}, \rho_{34} > 0$. Further,

$$(1.2) \qquad \rho_{\xi\eta} = \frac{\sigma_{\xi\eta}}{\sigma_\xi\ \sigma_\eta} , \qquad = \frac{\sigma_{ij}}{\sqrt{\sigma_{12}\sigma_{34}}} \quad (i = 1,2;\ j = 3,4) .$$

Now, we impose two conditions which apply to the whole of the remaining discussion, <u>viz</u>., $\sigma_{\varepsilon_1} = \sigma_{\varepsilon_2}$ and $\sigma_{\varepsilon_3} = \sigma_{\varepsilon_4}$. This implies $\sigma_1 = \sigma_2$ and $\sigma_3 = \sigma_4$, so that

$$(1.3) \qquad \rho_{\xi\eta} = \frac{\rho_{ij}}{\sqrt{\rho_{12}\rho_{34}}} \quad (i = 1,2;\ j = 3,4) .$$

Denote the variables $X_1 + X_2$ and $X_3 + X_4$ by Z_1 and Z_2, respectively. Then, using customary notation for correlations between variables, we have

$$\rho_{Z_1Z_2} = \frac{4\sigma_{\xi\eta}}{\sqrt{(4\sigma_\xi^2 + 2\sigma_{\varepsilon_1}^2)(4\sigma_\eta^2 + 2\sigma_{\varepsilon_3}^2)}} ,$$

(1.4)

$$= \frac{2\sigma_{13}}{\sqrt{(\sigma_\xi^2 + \sigma_1^2)(\sigma_\eta^2 + \sigma_3^2)}} ,$$

from which

(1.5) $$\rho_{\xi\eta} = \frac{\rho_{Z_1Z_2}}{\sqrt{\left(\frac{2\rho_{12}}{1+\rho_{12}}\right)\left(\frac{2\rho_{34}}{1+\rho_{34}}\right)}} .$$

1.3 <u>A generalization</u>

Generalization to the case of $p > 2$ variates in each set is straightforward. Let $X_1,\ldots,X_p;X_{p+1},\ldots,X_{2p}$ be random variables with structures

$$X_i = \begin{cases} \xi + \varepsilon_i & , \quad i = 1,\ldots,p \\ \eta + \varepsilon_i & , \quad i = p+1,\ldots,2p \end{cases}$$

and make assumptions which are direct extensions of (a) and (b) of paragraph 1.2. If $Z_1 = \sum_1^p X_i$ and $Z_2 = \sum_{p+1}^{2p} X_i$, then $\sigma_1 = \sigma_i$ $(i = 1,\ldots,p)$ and $\sigma_{p+1} = \sigma_i$ $(i = p+1,\ldots,2p)$ imply

$$\rho_{Z_1Z_2} = \frac{p^2\sigma_{\xi\eta}}{\sqrt{(p^2\sigma_\xi^2 + p\sigma_{\varepsilon_1}^2)(p^2\sigma_\eta^2 + p\sigma_{\varepsilon_{p+1}}^2)}} ,$$

(1.6)

$$= \frac{p^2\sigma_{1,p+1}}{\sqrt{[(p^2-p)\sigma_{12} + p\sigma_1^2][(p^2-p)\sigma_{p+1,p+2} + p\sigma_{p+1}^2]}} ,$$

from which

(1.7) $$\rho_{\xi\eta} = \frac{\rho_{1,p+1}}{\sqrt{\rho_{12}\,\rho_{p+1,p+2}}} ,$$

and

(1.8) $$\rho_{\xi\eta} = \frac{\rho_{Z_1Z_2}}{\sqrt{\left(\frac{p\,\rho_{12}}{1+(p-1)\rho_{12}}\right)\left(\frac{p\,\rho_{p+1,p+2}}{1+(p-1)\rho_{p+1,p+2}}\right)}} .$$

1.4 <u>Covariance matrix</u>

We assume hereafter, in all of the discussions, that $p \geq 2$, $\sigma_i = \sigma > 0$ $(i = 1,\ldots,2p)$, and adopt the notation $\rho_{12} = \rho_1$, $\rho_{p+1,p+2} = \rho_2$, and $\rho_{1,p+1} = \rho$, maintaining all assumptions of paragraph 1.3. The parameters ρ_1 and ρ_2 may be regarded as correlations between like forms of the same test and will be referred to as reliability coefficients. We denote the covariance matrix for the variables $X_1,\ldots,X_{2p}$ by $\tilde{\Sigma}$ and define

(1.9) $$\tilde{\Sigma} = \sigma^2 \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \tilde{\Sigma}_{22} \end{pmatrix} ,$$

where

$$\Sigma_{11} = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_1 \\ \cdots\cdots & & \cdots & \cdots \\ \rho_1 & \rho_1 & \cdots & 1 \end{pmatrix} , \tag{1.10}$$

$$\tilde{\Sigma}_{22} = \begin{pmatrix} 1 & \rho_2 & \cdots & \rho_2 \\ \rho_2 & 1 & \cdots & \rho_2 \\ \cdots\cdots & & \cdots & \cdots \\ \rho_2 & \rho_2 & \cdots & 1 \end{pmatrix} , \tag{1.11}$$

and

$$\Sigma_{12} = \Sigma_{21} = \begin{pmatrix} \rho & \rho & \cdots & \rho \\ \rho & \rho & \cdots & \rho \\ \cdots\cdots & & \cdots & \cdot \\ \rho & \rho & \cdots & \rho \end{pmatrix} , \tag{1.12}$$

specifying that

$$1 + (p-1)\,\frac{\rho_1 + \rho_2}{2} + p\rho > 0,\; 1 + (p-1)\,\frac{\rho_1 + \rho_2}{2} - p\rho > 0 .$$

Now consider an orthogonal transformation of the variables: Let X and Y be random column vectors with elements $X_1,\ldots,X_{2p}$ and $Y_1,\ldots,Y_{2p}$, respectively. Let these vectors be related by $Y = AX$, where A is the orthogonal matrix

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & 0 \\ 0 & A_4 \end{pmatrix} , \tag{1.13}$$

where A_1 and A_2 are the $2 \times p$ submatrices

$$(1.14) \qquad A_1 = \frac{1}{\sqrt{2p}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \end{bmatrix} ,$$

$$(1.15) \qquad A_2 = \frac{1}{\sqrt{2p}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \end{bmatrix} ,$$

A_3 and A_4 are the $(p-1) \times p$ submatrices

$$(1.16) \qquad A_3 = \begin{bmatrix} \frac{p-1}{a_1} & \frac{-1}{a_1} & \frac{-1}{a_1} & \cdots & \frac{-1}{a_1} & \frac{-1}{a_1} \\ 0 & \frac{p-2}{a_2} & \frac{-1}{a_2} & \cdots & \frac{-1}{a_2} & \frac{-1}{a_2} \\ 0 & 0 & \frac{p-3}{a_3} & \cdots & \frac{-1}{a_3} & \frac{-1}{a_3} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{bmatrix} ,$$

$$(1.17) \qquad A_4 = \begin{bmatrix} \frac{-1}{a_1} & \frac{-1}{a_1} & \cdots & \frac{-1}{a_1} & \frac{-1}{a_1} & \frac{p-1}{a_1} \\ \frac{-1}{a_2} & \frac{-1}{a_2} & \cdots & \frac{-1}{a_2} & \frac{p-2}{a_2} & 0 \\ \frac{-1}{a_3} & \frac{-1}{a_3} & \cdots & \frac{p-3}{a_3} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \cdots & 0 & 0 & 0 \end{bmatrix} ,$$

where $a_j = \sqrt{(p-j)(p-j+1)}$, $j = 1,\ldots,p-1$, and where all elements of A not defined by (1.13),...,(1.17) are zero. We denote the covariance matrix for the elements of Y by $\tilde{D}$ and, after noting that the means of $Y_1,\ldots,Y_{2p}$ are zero, state the following theorem:

<u>Theorem 1.1</u> If the elements of X have covariance matrix $\tilde{\Sigma}$, and other assumptions of this paragraph obtain,

$$(1.18)\qquad \tilde{D} = \begin{bmatrix} B & 0 & 0 \\ 0 & \sigma^2(1-\rho_1)I & 0 \\ 0 & 0 & \sigma^2(1-\rho_2)I \end{bmatrix}$$

where the zeros represent zero submatrices with appropriate numbers of rows and columns,

$$(1.19)\qquad B = \sigma^2 \begin{bmatrix} 1+(p-1)\frac{\rho_1+\rho_2}{2}+p\rho & (p-1)\frac{\rho_1-\rho_2}{2} \\ (p-1)\frac{\rho_1-\rho_2}{2} & 1+(p-1)\frac{\rho_1+\rho_2}{2}-p\rho \end{bmatrix}$$

and where I is a $(p-1)\times(p-1)$ identity matrix.

<u>Proof</u>: Since $Y_1 = \frac{1}{\sqrt{2p}}(Z_1 + Z_2)$, the definitions of Z_1 and Z_2, together with the structures of X_i $(i=1,\ldots,2p)$, permit us to write the variance of Y_1 as

$$(1.20)\qquad \frac{1}{\sqrt{2p}}[p(p-1)\{\rho_1\sigma^2+\rho_2\sigma^2\}+2p\sigma^2+2p^2\rho\sigma^2]$$
$$= \sigma^2[1+(p-1)\frac{\rho_1+\rho_2}{2}+p\rho] .$$

Since $Y_2 = \frac{1}{\sqrt{2p}}[Z_1 - Z_2]$ we have the variance of Y_2 as

$$(1.21)\qquad \frac{1}{\sqrt{2p}}[p(p-1)\{\rho_1\sigma^2+\rho_2\sigma^2\}+2p\sigma^2-2p^2\rho\sigma^2]$$
$$= \sigma^2[1+(p-1)\frac{\rho_1+\rho_2}{2}-p\rho] .$$

We have $Y_1Y_2 = \frac{1}{2p}(Z_1^2 - Z_2^2)$, with expectation

$$(1.22)\qquad \frac{1}{2p}[p(p-1)(\rho_1\sigma^2 - \rho_2\sigma^2)] = \sigma^2(p-1)\,\frac{\rho_1 - \rho_2}{2}\ ;$$

which establishes the matrix B. Using customary notation for variances and covariances, we have for $j = 1,\ldots,p-1$

$$(1.23)\qquad \sigma^2_{Y_{2+j}} = \frac{1}{a_j^2}\Big[(p-j)^2\sigma^2 + (p-j)^2\sigma^2 - 2\Big\{(p-j)^2\sigma_{12} - \frac{(p-j)(p-j-1)}{2}\sigma_{12}\Big\}\Big]\ ,$$

$$= \sigma^2(1-\rho_1)\ .$$

Similarly, for $j = p,\ldots,2(p-1)$, $\sigma_{Y_{2+j}} = \sigma^2(1-\rho_2)$, with the notation agreement (here and elsewhere) that $a_p = a_1$, $a_{p+1} = a_2$, etc.

The covariance of Y_{2+j} and Y_1 $(j = 1,\ldots,p-1)$ is

$$(1.24)\qquad \sigma_{Y_{2+j}\,Y_1} = \frac{\sigma^2}{a_j\sqrt{2p}}\big[(p-j)(p-1)\rho_1 + (p-j) + (p-j)p\rho - (p-j)\{(p-1)\rho_1 + 1 + p\rho\}\big]\ ,$$

$$= 0\ .$$

We consider next the covariance of Y_{2+j} and Y_{2+k} for $j,k < p$ and $j \neq k$. We have

$$(1.25)\qquad \sigma_{Y_{2+j}Y_{2+k}} = \frac{1}{a_j a_k}\big[2\{(p-j)(p-k)\sigma_{12} - (p-j-1)\sigma_{12}\}$$

$$- (p-k)\sigma^2 - 2\{(p-k)(p-j)\sigma_{12} + (p-j-1)(p-k)\sigma_{12}\}$$

$$+ (p-k)\sigma^2$$

$$= 0\ .$$

By similar considerations, when $j,k \geq p$ and $j \neq k$, $\sigma_{Y_{2+j}Y_{2+k}} = 0$.

The only remaining case is $j < p$ with $k \geq p$. The algebra is the same as in the previous case except that the signs are reversed, no variances appear, and σ_{12} (or $\sigma_{p+1,p+2}$ as the particular case demands) is replaced by $\sigma_{1,p+1}$. This gives

$$\sigma_{Y_{2+j}Y_{2+k}} = 0 \quad , \quad j < p \; ; \; k \geq p \quad .$$

This establishes Theorem 1.1 and a corollary follows immediately by inspection of the matrix B.

Corollary 1.1 If $\rho_1 = \rho_2$, the elements of Y are independent.

1.5 Illustration

It is appropriate to illustrate Theorem 1.1. We choose an example in which $p = 3$ and $\sigma^2 = 1$. The matrix A becomes

$$A = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ 0 & 0 & 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

and is orthogonal. After transformation by A, the covariance matrix for the new variables is, by Theorem 1.1,

$$\tilde{D} = \begin{vmatrix} 1+(\rho_1+\rho_2)+3\rho & \rho_1-\rho_2 & 0 & 0 & 0 & 0 \\ \rho_1-\rho_2 & 1+(\rho_1+\rho_2)-3\rho & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-\rho_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-\rho_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-\rho_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1-\rho_2 \end{vmatrix} .$$

We could have established the theorem by computing the product $A \tilde{\Sigma} A'$ for the covariance matrix of the Y elements. Hence we have the corollary [1]:

<u>Corollary 1.2</u> $\qquad A \tilde{\Sigma} A' = \tilde{D}$.

We denote the inverses of $\tilde{\Sigma}$ and $\tilde{D}$ by Λ and Ω, respectively. Then, by well known [2] properties of matrices and determinants

$$\Omega = A'^{-1} \Lambda A^{-1} ,$$

and since A is orthogonal,

<u>Theorem 1.2</u> $\qquad |\Omega| = |\Lambda|$.

CHAPTER II

DENSITY FUNCTIONS

2.1 Density considerations

Apart from the introduction of means, variances, and covariances, references to distributions have been avoided in Chapter I.

Let the variables $X_1, \ldots, X_{2p}$ have means zero and positive definite covariance matrix $\tilde{\Sigma}$, the form of which is specified by (1.9). Let the joint density be

$$(2.1) \qquad \frac{|\Lambda|^{\frac{1}{2}}}{(2\pi)^p} e^{-\frac{1}{2}\sum_{i,j=1}^{2p} \lambda_{ij} x_i x_j} dx_1 \ldots dx_{2p} \quad ,$$

where Λ has elements λ_{ij} and is the inverse of the matrix $\tilde{\Sigma}$. That is to say, we assume the variables have a joint (2p)-variate normal distribution, with parameters as specified above. Let X be a column vector with elements $X_1, \ldots, X_{2p}$ and x be a column vector with elements $x_1, x_2, \ldots, x_{2p}$. If $Y = AX$, where A is defined by (1.13), the joint density of the elements of Y is obtained from (2.1) by means of the transformation $y = Ax$. A is orthogonal so that the Jacobian of the transformation is unity. Thus, the elements of Y have joint density

$$(2.2) \qquad \frac{|\Omega|^{\frac{1}{2}}}{(2\pi)^p} e^{-\frac{1}{2}\sum_{i,j=1}^{2p} \omega_{ij} y_i y_j} dy_1 \ldots dy_{2p} \quad ,$$

where Ω is the inverse of $\tilde{D}$, (1.18), and whose elements are denoted by ω_{ij}. Here we have used Theorems 1.1 and 1.2. By a well known theorem in multivariate analysis (page 29 of [1]), the joint density of Y_1 and Y_2 is, therefore,

$$(2.3) \qquad \frac{|B^{-1}|^{\frac{1}{2}}}{2\pi} e^{-\frac{1}{2}\sum_{i,j=1}^{2} c_{ij} y_i y_j} dy_1 \, dy_2 \quad ,$$

where B^{-1} is the inverse of B, (1.19), and has elements c_{ij}.

Let the respective variances of Y_1 and Y_2 be v_1^2 and v_2^2 and the correlation between Y_1 and Y_2 be τ. Theorem 1.1 indicates that

$$(2.4) \qquad v_1^2 = \sigma^2[1 + (p-1)\frac{\rho_1 + \rho_2}{2} + p\rho] \quad ,$$

$$v_2^2 = \sigma^2[1 + (p-1)\frac{\rho_1 + \rho_2}{2} - p\rho] \quad ,$$

and

$$\tau = \frac{p-1}{v_1 v_2}\frac{\rho_1 - \rho_2}{2} \; .$$

Our assumption that $\tilde{\Sigma}$ is positive definite requires the right hand sides of equations (2.4) to be positive.

Let the columns of the $2 \times n$ matrix

$$(2.5) \qquad \begin{bmatrix} Y_{11} & \ldots & Y_{1n} \\ Y_{21} & \ldots & Y_{2n} \end{bmatrix}$$

be independently and identically distributed according to (2.3). The joint density of the variables is

$$(2.6) \qquad \frac{|B^{-1}|^{\frac{n}{2}}}{(2\pi)^n} e^{-\frac{1}{2}\sum_{\alpha=1}^{n}\sum_{i,j=1}^{2} c_{ij}y_{i\alpha}y_{j\alpha}} dy_{11}\ldots dy_{2n} \quad .$$

If $\tau = 0$ and

$$(2.7) \qquad \tilde{r} = \frac{\sum_{\alpha=1}^{n} Y_{1\alpha}Y_{2\alpha}}{\sqrt{\left(\sum_{\alpha=1}^{n} Y_{1\alpha}^2\right)\left(\sum_{\alpha=1}^{n} Y_{2\alpha}^2\right)}} ,$$

we have, from the theory of simple correlation (page 64 of[1]), the distribution of $\frac{\sqrt{n-1}\,\tilde{r}}{\sqrt{1-\tilde{r}^2}}$ as the "Student" t-distribution with $n-1$ degrees of freedom. A test of $\tau = 0$ versus $\tau \neq 0$ is a test of $\rho_1 = \rho_2$ versus $\rho_1 \neq \rho_2$ and may be made by means of the standard "two-tail" t-test with $n-1$ degrees of freedom (page 65, [1]).

Later in the discussion we shall adopt a model which requires $\rho_1 = \rho_2$ so that the above considerations provide a partial test of that model.

2.2 Canonical correlations

Hotelling [5] has shown that canonical correlations are the non-negative roots of the determinant equation which, in terms of the present model, is

$$(2.8) \qquad \begin{vmatrix} \lambda\Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \lambda\tilde{\Sigma}_{22} \end{vmatrix} = 0 \quad ,$$

or, by [1],

$$|\lambda\Sigma_{11} - \Sigma_{12}(\lambda\tilde{\Sigma}_{22})^{-1}\Sigma_{21}| \cdot |\lambda\tilde{\Sigma}_{22}| = 0 \quad .$$

From this

$$|\Sigma_{11}^{-1}| \cdot |\lambda\Sigma_{11} - \frac{1}{\lambda}\, \Sigma_{12}\tilde{\Sigma}_{22}^{-1}\Sigma_{21}| = 0 \quad ,$$

and

$$|\lambda^2 I - \Sigma_{11}^{-1}\, \Sigma_{12}\, \tilde{\Sigma}_{22}^{-1}\, \Sigma_{21}| = 0 \quad ,$$

a result frequently used by Roy [15], and in which I is a $p \times p$ identity matrix.

To obtain roots of (2.9), we first examine the matrix $\Sigma_{11}^{-1}\, \Sigma_{12}\, \tilde{\Sigma}_{22}^{-1}\, \Sigma_{21}$. Let j be a $1 \times p$ matrix with unit elements. Then $j'j$ is a $p \times p$ matrix with unit elements, $jj' = p$, and $j'jj'j = pj'j$. From (1.10), (1.11), and (1.12), we have

$$\Sigma_{11} = (1 - \rho_1)I + \rho_1 j'j \quad ,$$

$$(2.10) \qquad \tilde{\Sigma}_{22} = (1 - \rho_2)I + \rho_2 j'j \quad ,$$

$$\Sigma_{12} = \Sigma_{21} = \rho j'j \quad .$$

Inverses of Σ_{11} and $\tilde{\Sigma}_{22}$ will be obtained with the aid of a result by Roy and Sarhan [16]. Let D be a $(k \times k)$ diagonal matrix with diagonal elements p_i ; let q and r be $1 \times k$ matrices having elements q_i and r_i , respectively; and let λ be a scalar. Paraphrasing the result in paragraph 4 of [16], we have

Lemma 2.1 If

$$C = D + \lambda q' r \quad , \tag{2.11}$$

then

$$C^{-1} = D^{-1} - \frac{\lambda}{1 + \lambda \sum_{1}^{k} \frac{q_i r_i}{p_i}} (\tilde{q}' \tilde{r}) \quad , \tag{2.12}$$

where $\tilde{q}$ and $\tilde{r}$ are row matrices $(1 \times k)$ whose elements are q_i/p_i and r_i/p_i , respectively.

Direct application of the lemma gives

$$\Sigma_{11}^{-1} = \frac{1}{1-\rho_1} I - \frac{\rho_1}{1 + \frac{p\rho_1}{1-\rho_1}} \left(\frac{1}{1-\rho_1}\right)^2 j'j \quad ,$$

$$= \frac{1}{1-\rho_1} [I - \frac{\rho_1}{1+(p-1)\rho_1} j'j] \quad , \tag{2.13}$$

and similarly

$$\Sigma_{22}^{-1} = \frac{1}{1-\rho_2} [I - \frac{\rho_2}{1+(p-1)\rho_2} j'j] \quad . \tag{2.14}$$

These relations together with (2.10) give

$$\Sigma_{11}^{-1} \Sigma_{12} \tilde{\Sigma}_{22}^{-1} \Sigma_{21} \tag{2.15}$$

$$= \frac{\rho^2}{(1-\rho_1)(1-\rho_2)}[(I - \frac{\rho_1}{1+(p-1)\rho_1} j'j)j'j(I - \frac{\rho_2}{1+(p-1)\rho_2} j'j)j'j] \quad ,$$

$$= \frac{\rho^2}{(1-\rho_1)(1-\rho_2)}[(j'j - \frac{p\rho_1}{1+(p-1)\rho_1} j'j)(j'j - \frac{p\rho_2}{1+(p-1)\rho_2} j'j)] \quad ,$$

$$= \frac{\rho^2}{(1-\rho_1)(1-\rho_2)}[(1 - \frac{p\rho_1}{1+(p-1)\rho_1})(1 - \frac{p\rho_2}{1+(p-1)\rho_2}) j'jj'j] \quad ,$$

$$= \frac{p\rho^2}{(1-\rho_1)(1-\rho_2)} \cdot \frac{(1-\rho_1)(1-\rho_2)}{[1+(p-1)\rho_1][1+(p-1)\rho_2]} j'j \quad ,$$

$$= \frac{p\rho^2}{[1+(p-1)\rho_1][1+(p-1)\rho_2]} j'j \quad .$$

Thus the positive roots of (2.8) are simple functions (square roots) of the characteristic roots of the matrix given by the last line of (2.15). We denote the scalar associated with that matrix by φ and write (2.9) as

$$|\lambda^2 I - \varphi j'j| = |\varphi(\frac{\lambda^2}{\varphi} I - j'j)| = 0 \quad .$$

Hence the square of any root of (2.8) is proportional to some characteristic root of the matrix j'j, the proportionality factor being φ . The general result A.1.18 of [15] applies; so that the non-negative roots of $j'j$ are the same as those for $jj' = p$. Hence $\lambda^2 = \varphi p$, and

$$(2.16) \qquad \lambda = \frac{p|\rho|}{\sqrt{[1+(p-1)\rho_1][1+(p-1)\rho_2]}}$$

is the only positive root of (2.8). We have proved:

Theorem 2.1 Provided $-\frac{1}{p-1} < \rho_1, \rho_2 < 1$, there is a unique canonical correlation ζ , say, where

$$(2.17) \qquad \zeta = \frac{p|\rho|}{\sqrt{[1+(p-1)\rho_1][1+(p-1)\rho_2]}} \quad ,$$

and the positive square root is taken.

Also if $\rho_1 \neq \rho_2$ and $\rho = 0$, we have $\zeta = 0$.

2.3 Canonical variates and quasi-canonical correlation

Assumptions of this chapter are compatible with those of paragraph 1.3 provided $\sigma_1 = \sigma_{p+1} = \sigma$ and the new notation $\rho_{12} = \rho_1$, $\rho_{p+1,p+2} = \rho_2$, and $\rho_{1,p+1} = \rho$ is adopted. Now assume that the variables $X_1,\dots,X_{2p}$ satisfy other requirements of this chapter and may be written

$$X_i = \begin{cases} \xi + \varepsilon_i , & i = 1,\dots,p \\ \eta + \varepsilon_i , & i = p+1,\dots,2p , \end{cases}$$

where $\varepsilon_1,\dots,\varepsilon_{2p}$, ξ, and η are defined by paragraph 1.3 with $\sigma_1 = \sigma_{p+1} = \sigma$. Equation (1.3) applies and may be written

$$\rho_{Z_1Z_2} = \frac{p\rho\sigma^2}{\sqrt{[(p-1)\rho_1\sigma^2 + \sigma^2][(p-1)\rho_2\sigma^2 + \sigma^2]}} ,$$

$$= \frac{p\rho}{\sqrt{[1 + (p-1)\rho_1][1 + (p-1)\rho_2]}} ,$$

where $Z_1 = \sum_1^p X_i$ and $Z_2 = \sum_{p+1}^{2p} X_i$. By Theorem 2.1, $|\rho_{Z_1Z_2}| = \zeta$, which is a unique non-vanishing, canonical correlation if $\rho \neq 0$. It is evident that, for the present model, the population canonical variates are independent of σ^2, ρ_1, and ρ_2 ; depending only upon general properties of ρ . We arbitrate possible ambiguities by the following conventions: If ρ is non-negative we take Z_1 and Z_2 as population canonical variates. Otherwise, we take Z_1 and $(-Z_2)$.

We shall call the random variables Z_1 and Z_2 quasi-canonical variates and their correlation (regardless of sign) the quasi-canonical correlation, denoted by ζ^* .

CHAPTER III

MAXIMUM LIKELIHOOD ESTIMATES

3.1 Likelihood function

Let the columns of the matrix

$$(3.1) \qquad \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ x_{2p,1} & x_{2p,2} & \cdots & x_{2p,n} \end{bmatrix}$$

reprèsent elements of a random sample of size n from a population specified by (2.1) in which the population means are known to be zero. Let the covariance matrix $\tilde{\Sigma}$ be positive definite, be defined by (1.9) with the restriction $\rho_1 = \rho_2$, and have elements whose values are unknown. We denote the covariance matrix (with $\rho_1 = \rho_2$) by Σ and have

$$(3.2) \qquad \Sigma = \sigma^2 \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{11} \end{bmatrix} ,$$

where Σ_{11}, Σ_{12}, and Σ_{21} are defined by (1.10) and (1.12).

The likelihood function [1] of the sample is

$$(3.3) \qquad \frac{|\Lambda|^{\frac{n}{2}}}{(2\pi)^{np}} e^{-\frac{1}{2} \sum_{i,j=1}^{2p} \lambda_{ij} s_{ij}}$$

where Λ, with elements λ_{ij}, is the inverse of Σ and

$S_{ij} = \sum_{\alpha=1}^{n} x_{i\alpha} x_{j\alpha}$. In this chapter, maximum likelihood estimates of the parameters σ^2, ρ_1, and ρ are derived.

3.2 A lemma

Use of the symbol Λ here differs slightly from the previous use. We have imposed the condition $\rho_1 = \rho_2$; otherwise, the meaning of Λ is unchanged. We shall use the temporary expedient of denoting the elements of Σ by σ_{ij} .

Lemma 3.1

If in Σ we have $\sigma_{ij} = \sigma_{k\ell}$, $\lambda_{ij} = \lambda_{k\ell}$.

Proof: Consider the matrix

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_1 & \rho & \cdots & \rho \\ \rho_1 & 1 & \cdots & \rho_1 & \rho & \cdots & \rho \\ \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot \\ \rho_1 & \rho_1 & \cdots & 1 & \rho & \cdots & \rho \\ \rho & \rho & \cdots & \rho & 1 & \cdots & \rho_1 \\ \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot \\ \rho & \rho & \cdots & \rho & \rho_1 & \cdots & 1 \end{pmatrix} . \tag{3.4}$$

Cofactors of its elements (σ_{ij}) are proportional to the elements (λ_{ji}) of Λ . The cofactors of σ_{ii} for $i = 1,\ldots,p$ are identical by inspection of (3.4). When $i > p$, an even number of row and column interchanges in the matrix which determines the cofactor produces a matrix which is identical with the matrix determining the cofactor of σ_{ii} for $i \leq p$. Hence, the cofactor of σ_{ii} equals that for σ_{jj} , $i,j = 1,\ldots,2p$.

Consider two elements σ_{ij} and $\sigma_{i,j+1}$ which lie above the main diagonal and which are equal by definition (3.2).

The cofactor of the former is based upon the sub-matrix obtained from Σ by deleting the i^{th} row and j^{th} column. The i^{th} row and $(j+1)^{st}$ columns are deleted in connection with the element $\sigma_{i,j+1}$. It is observed that the corresponding sub-matrices differ only in two rows and that they would be equal if the appropriate two rows in one of them were interchanged. Hence, except for sign the corresponding determinants are equal. The cofactors of σ_{ij} and $\sigma_{i,j+1}$ are equal if both elements lie above the main diagonal and are equal by (3.2).

Reference to the symmetry of Σ completes the proof.

3.3 Estimation

The maximum of the likelihood function and its logarithm occur at the same point in the parameter space. Let us denote the logarithm of (3.3) by L, adhering to the notation σ_{ij} and λ_{ij} for the elements of Σ and its inverse, respectively. We have

$$(3.5) \qquad L = -np \log 2\pi + \frac{n}{2} \log |\Lambda| - \frac{1}{2} \sum_{i,j=1}^{2p} \lambda_{ij} s_{ij} ,$$

and, following customary maximum likelihood technique, put

$$(3.6) \qquad \frac{\partial L}{\partial \sigma^2} = \frac{n}{2} \sum_{i,j=1}^{2p} \frac{\frac{\partial |\Lambda|}{\partial \lambda_{ij}} \frac{\partial \lambda_{ij}}{\partial \sigma^2}}{|\Lambda|} - \frac{1}{2} \sum_{i,j=1}^{2p} s_{ij} \frac{\partial \lambda_{ij}}{\partial \sigma^2} = 0 .$$

The equations $\frac{\partial L}{\partial \rho} = 0$ and $\frac{\partial L}{\partial \rho_1} = 0$ have a similar form. By [1], (Appendix 1, Theorem 7)

(3.7) $$\frac{\partial|\Lambda|}{\partial\lambda_{ii}} = \Lambda_{ii}$$

and

(3.8) $$\frac{\partial|\Lambda|}{\partial\lambda_{ij}} = 2\,\Lambda_{ij} \quad ,$$

where Λ_{ij} is the cofactor of λ_{ij} , $(i,j = 1,\ldots,2p)$. The matrix Σ is symmetric and is the inverse of Λ, so that

(3.9) $$\frac{\Lambda_{ij}}{|\Lambda|} = \sigma_{ji} = \sigma_{ij} \quad .$$

Lemma 3.1 establishes various equalities between elements λ_{ij} of Λ. Using those results, together with (3.7), (3.8), and (3.9), the equation (3.6) may be written

(3.10) $$\frac{\partial\lambda_{11}}{\partial\sigma^2}\sum_{i,j=1}^{2p}\left(\sigma_{ii}-\frac{S_{ii}}{n}\right)+\frac{\partial\lambda_{12}}{\partial\sigma^2}\sum_{i,j\varepsilon R_1}\left(\sigma_{ij}-\frac{S_{ij}}{n}\right)+$$
$$\frac{\partial\lambda_{p,p+1}}{\partial\sigma^2}\sum_{i,j\varepsilon R_2}\left(\sigma_{ij}-\frac{S_{ij}}{n}\right) = 0 \ ,$$

where the notation $\sum_{i,j\varepsilon R_1}$ indicates summation over those values of the subscripts which satisfy either of the restrictions

(3.11) $$\begin{gathered} i,j \leq p \ , \text{ with } \ i \neq j \ , \\ \text{or} \\ i,j > p \ , \text{ with } \ i \neq j \ , \end{gathered}$$

and $\sum_{i,j\varepsilon R_2}$ indicates similar summation for either

$$(3.12)\qquad \begin{array}{c} i \le p \ \text{ and } \ j > p\ , \\ \text{or} \\ i > p \ \text{ and } \ j \le p\ . \end{array}$$

Analogous development for $\frac{\partial L}{\partial \rho_1}$ and $\frac{\partial L}{\partial \rho}$ produce results similar to (3.10). Those equations and (3.10) are satisfied if

$$\sum_{i=1}^{2p}\left(\sigma_{ii} - \frac{S_{ii}}{n}\right) \equiv 2p\sigma^2 - \sum_{i=1}^{2p} \frac{S_{ii}}{n} = 0 \quad ,$$

$$(3.13)\qquad \sum_{i,j\varepsilon R_1}\left(\sigma_{ij} - \frac{S_{ij}}{n}\right) \equiv 2p(p-1)\rho_1\sigma^2 - \sum_{i,j\varepsilon R_1} \frac{S_{ij}}{n} = 0 \ , \text{ and}$$

$$\sum_{i,j\varepsilon R_2}\left(\sigma_{ij} - \frac{S_{ij}}{n}\right) \equiv 2p(p-1)\rho_2\sigma^2 - \sum_{i,j\varepsilon R_2} \frac{S_{ij}}{n} = 0 \quad .$$

Equations (3.13) establish

<u>Theorem 3.1</u> Under our assumptions, the maximum likelihood estimates of σ^2, ρ_1, and ρ are

$$(3.14)\qquad \hat{\sigma}^2 = \frac{\sum_{i=1}^{2p} S_{ii}}{2pn} = \frac{S_1^2 + S_2^2 + \ldots + S_{2p}^2}{2pn} \quad ,$$

$$(3.15)\qquad \hat{\rho}_1 = \frac{2}{p-1}\,\frac{S_{12} + S_{13} + \ldots + S_{p-1,p} + S_{p+1,p+2} + \ldots + S_{2p-1,2p}}{S_1^2 + S_2^2 + \ldots + S_{2p}^2} ,$$

$$(3.16)\qquad \hat{\rho} = \frac{2}{p}\,\frac{S_{1,p+1} + S_{1,p+2} + \ldots + S_{p,2p}}{S_1^2 + S_2^2 + \ldots + S_{2p}^2} \quad ,$$

where $S_i^2 = S_{ii}$ $(i = 1,\ldots,2p)$.

The numerators of the rightmost expressions in (3.14), (3.15), and (3.16) occur frequently in subsequent discussion and will be denoted by S, $2(\overline{S}_1 + \overline{S}_2)$, and $2\overline{S}_3$, respectively, where $\overline{S}_1 = S_{12} + S_{13} + \dots + S_{p-1,p}$ and

$$\overline{S}_2 = S_{p+1,p+2} + S_{p+1,p+3} + \dots + S_{2p-1,2p} \ .$$

In passing it should be noted that the estimators $\hat{\sigma}^2$, $\hat{\rho}_1$, and $\hat{\rho}$ give the maximum of (3.3) without regard to the inequalities $\rho_1 > 0$ and $|\rho/\rho_1| < 1$. Hence, it may happen that $\hat{\rho}_1 < 0$ or $|\hat{\rho}/\hat{\rho}_1| > 1$ or both.

3.4 Estimates using new variates

Regard the α^{th} column of (3.1) as a column vector, $\tilde{x}_\alpha$ $(\alpha = 1,\dots,n)$ and apply the transformation $y_\alpha = A\,\tilde{x}_\alpha$, where A is defined by (1.13). This gives

(3.17)
$$\sqrt{2p}\; y_{1\alpha} = \sum_{i=1}^{2p} x_{i\alpha} \ ,$$
$$\sqrt{2p}\; y_{2\alpha} = \sum_{i=1}^{p} x_{i\alpha} - \sum_{i=p+1}^{2p} x_{i\alpha} \ ,$$

with

(3.18)
$$2p\; y_{1\alpha}^2 = \sum_{i=1}^{2p} x_{i\alpha}^2 + 2(T_{1\alpha} + T_{2\alpha} + T_{3\alpha})$$
$$2p\; y_{2\alpha}^2 = \sum_{i=1}^{2p} x_{i\alpha} + 2(T_{1\alpha} + T_{2\alpha} - T_{3\alpha})$$

where

(3.19
$$T_{1\alpha} = x_{1\alpha}x_{2\alpha} + x_{1\alpha}x_{3\alpha} + \dots + x_{p-1,\alpha}x_{p\alpha} \ ,$$
$$T_{2\alpha} = x_{p+1,\alpha}x_{p+2,\alpha} + \dots + x_{2p-1,\alpha}x_{2p,\alpha} \ ,$$
$$T_{3\alpha} = x_{1\alpha}x_{p+1,\alpha} + x_{1\alpha}x_{p+2,\alpha} + \dots + x_{p\alpha}x_{2p,\alpha} \ .$$

The subscript α will be deleted when no ambiguity arises.

We also have

$$\sqrt{(p-1)p}\; y_3 = (p-1)x_1 - \sum_{i=2}^{p} x_i$$

$$\cdots$$

$$\sqrt{(p-j)(p-j+1)}\; y_{2+j} = (p-j)x_j - \sum_{i=j+1}^{p} x_i$$

$$\sqrt{(p-j)(p-j+1)}\; y_{p+j} = (p-j)x_{2p-j+1} - \sum_{i=p+1}^{2p-j} x_i$$

for $j = 1,\ldots,p-1$.

Let us examine $p!\sum_{i=1}^{p-1} y_{2+i}^2$ and $p!\sum_{i=1}^{p-1} y_{p+i}^2$. We have

$$(p-1)p\; y_3^2 = (p-1)^2x_1^2 + \sum_{i=2}^{p} x_i^2$$

$$- 2[(p-1)\sum_{i=2}^{p} x_1x_i - \sum_{i=2}^{p-1}\sum_{k=i+1}^{p} x_ix_k] \;,$$

(3.21) $$\cdots \qquad \cdots \qquad \cdots$$

$$(p-j)(p-j+1)y_{2+j}^2 = (p-j)x_j^2 + \sum_{i=j+1}^{p} x_i^2$$

$$- 2[(p-j)\sum_{i=j+1}^{p} x_jx_i - \sum_{i=j+1}^{p-1}\sum_{k=i+1}^{p} x_ix_k] \;,$$

for $j = 1,\ldots,p-1$. We multiply the equation involving y_{2+j}^2 by $p(p-1)\ldots(p-j+2)\cdot(p-j-1)!$ $j = 1,\ldots,p-1$. The coefficient of x_j^2 in $p!\sum_{i=1}^{p-1} y_{2+i}^2$ is b_j , say, where

$$
\begin{aligned}
b_j &= (p-2)! + p(p-3)! + p(p-1)(p-4)! + \ldots + \\
&\quad [p(p-1)\ldots(p-j+3)](p-j)! + p(p-1)\ldots(p-j+2) \times \\
&\qquad (p-j-1)!(p-j)^2 ,
\end{aligned} \tag{3.22}
$$

with

$$
\begin{aligned}
b_{j+1} &= b_j - p(p-1)\ldots(p-j+2)(p-j-1)!(p-j)^2 \\
&\quad + p(p-1)\ldots(p-j+2)(p-j-1)! \\
&\quad + p(p-1)\ldots(p-j+1)(p-j-2)!(p-j-1)^2 , \\
&= b_j - p(p-1)\ldots(p-j+2)(p-j-1)![(p-j)^2 - \\
&\qquad \{1 + (p-j+1)(p-j-1)\}] , \\
&= b_j .
\end{aligned} \tag{3.23}
$$

But

$$
\begin{aligned}
b_1 &= (p-2)!(p-1)^2 = (p-1)(p-1)! , \\
b_2 &= (p-2)! + p(p-3)!(p-2)^2 = (p-2)![1 + p(p-2)] , \\
&= (p-1)(p-1)!
\end{aligned} \tag{3.24}
$$

We have completed an induction which shows that the squared terms of $p!\sum_{i=1}^{p-1} y_{2+i}^2$ each have coefficient $(p-1)(p-1)!$

Similarities in the last two equations of (3.20) show that the squared terms of $p!\sum_{i=1}^{p-1} y_{p+i}^2$ have coefficient $(p-1)(p-1)!$.

Consider the coefficient of $x_j x_\ell$ for $\ell > j$ in the sum $p!\sum_{i=1}^{p-1} y_{2+i}^2$. We follow the previous pattern and denote this coefficient by c_j, the subscript ℓ being unimportant so long as $\ell > j$. We have

$$\frac{c_{j+1}}{2} = \frac{c_j}{2} + p(p-1)\ldots(p-j+2)\cdot(p-j-1)!(p-j)$$

(3.25)
$$+ p(p-1)\ldots(p-j+2)\cdot(p-j-1)!$$

$$- p(p-1)\ldots(p-j+1)\cdot(p-j-2)!(p-j-1) ,$$

$$= \frac{c_j}{2} + p(p-1)\ldots(p-j+2)\cdot(p-j-1)![p-j+1-(p-j+1)] ,$$

$$= \frac{c_j}{2} .$$

This, together with

$$\frac{c_1}{2} = -(p-1)(p-2)! = -(p-1)! ,$$

(3.26)
$$\frac{c_2}{2} = (p-2)! - (p-2)p(p-3)! = (p-2)![1 - p] ,$$

$$= -(p-1)!$$

completes an induction which shows the coefficient of each cross-product term in $p!\sum_{i=1}^{p-1} y_{2+i}^2$ to be $[-2(p-1)!]$. Hence

(3.27)
$$p!\sum_{i=1}^{p-1} y_{2+i}^2 = (p-1)![(p-1)\sum_{i=1}^{p} x_i^2 - 2\sum_{i=1}^{p-1}\sum_{k=i+1}^{p} x_i x_k] .$$

The previously mentioned similarities in equation (3.20) give

(3.28)
$$p!\sum_{i=1}^{p-1} y_{p+q}^2 = (p-1)![(p-1)\sum_{i=p+1}^{2p} x_i^2 - 2\sum_{i=p+1}^{2p-1}\sum_{k=i+1}^{2p} x_i x_k].$$

We restore the subscript α and use (3.19) to write

(3.29)
$$\frac{p}{p-1}\sum_{i=3}^{2p} y_{i\alpha}^2 = \sum_{i=1}^{2p} x_{i\alpha}^2 - 2\,\frac{T_{1\alpha} + T_{2\alpha}}{p-1} ,$$

based upon (3.27) and (3.28). Also,

$$(3.30)\qquad \begin{aligned} p(y_{1\alpha}^2 + y_{2\alpha}^2) &= \sum_{i=1}^{2p} x_{i\alpha}^2 + 2(T_{1\alpha} + T_{2\alpha}) \ , \\ p(y_{1\alpha}^2 - y_{2\alpha}^2) &= 2T_{3\alpha} \ , \end{aligned}$$

from (3.18).

The paragraph following (3.16) defined the sums $S, \bar{S}_1, \bar{S}_2, \bar{S}_3$, which refer to summation upon the subscript α. We use the additional notation

$$2t = \sum_{\alpha=1}^{n} y_{1\alpha}^2 \ ,$$

$$2u = \sum_{\alpha=1}^{n} y_{2\alpha}^2 \ ,$$

$$2v = \sum_{\alpha=1}^{n} \sum_{i=3}^{2p} y_{i\alpha}^2 \ .$$

It is further observed that $\sum_{\alpha=1}^{n} \sum_{i=1}^{2p} x_{i\alpha}^2 = S$ and

$$(3.32)\qquad \sum T_{k\alpha} = \bar{S}_k \ , \qquad k = 1,2,3 \ .$$

By summation on both sides of each equation, (3.29) and (3.30) yield

$$(3.33)\qquad \begin{aligned} 2p(t + u) &= S + 2(\bar{S}_1 + \bar{S}_2) \ , \\ 2p(t - u) &= 2\bar{S}_3 \ , \qquad \text{and} \\ \frac{2p}{p-1} v &= S - 2\,\frac{\bar{S}_1 + \bar{S}_2}{p-1} \ . \end{aligned}$$

Then

$$\frac{S}{2} = t + u + v \quad ,$$

(3.34) $$\frac{\bar{S}_1 + \bar{S}_2}{p - 1} = t + u - \frac{v}{p - 1} \quad , \text{ and}$$

$$\frac{S_3}{p} = t - u \quad .$$

Equation (3.34) and Theorem 3.1 establish the following theorem:

Theorem 3.2 The maximum likelihood estimates of σ^2, ρ_1 and ρ may be written:

$$\hat{\sigma}^2 = \frac{t + u + v}{pn} \quad ,$$

(3.35) $$\hat{\rho}_1 = \frac{t + u - \frac{v}{p-1}}{t + u + v} \quad ,$$

$$\hat{\rho} = \frac{t - u}{t + u + v} \quad .$$

Since $\rho_{\xi\eta} = \frac{\rho}{\sqrt{\rho_1 \rho_2}} = \rho/\rho_1$, the ratio $\hat{\rho}/\hat{\rho}_1$ will be denoted by w and proposed as an estimate of $\rho_{\xi\eta}$, the correlation corrected for attenuation. We have

(3.36) $$w = \frac{t - u}{t + u - \frac{v}{p-1}}$$

which reduces to

(3.37) $$w = \frac{t - u}{t + u - v}$$

in the special case $p = 2$.

3.5 Asymptotic Distributions

The statistics $\hat{\sigma}^2$, $\hat{\rho}_1$, and $\hat{\rho}$ are functions of the sample moments and are consistent estimates of σ^2, ρ_1, and ρ [1]. The same is true for w as an estimate of $\rho/\rho_1 = \rho_{\xi\eta}$. Hence, the means in the asymptotic distributions of $\hat{\sigma}^2$, $\hat{\rho}_1$, and $\hat{\rho}$ are σ^2, ρ_1, and ρ, respectively.

The remainder of this chapter is devoted to the variance of $\hat{\sigma}^2$ and the asymptotic distributions of $\hat{\sigma}^2$, $\hat{\rho}_1$, and $\hat{\rho}$. The elements of the sample matrix (3.1) and variables obtained from them by transformation may be regarded as random variables. The marginal distributions of the elements in (3.1) are normal [1] with means zero and variances σ^2. Hence each variable y_{ij} obtained in paragraph (3.4) is a linear combination of normal variates and is normally distributed with mean zero and variance according to the matrix $\tilde{D}$ (1.18), with $\rho_1 = \rho_2$. We define α to be the variance of y_{1j} $(j = 1,\ldots,n)$, β to be the variance of y_{2j} $(j = 1,\ldots,n)$, and γ to be the common value of the variances of y_{3j}, y_{4j}, $\ldots$, $y_{2p,j}$ $(j = 1,\ldots,n)$. Reference to Theorem 1.1 shows that, with $\rho_1 = \rho_2$

$$\alpha = \sigma^2[1 + (p-1)\rho_1 + p\rho] \quad ,$$

$$\beta = \sigma^2[1 + (p-1)\rho_1 - p\rho] \quad , \text{ and} \tag{3.38}$$

$$\gamma = \sigma^2(1 - \rho_1) \quad .$$

Then $\frac{y_{1j}^2}{\alpha}$, $\frac{y_{2j}^2}{\beta}$, $\frac{y_{3j}^2}{\gamma}$, $\frac{y_{4j}^2}{\gamma}$, $\ldots$, $\frac{y_{2p,j}^2}{\gamma}$ are independently

(Corollary 1.1) distributed as χ^2 variables, each with one degree of freedom. The variables 2t, 2u, and 2v as defined by (3.31) are sums of independent sets of independent χ^2 variables, the individual χ^2 variables having one degree of freedom. Hence we have:

Theorem 3.4 The random variables 2t, 2u, 2v are independent as a set and

a. $\frac{t}{\alpha}$ is distributed as $\frac{1}{2}\chi^2$ with n degrees of freedom,

b. $\frac{u}{\beta}$ is distributed as $\frac{1}{2}\chi^2$ with n degrees of freedom,

c. $\frac{v}{\gamma}$ is distributed as $\frac{1}{2}\chi^2$ with 2(p-1)n degrees of freedom.

The frequency functions of t, u, and v are therefore of the form

$$(3.39) \qquad \frac{x^{\nu}\, e^{-\frac{x}{\theta}}}{\Gamma(\nu+1)\ \theta^{\nu+1}} \quad ; \quad x \geq 0 \ , \quad \theta > 0 \ , \quad \nu \geq 0 \ .$$

This "gamma" type distribution [13] has mean and variance

$$\theta(\nu + 1) \ , \text{ and}$$

$$\theta^2(\nu + 1) \ ,$$

respectively. The variables t, u, and v require $\theta = \alpha, \beta, \gamma$ and $\nu = \frac{n}{2}, \frac{n}{2}, \frac{2(p-1)n}{2}$ (the order is t,u,v).

Table 3.1 contains the means and variances of some functions which appear in (3.35) and (3.36) and will be used in computing the variances of $\hat{\sigma}^2$, $\hat{\rho}_1$, and $\hat{\rho}$.

Table 3.1

Means and Variances

(Ancillary)

Function	Mean	Variance
$t - u$	$\frac{n}{2}(\alpha - \beta)$	$\frac{n}{2}(\alpha^2 + \beta^2)$
$t + u + v$	$\frac{n}{2}[\alpha + \beta + 2(p-1)\gamma]$	$\frac{n}{2}[\alpha^2 + \beta^2 + 2(p-1)\gamma^2]$
$t + u + \frac{v}{p-1}$	$\frac{n}{2}[\alpha + \beta - 2\gamma]$	$\frac{n}{2}[\alpha^2 + \beta^2 + \frac{2\gamma^2}{p-1}]$

The following covariances are useful also:

$$\text{cov}(t - u,\ t + u + v) = \frac{n}{2}(\alpha^2 - \beta^2) \quad ,$$

$$\text{cov}(t + u - \frac{v}{p-1},\ t + u + v) = \frac{n}{2}(\alpha^2 + \beta^2 - 2\gamma^2) \quad ,$$

$$\text{cov}(t - u,\ t + u - \frac{v}{p-1}) = \frac{n}{2}(\alpha^2 - \beta^2) \quad .$$

Using Theorem 3.2, the variance of $\hat{\sigma}^2$ is

$$\text{(3.40)} \qquad \text{Var}(\hat{\sigma}^2) = \frac{1}{2p^2n}[\alpha^2 + \beta^2 + 2(p-1)\gamma^2] \quad ,$$

an exact expression. By Theorem 3.2 and equations (3.17), (3.18), and (3.31), it is seen that $\hat{\sigma}^2$ is a linear function of sample second order moments from a multivariate normal distribution. It follows [3] that, for large n, $\hat{\sigma}^2$ is approximately normally distributed about mean σ^2, with variance given by (3.40).

The asymptotic distributions of the estimators $\hat{\rho}_1$, $\hat{\rho}$, and w will be obtained by means of a lemma which is a special case of a general result by Hoeffding and Robbins [4].

Let

(3.41) $(X_1^*, Y_1^*), (X_2^*, Y_2^*), \ldots$

be a sequence of random vectors (real elements) which are independently and identically distributed. For all values of i, let the vector elements X_i^* and Y_i^* have zero means, finite third absolute moments, variances σ_1^{*2} and σ_2^{*2} (respectively), and covariance σ_{12}^*. Finally, let

$$H(x,y) = \frac{x + \theta_1}{y + \theta_2}$$

where θ_1 and θ_2 are constants with $\theta_2 \neq 0$. H(x,y) and its derivatives are continuous at the point (0,0). All conditions of Theorem 4 from the cited research [4] are satisfied. Hence,

<u>Lemma 3.1</u> As $n \to \infty$, the function

$$\sqrt{n}\left\{H\left(\frac{\sum^n X_i^*}{n}, \frac{\sum^n Y_i^*}{n}\right) - H(0,0)\right\}$$

has a limiting normal distribution in which the mean is zero and the variance is

$$H_1\sigma_1^{*2} + H_1H_2\sigma_{12}^* + H_2^2\sigma_2^{*2} \quad ,$$

where H_1 and H_2 are the first order derivatives with respect to x and y, evaluated at the point (0,0).

Each of the estimators $\hat{\rho}_1$, $\hat{\rho}$, and w may be written in the form

$$Q(N,D) = \frac{N}{D} \quad ,$$

$$= \frac{\frac{1}{n}(N - n\nu) + \nu}{\frac{1}{n}(D - n\delta) + \delta} \quad ,$$

where $(n\nu)$ and $(n\delta)$ are the respective means of N and D and each of the functions $(N - n\nu)$ and $(D - n\delta)$ is a sum of n identically and independently distributed random variables whose means are zero and third absolute moments are finite. This may be verified by inspection of equations (3.35), (3.36), (3.17) and (3.18). By appropriate pairing of a random variable from $(N - n\nu)$ with a random variable from $(D - n\delta)$, we can construct a sequence of random vectors with the properties of (3.41).

We have assumed that the covariance matrix Σ is positive definite so that $\alpha,\beta,\gamma > 0$. By Table 3.1, then, $\delta > 0$ in the cases of $\hat{\rho}_1$ and $\hat{\rho}$. The estimator w is used only when $\rho_1 > 0$ and $|\rho/\rho_1| < 1$. These inequalities lead to

$$1 - \rho_1 < 1 + (p-1)\rho_1 - p|\rho| \leq 1 + (p-1)\rho_1 + p|\rho| \quad .$$

Hence, we have $\alpha,\beta > \gamma > 0$ and $\alpha + \beta - 2\gamma > 0$ in cases for which w is appropriate; and $\delta > 0$.

The conditions of Lemma 3.1 are satisfied in all three cases.

The variance of N (and D), in the present context, is n times the variance of one of its n independent components.

In the sequence (3.41) which is established from $(N - n\nu)$ and $(D - n\delta)$, denote the vector elements from the former by X_i^* and those from the latter by Y_i^*. We observe that, if $i \neq j$, the expected value of $(X_i^* Y_j^*)$ is zero. Hence, the covariance of N and D is n times the covariance of X_i^* and Y_i^*.

This completes preparation for applying Lemma 3.1 to the functions Q given by (3.42). We have:

<u>Lemma 3.2</u>

$$\sqrt{n}\ \{Q(N,D) - Q(n\nu,\ n\delta\}$$

has a limiting normal distribution in which the mean is zero and the variance is

$$\frac{1}{n\delta^2}\sigma_N - 2\frac{\nu}{n\delta^3} + \frac{\nu^2}{n\delta^4}\sigma_D^2$$

(3.43)

$$= \frac{1}{\delta^2}\left[\frac{\sigma_N^2}{n} - 2\frac{\nu}{\delta}\left(\frac{\sigma_{ND}}{n}\right) + \left(\frac{\nu}{\delta}\right)^2\frac{\sigma_D^2}{n}\right] .$$

Let us apply this result to $\hat{\rho}_1$ with the aid of equation 3.35, Table 3.1, and the covariance information which follows that table. The variance in the asymptotic distribution of $\sqrt{n}(\hat{\rho}_1 - \rho_1)$ is

$$\frac{2}{\alpha + \beta + 2(p-1)\gamma^2}\left[\alpha^2 + \beta^2 + \frac{2\gamma^2}{p-1} - 2\rho_1(\alpha^2 + \beta^2 - 2\gamma^2) + \rho_1^2(\alpha^2 + \beta^2 + 2(p-1)\gamma^2)\right]$$

(3.44)

$$= \frac{2}{[\alpha + \beta + 2(p-1)\gamma]^2}\left\{(\alpha^2 + \beta^2)(1 - \rho_1)^2 + \frac{2\gamma^2}{p-1}[1 + (p-1)\rho_1]^2\right\}$$

An interesting sidelight is the fact that if $p = 2$ and $\rho_{\xi\eta} = 0$, the variance (3.44) reduces to $\frac{(1 - \rho_1^2)^2}{2}$. It follows that the variance in the large sample distribution of $\hat{\rho}_1$ under this set of special conditions is $\frac{(1 - \rho_1^2)^2}{2n}$. This expression is equivalent to the asymptotic variance of a sample product-moment correlation coefficient for a sample of size 2n from a bivariate normal distribution.

Similarly, the respective variances in the asymptotic distributions of $\sqrt{n}(\hat{\rho} - \rho)$ and $\sqrt{n}(w - \rho_{\xi\eta})$ are

$$(3.45) \quad \frac{2}{[\alpha + \beta + 2(p-1)\gamma]^2} \left\{\alpha^2(1 - \rho)^2 + \beta^2(1 + \rho)^2 + 2(p-1)\gamma^2\rho^2\right\}$$

and

$$(3.46) \quad \frac{2}{[\alpha + \beta - 2\gamma]^2} \left\{\alpha^2(1 - \rho_{\xi\eta})^2 + \beta^2(1 + \rho_{\xi\eta})^2 + \frac{2\gamma^2}{p-1} \rho_{\xi\eta}^2\right\} .$$

The preceding discussion is summarized by

<u>Theorem 3.5</u> If $\rho_1 > 0$ and $|\rho/\rho_1| < 1$, each of the variables

$$\sqrt{n}(\hat{\rho}_1 - \rho_1) ,$$

$$\sqrt{n}(\hat{\rho} - \rho) , \text{ and}$$

$$\sqrt{n}(w - \rho_{\xi\eta})$$

has a limiting normal distribution as $n \to \infty$; the means in these respective distributions are zero; and the corresponding variances are given by (3.44), (3.45) and (3.46).

It follows that the asymptotic distributions for $\hat{\rho}_1$, $\hat{\rho}$, and w are normal, have means ρ_1, ρ, and $\rho_{\xi\eta}$, respectively, the variances being easily obtained from (3.44), (3.45) and (3.46).

Theorems (3.2) and (3.4) together with (3.36) and the definitions of α, β, and γ indicate that $\hat{\rho}_1$, $\hat{\rho}$, and w may be expressed directly as functions of $(\frac{1}{2}\chi^2)$-variables and that each is independent of σ^2. Hence when distributions for $\hat{\rho}_1$, $\hat{\rho}$, and w are discussed, there is no loss of generality in assuming that $\sigma^2 = 1$.

3.6 Distribution of the sample quasi-canonical correlation

Developments in the present chapter permit us to comment upon paragraph (2.3) and, earlier, Theorem 2.1. Expression (2.17) gave the canonical correlation as a function of the parameters ρ_1, ρ_2, and $|\rho|$. Paragraph (2.3) gives the quasi-canonical variates Z_1 and Z_2 when $\rho > 0$. It was assumed in those discussions that the basic variables $X_1, \ldots, X_{2p}$ had the particular structure given in paragraph (2.3). We make the corresponding assumption now for the elements of (3.1).

Let ρ be positive and the columns of (3.1) be transformed by $z_{1j} = \sum_{i=1}^{p} x_{ij}$, $z_{2j} = \sum_{i=p+1}^{2p} x_{ij}$, and $z_{ij} = x_{ij}$ $(i = 3, \ldots, 2p)$ for $j = 1, \ldots, n$. The joint distribution of the variables z_{ij} $(i = 1, \ldots, 2p;\ j = 1, \ldots, n)$ is multivariate normal with means zero and covariance matrix Ψ , say. Then by [1] (page 24), z_{1j} and z_{2j} have a joint bivariate normal distribution in which the means are zero, the variances can be specified, and the correlation coefficient is the quasi-canonical correlation (2.18). We denote this canonical correlation by ζ^*. The parameter ζ^* is estimated by z , say, where

$$(3.48) \qquad z = \frac{\sum_{j=1}^{n} z_{1j}\, z_{2j}}{\sqrt{\left(\sum_{j=1}^{n} z_{1j}^2\right)\left(\sum_{j=1}^{n} z_{2j}^2\right)}}$$

is distributed like a standard sample product moment correlation coefficient, the distribution having parameters ζ^* and n.

In addition to the above assumptions, let $\rho_1 = \rho_2$. Then the variances of z_1 and z_2 are equal. It is easy to see that [15] $(z_1 - \zeta^* z_2)$ and z_2 are uncorrelated and have a joint bivariate normal distribution with means zero and positive variances. Let r^* be the sample product moment correlation between $(z_1 - \zeta^* z_2)$ and z_2. Then, following the development of Roy [15, page 93], we note that

$$\frac{r^*\sqrt{n-1}}{\sqrt{1-r^{*2}}} = t \text{ , say,}$$

has a "Student" t-distribution with $n-1$ degrees of freedom and that we can set confidence intervals about ζ^*. Let $\tilde{s}_1^2$ and $\tilde{s}_1^2$ be sample variances for z_1 and z_2 respectively. Then with confidence $1-\alpha$

$$\frac{\tilde{s}_1}{\tilde{s}}\left[z - \frac{t_\theta\sqrt{1-z^2}}{\sqrt{n-1}}\right] \leq \zeta^* \leq \frac{\tilde{s}_1}{\tilde{s}_2}\left[z + \frac{t_\theta\sqrt{1-z^2}}{\sqrt{n-1}}\right] ,$$

where t_α is the value of t (in the t-distribution) for which

$$P(t > t_\alpha) = \frac{\theta}{2} ,$$

when $(n-1)$ degrees of freedom are used.

3.7 Alternate derivation of estimates

It was seen in Theorem 3.4 that the statistic w may be expressed as a function of independent $\frac{1}{2}\chi^2$ variables with specified parameters. These variables are t, u, and v and we denote their joint frequency by

$$g = \frac{(tu)^{\frac{n}{2}-1} v^{\frac{m}{2}-1} e^{-(\frac{t}{\alpha}+\frac{u}{\beta}+\frac{v}{\gamma})}}{[\Gamma(\frac{n}{2})]^2\Gamma(\frac{m}{2})(\alpha\beta)^{\frac{n}{2}}\gamma^{\frac{m}{2}}} , \tag{3.52}$$

with $t, u, v \geq 0$, $\alpha, \beta, \gamma > 0$, and $m = 2(p-1)n$. Consider the space of the parameters α, β, and γ. The function g is continuous with respect to the parameters and its derivatives with respect to the parameters are continuous. Further, as one or more of the parameters tends to 0 or tends to ∞ , $g = 0$. Since g is non-negative it therefore must have a smooth maximum at which its derivatives with respect to α, β, and γ are zero. In fact we shall see that g has a unique maximum in the octant $\alpha,\beta,\gamma > 0$.

Let $\log(g) = M$ so that

$$M = -\log[\Gamma(\tfrac{n}{2})]^2\,\Gamma(\tfrac{m}{2}) - \frac{n}{2}\log\alpha - \frac{n}{2}\log\beta$$

$$-\frac{m}{2}\log\gamma - (\frac{t}{\alpha}+\frac{u}{\beta}+\frac{v}{\gamma}) + (\frac{n}{2}-1)\log tu + (\frac{m}{2}-1)\log v,$$

and

$$\frac{\partial M}{\partial\alpha} = -\frac{n}{2}\frac{1}{\alpha} + \frac{t}{\alpha^2} ,$$

(3.53)
$$\frac{\partial M}{\partial\beta} = -\frac{n}{2}\frac{1}{\beta} + \frac{u}{\beta^2} ,$$

$$\frac{\partial M}{\partial\gamma} = -\frac{m}{2}\frac{1}{\gamma} + \frac{v}{\gamma^2} .$$

Since g and $\log g$ have maxima at the same point in the parameter space, we maximize g by setting equations (3.5) equal to zero and solving for α, β, and γ. The unique solution is

$$\hat{\alpha} = \frac{2t}{n} ,$$

(3.54)
$$\hat{\beta} = \frac{2u}{n} ,$$

$$\hat{\gamma} = \frac{2v}{m} .$$

Let us put

$$\frac{2t}{n} = \sigma^2[1 + (p-1)\rho_1 + p\rho] ,$$

(3.55)
$$\frac{2u}{n} = \sigma^2[1 + (p-1)\rho_1 - p\rho] ,$$

$$\frac{2v}{m} = \sigma^2[1 - \rho_1] ,$$

and determine values of σ^2, ρ_1, and ρ which satisfy these equations. Direct substitution from equations (3.35) verifies that $\sigma^2 = \hat{\sigma}^2$, $\rho_1 = \hat{\rho}_1$, and $\rho = \hat{\rho}$ provide a solution.

3.8 Anomalies of $\hat{\rho}_1$ and w

Before proceeding to the distribution of w, we shall examine the function g in the ρ_1,ρ plane. It has been pointed out that g vanishes as α, β, and γ tend to zero either jointly or independently. This fact and the definitions of α, β, and γ indicate that g must vanish along the lines

$$\rho = -\frac{1 + (p-1)\rho_1}{p}$$

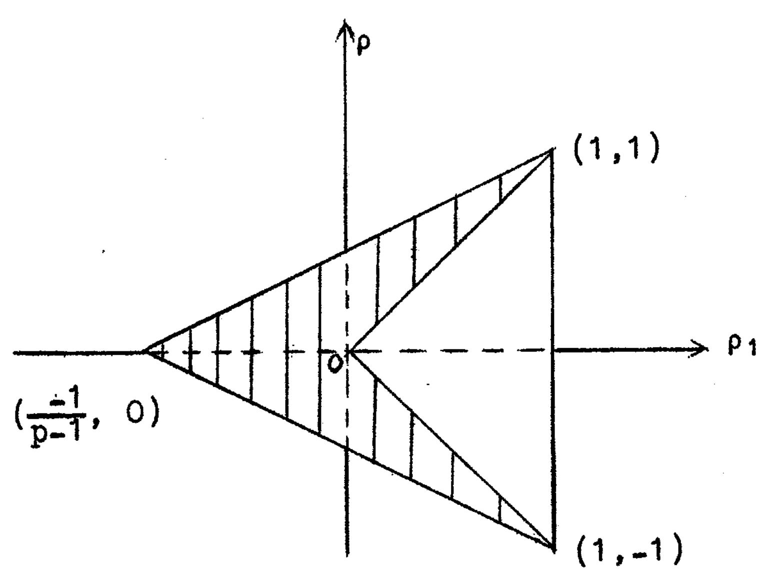
$$\rho = \frac{1 + (p-1)\rho_1}{p} \qquad (3.56)$$

$$\rho_1 = 1$$

in the ρ_1,ρ plane. These lines form a triangle illustrated by the largest triangle shown in Figure 3.1.

Figure 3.1

Region in which $g > 0$



Outside and on the boundaries of the largest triangle in the figure, $g = 0$. Within that triangle $\alpha, \beta, \gamma > 0$ and $g > 0$. The smaller, unshaded, isosceles triangle indicates the region in which $|\rho_{\xi\eta}| \equiv |\rho/\rho_1| \leq 1$, the equality holding on the non-vertical boundaries. Inside and on the boundaries of this smaller triangle $\alpha,\beta \geq \gamma$.

It is evident that $|\rho/\rho_1| \leq 1$ does not imply $|\hat{\rho}/\hat{\rho}_1| \leq 1$ necessarily. Either or both of the following inequalities may be satisfied:

a. $\hat{\rho}_1 < 0$

b. $|\hat{\rho}/\hat{\rho}_1| > 1$.

The structure which has been assumed prevents $\rho_1 < 0$, so that $\hat{\rho}_1 < 0$ cannot be used directly to estimate $\rho_1 > 0$. Likewise, w cannot be used directly to estimate $\rho_{\xi\eta}$ when we have $|\hat{\rho}/\hat{\rho}_1| \equiv |w| > 1$ or when $\hat{\rho}_1 < 0$.

It may be possible to derive estimates which do not possess the present anomalies--perhaps, by combining the method of maximum likelihood with the use of some method to introduce the conditions $\rho_1 > 0$ and $|\rho/\rho_1| \leq 1$ (or alternately $\alpha,\beta \geq \gamma$).

The practicing statistician may be able to alleviate the trouble by taking advantage of another property shown by Figure 3.1. The left vertex of the largest triangle is at the point $(\frac{-1}{p-1}, 0)$, so that increasing values of p diminish the size of the shaded area. However, the precise effects of increasing the value of p have not been investigated.

Further guidance in these matters is provided by the consistency property of each of the present estimates. As the sample size, n, increases, $\hat{\rho}_1$ and w converge in probability to ρ_1 and $\rho_{\xi\eta}$, respectively.

CHAPTER IV

DISTRIBUTION OF THE STATISTIC w

4.1 General

Much of Chapter III was devoted to expressing the statistic w in terms of independent variables t, u, and v. The development culminated in (3.36), giving

$$w = \frac{t - u}{t + u - \frac{v}{p-1}} . \tag{4.1}$$

We shall use capital letters to denote random variables which correspond to t, u, v, and w so that the smaller letters may be used in integration processes. For example, we let the cumulative probability function of W be $F(w) = P(W \leq w)$.

The frequency functions of T, U, and V are given by (3.39) and their joint density is

$$\frac{(\alpha\beta)^{-\frac{n}{2}} \gamma^{-\frac{m}{2}}}{[\Gamma(\frac{n}{2})]^2 \Gamma(\frac{m}{2})} (tu)^{\frac{n}{2}-1} v^{\frac{m}{2}-1} e^{-(\frac{t}{\alpha}+\frac{u}{\beta}+\frac{v}{\gamma})} dtdudv , \tag{4.2}$$

for $t,u,v \geq 0$; $\alpha,\beta,\gamma > 0$; and $m = 2(p-1)n$.

Let $t = t'$, $u = u'$, and $v = (p-1)v'$, with Jacobian $(p-1)$. Then (4.2) becomes

$$(4.3) \qquad K(t'u')^{\frac{n}{2}-1}(v')^{\frac{m}{2}-1} e^{-(\frac{t'}{\alpha}+\frac{u'}{\beta}+\frac{v'}{\gamma'})} dt'du'dv' \quad ,$$

where $\gamma' = \gamma/(p-1)$ and $K = \{\alpha\beta\gamma'[\Gamma(\frac{n}{2})]^2\Gamma(\frac{m}{2})\}^{-1}$. We have

$$P(W \leq w) = P\left(\frac{T - U}{T + U - V'} \leq w\right)$$

where $V' = V/(p-1)$. The ratio $(T - U)/(T + U - V')$ is positive or negative according to whether numerator and denominator have like or unlike signs. The ratio is real valued, having range $\pm\infty$.

We shall find it convenient to omit the prime, ('), associated with the variables and γ'. The notation will be renewed when clarity of expression requires it. Section 4.4 of this chapter is the first in which the primes reoccur.

Now let us put $w > 0$. Then $(T - U)/(T + U - V) \leq w$ if the ratio is negative or if either of the following sets of inequalities is satisfied

$$(4.5) \qquad \begin{aligned} T &\geq U \quad , \\ T + U &> V \quad , \\ (w-1)T + (w+1)U &\geq Vw \quad ; \end{aligned}$$

or

$$(4.6) \qquad \begin{aligned} T &< U \quad , \\ T + U &< V \quad , \\ (w-1)T + (w+1)U &< wV \quad . \end{aligned}$$

4.2 Geometry

The above sets of restrictions define boundaries of regions in the t, u, v space over which (4.3) must be integrated to derive the probability (4.4). The above discussion, taken with late developments in Chapter III, reduce the multiplicity of the integration to 3, so that three dimensional geometric figures can be used to summarize the problem and to guide the integration processes.

Decreasing or increasing the value of w rotates the plane

$$(4.7) \qquad (w-1)t + (w+1)u - wv = 0$$

about the common intersection--clockwise for decreasing w and counterclockwise for increasing w. At $w = 0$, the plane coincides with $t - u = 0$; as $w \to \infty$ it approaches coincidence with $t + u - v = 0$. When $0 < w < 1$, the intersection of $v = 0$ and (4.7) is the line $u = t(1-w)/(1+w)$, $v = 0$.

We have, for $V < T + U$,

$$(4.8) \qquad P(W \geq -|w|) = P[|w|V \leq (|w|-1)U + (|w|+1)T] \ .$$

This could have been obtained from (4.6) by interchanging the roles of T and U and substituting $|w|$ for w. The relation (4.8), the illustration of Figure 4.1, and the density (4.3) each suggests that integration results for $w > 0$ can be modified to give results for the case $w < 0$. Specifically, if we write $F(w) = F(\alpha,\beta,\gamma;\ w)$ and have $w < 0$, $F(w) = 1 - F(\beta,\alpha,\gamma;\ |w|)$.

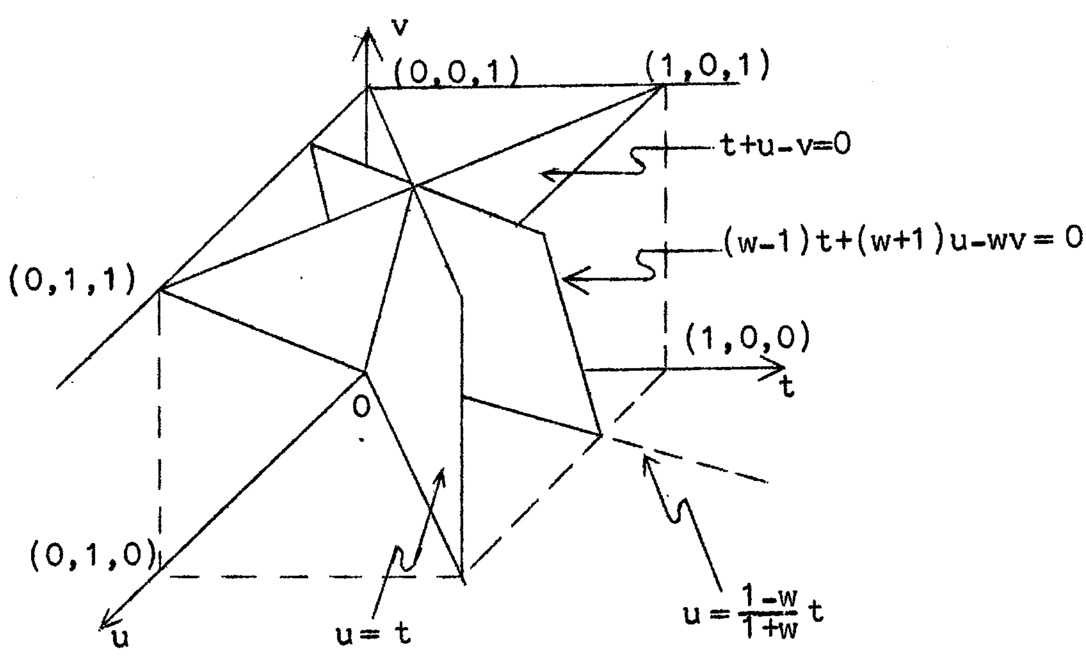
Only the positive octant of the t,u,v space concerns us. Boundary planes such as those defined by (4.5) and (4.6) extend indefinitely in the positive directions. To depict the situation, we truncate the planes

$$\begin{aligned} t &= u \\ v &= t + u \qquad (4.9) \\ wV &= (w-1)t + (w+1)u \end{aligned}$$

in Figure 4.1, showing only the unit cube.

Figure 4.1

Boundaries for Integration Regions



The origin and the point (1,1,2) satisfy equations (4.9). Hence, the corresponding planes of (4.9) intersect in a line.

Suppose $0 < w < 1$. The probability that W does not exceed w is given by

$$
(4.10)\quad
\begin{aligned}
P(W \le w) &= \int_0^{\infty}\int_{at}^{t}\int_0^{bt+cu} f\,dt\,du\,dv \\
&+ \int_0^{\infty}\int_t^{\infty}\int_{bt+cu}^{\infty} f\,dt\,du\,dv \\
&+ \int_0^{\infty}\int_t^{\infty}\int_0^{t+u} f\,dt\,du\,dv \\
&+ \int_0^{\infty}\int_0^{t}\int_{t+u}^{\infty} f\,dt\,du\,dv \quad ,
\end{aligned}
$$

where f is the frequency function determined from (4.3) by deleting the differential elements and recalling that the primes (') are omitted temporarily. Here, $a = \frac{1 - w}{1 + w}$, $b = \frac{w - 1}{w}$, and $c = \frac{w + 1}{w}$.

We shall use the first integral on the right hand side of (4.10) to illustrate other geometric concepts. We have

$$
(4.11)\quad \int_0^{\infty}\int_{at}^{t}\int_0^{bt+cu} f\,dt\,du\,dv \ ,
$$

and transform by

$$
(4.12)\quad
\begin{aligned}
t &= \alpha z x \quad , \\
u &= \beta z y \quad , \\
v &= \gamma z(1 - x - y) \quad ,
\end{aligned}
$$

which has Jacobian $\alpha\beta\gamma z^2$. The integrand becomes

$$(4.13)\qquad \frac{(xy)^{\frac{n}{2}-1}(1-x-y)^{\frac{m}{2}-1}}{[\Gamma(\frac{n}{2})]^2\Gamma(\frac{m}{2})}\, z^{n+\frac{m}{2}-1}\, e^{-z} \quad .$$

The transformation was chosen to yield $(t/\alpha)+(u/\beta)+(v/\gamma)=z$. Before the transformation, the region of integration was defined by

$$(4.14)\qquad \begin{aligned} &t,u,v \geq 0 \quad , \\ &at \leq u < t \quad , \\ &0 \leq v < bt + cu \quad . \end{aligned}$$

Substituting from (4.12), these inequalities define the new region as

$$(4.15)\qquad \begin{aligned} &x,y,z \geq 0 \quad , \\ &y < 1 - x \quad , \\ &\frac{ax}{\beta} \leq y < \frac{\alpha}{\beta}x \quad , \\ &y \geq \frac{\gamma}{\gamma + c\beta} - \frac{\gamma + b\alpha}{\gamma + c\beta}x \quad . \end{aligned}$$

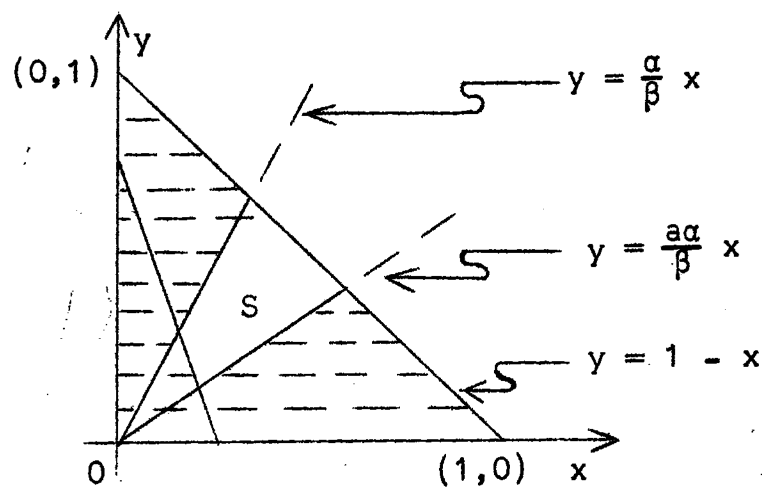
The variable z is restricted only by $z \geq 0$. Hence, we eliminate z by integration over the entire positive range. The remaining integration becomes

$$(4.16)\qquad \frac{\Gamma(n+\frac{m}{2})}{[\Gamma(\frac{n}{2})]^2\Gamma(\frac{m}{2})} \int\int_S (xy)^{\frac{n}{2}-1}(1-x-y)^{\frac{m}{2}-1}\,dx\,dy \quad ,$$

the region S being illustrated in Figure 4.2 by the unshaded area of the largest triangle shown.

Figure 4.2

Region S in the xy-plane



We note that (4.16) is proportional to

$$(4.17) \qquad \int_S\int \int_0^{1-x-y} (xy)^{\frac{n}{2}-1} z^{\frac{m}{2}-2} \, dx\, dy\, dz \quad .$$

If the region S contained all of the area between the co-ordinate axes and the line $y = 1 - x$, the integral (4.17) could be evaluated as a Dirichlet type integral. Hence we regard (4.17) as a truncated Dirichlet integral in the same sense that an incomplete Beta-function may be regarded as a 2-dimensional truncated Dirichlet integral. There is little hope for a tidy, easily computed solution to (4.16), but the integral can be expressed as a function of incomplete Beta-functions--as will be shown in paragraph 4.3.

If the variables T, U, and V are transformed in the same manners as (4.12),

$$(4.18)\qquad \frac{T - U}{T + U - V} = \frac{\alpha x - \beta y}{(\alpha+\gamma)x + (\beta+\gamma)y - \gamma},$$

and (4.13) indicates that x and y (as random variables) are not independent. However, if we transform by

$$r = \frac{x}{1 - y},$$

with Jacobian (1 - S), the integrand, (4.13), becomes proportional to

$$(4.19)\qquad r^{\frac{n}{2}-1}(1 - r)^{n-1} \cdot S^{\frac{n}{2}-1}(1 - S)^{n+\frac{m}{2}-1}.$$

Hence, random variables corresponding to r and s are independent. The new region of integration is bounded by non-linear curves and will not be used in further developments. Also

$$(4.20)\quad \frac{\alpha x - \beta y}{(\alpha+\gamma)x + (\beta+\gamma)y - \gamma} = \frac{\alpha r(1 - S) - \beta S}{(\alpha+\gamma)r(1-S) + (\beta+\gamma)S - \gamma}$$

is a result of the transformation, but will receive no further study in this paper.

4.3 Lemmas for the distribution of w

Reference is made to Figure 4.1, observations which follow equation (4.9), and to the example provided by equation (4.10). Cumulative probabilities for $W = (T - U)/(T + U - V)$ can be expressed as functions of four types of integrals, viz.,

$$(4.21)\quad \int_0^\infty\int_0^\infty\int_0^\infty,\ \int_0^\infty\int_0^{at}\int_0^\infty,\ \int_0^\infty\int_0^\infty\int_0^{bt+cu},\ \int_0^\infty\int_0^{at}\int_0^{bt+cu}.$$

The first of these is unity since the integrand is a frequency function. The remaining three will be evaluated in sequence and will be denoted by

$$A_{n,m}(\alpha,\beta,\gamma;\ a)$$

(4.22) $$C_{n,m}(\alpha,\beta,\gamma;\ b,c)\ , \quad \text{and}$$

$$D_{n,m}(\alpha,\beta,\gamma;\ a,b,c)\ .$$

The symbols (4.22) will be abbreviated A, C, and D when no ambiguity results; and the incomplete Beta-function,

(4.23) $$\int_0^{\varphi} x^{\lambda-1}(1-x)^{\mu-1}\,dx \quad ,$$

will be denoted by $B(\lambda, \mu)$. The density (4.3) is repeated (without the primes) here for future reference:

(4.24) $$K(tu)^{\frac{n}{2}-1}\ v^{\frac{m}{2}-1}\ e^{-(\frac{t}{\alpha}+\frac{u}{\beta}+\frac{v}{\gamma})}dt\,du\,dv \quad ,$$

where $m = 2(p-1)n$ and $K = \left\{(\alpha\beta)^{\frac{n}{2}}\ \gamma^{\frac{m}{2}}[\Gamma(\frac{n}{2})]^2\,\Gamma(\frac{m}{2})\right\}^{-1}$, and we shall write (4.22) as $f\ dt\,du\,dv$ when convenient.

4.3.1 <u>The integral A</u> We have

(4.25)
$$A = \int_0^{\infty}\int_0^{at}\int_0^{\infty} f\ dt\,du\,dv \quad ,$$
$$= K_A\int_0^{\infty}\int_0^{at}(tu)^{\frac{n}{2}-1}\ e^{-(\frac{t}{\alpha}+\frac{u}{\beta})}dt\,du \quad ,$$

where $K_A = \left\{(\alpha\beta)^{\frac{n}{2}}\ [\Gamma(\frac{n}{2})]^2\right\}^{-1}$. Now let

$$(4.26)\qquad \begin{aligned} t &= \alpha\, s(1 - r) \quad , \\ u &= \beta\, s\, r \quad , \end{aligned}$$

the Jacobian being $\alpha\beta s$. This gives

$$A = \frac{1}{[\Gamma(\frac{n}{2})]^2} \int_0^{\infty} \int_0^{\frac{a\alpha}{a\alpha+\beta}} r^{\frac{n}{2}-1} (1-r)^{\frac{n}{2}-1} s^{n-1} e^{-s} dr\, ds ,$$

$$= \frac{\Gamma(n)}{[\Gamma(\frac{n}{2})]^2} \int_0^{\frac{a\alpha}{a\alpha+\beta}} r^{\frac{n}{2}-1} (1-r)^{\frac{n}{2}-1} dr \quad .$$

Hence,

<u>Lemma 4.1</u>

$$(4.27)\qquad A = \frac{\Gamma(n)}{[\Gamma(\frac{n}{2})]^2} B_{\frac{a\alpha}{a\alpha+\beta}} (\tfrac{n}{2}, \tfrac{n}{2}) \quad .$$

4.3.2 <u>The Integral C</u>. Expressions (4.12), (4.13) and (4.16) indicate that we may write

$$*4.28)\qquad C = K_C \int_{S'}\int (xy)^{\frac{n}{2}-1} (1-x-y)^{\frac{m}{2}-1} dx\, dy ,$$

where $K_C = \Gamma(n+\frac{m}{2}) \Big/ [\Gamma(\frac{n}{2})]^2 \Gamma(\frac{m}{2})$ and S' will be defined.

The integration limits for C give

$$\begin{aligned} &t,u,v \geq 0 \quad , \\ &0 \leq v < bt + cu , \end{aligned}$$

and substitution from (4.12) leads to

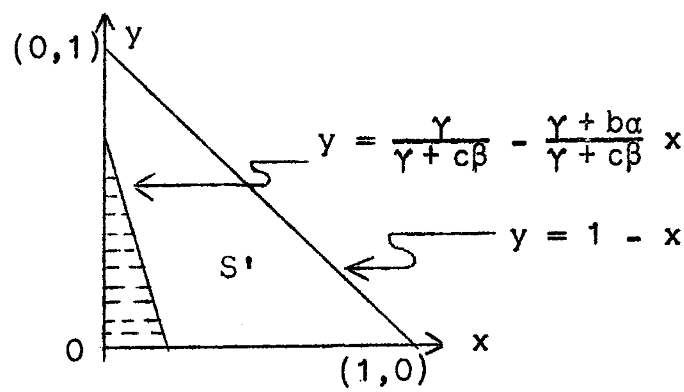
$$x, y \geq 0 \quad ,$$

$$1 - x > y \geq \frac{\gamma}{\gamma + c\beta} - \frac{\gamma + b\alpha}{\gamma + c\beta} x \quad ,$$

as illustrated by Figure 4.3.

Figure 4.3

Region S' in the x,y-plane



It is recalled that $m = 2(p - 1)n$ so that $\frac{m}{2}$ must be an integer. This fact permits the expansion of $(1 - x - y)^{\frac{m}{2} - 1}$ in the terminating series

$$(4.29) \quad \sum_{i=0}^{\frac{m}{2}-1} \binom{\frac{m}{2} - 1}{i} (x + y)^i (-1)^i$$

$$= \sum_{i=0}^{\frac{m}{2}-1} \sum_{j=0}^{i} (-1)^i \binom{\frac{m}{2} - 1}{i} \binom{i}{j} x^{i-j} y^j \quad .$$

Multiplying by $K_C(x\ y)^{\frac{n}{2} - 1}$ and integrating term by term, the integrals involved are of the form

(4.30) $$\int_{S'}\int x^{\ell-1}\, y^{q-1}\, dx\, dy \quad ,$$

where

(4.31) $$\ell = \frac{n}{2} + i - j \ , \quad \text{and}$$
$$q = \frac{n}{2} + j \quad .$$

Figure 4.3 indicates that (4.30) is the difference between a Beta-function and an integral which, except for the matter of scale, is a Beta-function. With reference to the second integral, we shall have

(4.32) $$x,y \geq 0$$
$$y \leq \frac{\gamma}{\gamma + c\beta} - \frac{\gamma + b\alpha}{\gamma + c\beta}\, x \quad .$$

We make the change of scale by letting

$$r = \frac{\gamma + b\alpha}{\gamma}\, x \quad ,$$

$$s = \frac{\gamma + c\beta}{\gamma}\, y \quad ,$$

with Jacobian $[\gamma/(\gamma + b\alpha)][\gamma/(\gamma + c\beta)]$. Use of that transformation in (4.30) and combining the result with (4.28) and (4.29) gives

<u>Lemma 4.2</u>

(4.33) $$C = \frac{\Gamma(n+\frac{m}{2})}{[\Gamma(\frac{n}{2})]^2\Gamma(\frac{m}{2})} \sum_{i=0}^{\frac{m}{2}-1} \sum_{j=0}^{i} \left\{ (-1)^i \binom{\frac{m}{2}-1}{i} \binom{i}{j} \times \frac{\Gamma(\ell)\,\Gamma(q)}{\Gamma(\ell+q+1)} \left[1 - \left(\frac{\gamma}{\gamma + b\alpha}\right)^{\ell} \left(\frac{\gamma}{\gamma + c\beta}\right)^{q} \right] \right\} .$$

4.3.3 The Integral D. Evaluation of D follows the same development. We begin with

$$(4.34) \qquad D = K_C \int_{S''}\!\!\int (xy)^{\frac{n}{2}-1} (1 - x - y)^{\frac{m}{2}-1} \, dx \, dy \ ,$$

where K_C is defined following (4.28) and the region S" is obtained from

$$(4.35) \qquad \begin{aligned} t,u,v &\geq 0 \ , \\ 0 \leq u &< at \ , \\ 0 \leq v &< bt + cu \ , \end{aligned}$$

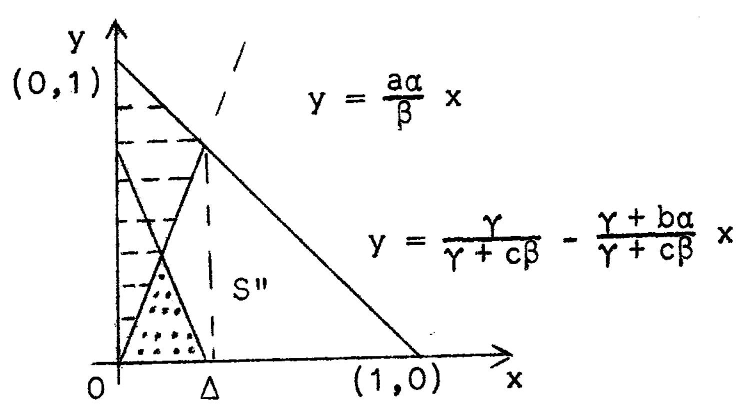
which lead to

$$(4.36) \qquad \begin{aligned} x,y &\geq 0 \ , \\ 0 \leq y &< \frac{a\alpha}{\beta} x \ , \\ \frac{\gamma}{\gamma + c\beta} - \frac{\gamma + b\alpha}{\gamma + c\beta} x \leq y &< 1 - x \ . \end{aligned}$$

The region defined by (4.36) is illustrated by the unshaded region of Figure 4.4.

Figure 4.4

Region S" in the x,y-plane



The region T, shaded with dots, and the point $(\Delta, 0)$ are shown because of their usefulness in subsequent development.

In fact, integration over S" may be accomplished by subtracting the result of integrating over T from the result of integrating over S" + T .

We have, after duplicating the step indicated by (4.29) and (4.30)

$$\int\int_{S''+T} x^{\ell-1} y^{q-1}\, dx\, dy$$

$$= \int_0^{\Delta}\int_0^{\frac{\alpha a}{\beta}x} x^{\ell-1}y^{q-1}dx\, dy + \int_{\Delta}^{1}\int_0^{1-x} x^{\ell-1}y^{q-1}dx\, dy\ ,$$

(4.37)
$$= \left(\frac{a\alpha}{\beta}\right)^q \frac{1}{q}\int_0^{\Delta} x^{\ell+q-1}dx + \int_0^{\Delta}\int_0^{x} (1-x)^{\ell-1}y^{q-1}dx\, dy\ ,$$

$$= \left(\frac{a\alpha}{\beta}\right)^q \frac{\Delta^{\ell+q}}{q(\ell+q)} + \frac{1}{q} B_{\Delta}(\ell,\ q+1)\quad ,$$

where Δ is the x-coordinate of the intersection of the lines $y = \frac{a\alpha}{\beta}x$ and $y = 1 - x$. Specifically,

(4.38)
$$\Delta = \frac{\beta}{\beta + a\alpha}\quad .$$

After a change of scale for each variable, integration over the region T is similar to that for the region S" + T . Hence, we let

(4.39)
$$r = \frac{\mu}{\lambda}x\quad ,$$
$$s = \frac{y}{\lambda}\quad ,$$

with Jacobian $\frac{\lambda^2}{\mu}$, where

$$\lambda = \frac{\gamma}{\gamma + c\beta} \,,$$

(4.40)

$$\mu = \frac{\gamma + b\alpha}{\gamma + c\beta} \,.$$

The line $y = \frac{a\alpha}{\beta} x$ transforms into $s = \frac{a\alpha}{\beta} \lambda r$ and the r coordinate of its intersection with the line $s = 1 - r$ is $r = \Delta_1 = \frac{\beta}{\beta + a\alpha\lambda}$, the last relation defining Δ_1.

Then

$$\int\int_T x^{\ell-1} y^{q-1} dx\, dy$$

(4.41)

$$= \left(\frac{\lambda}{\mu}\right)^{\ell} \lambda^q \left\{ \int_0^{\Delta_1} \int_0^{\frac{a\alpha\lambda}{\beta} r} r^{\ell-1} s^{q-1} dr\, ds + \int_0^{\Delta_1} \int_0^{r} (1 - r)^{\ell-1} s^{q-1} dr\, ds \right\} ,$$

$$= \left(\frac{\lambda}{\mu}\right)^{\ell} \lambda^q \left\{ \left(\frac{a\alpha}{\beta} \lambda\right)^q \frac{\Delta_1^{\ell+q}}{q(\ell+q)} + \frac{1}{q} B_{\Delta_1}(\ell, q+1) \right\} .$$

The results (4.37) and (4.41), combined with steps paralleling those which lead to (4.29), give

Lemma 4.3

(4.42)

$$D = \frac{\Gamma(n + \frac{m}{2})}{[\Gamma(\frac{n}{2})]^2\, \Gamma(\frac{m}{2})} \sum_{i=0}^{\frac{m}{2}-1} \sum_{j=0}^{i} \left\{ (-1)^i \binom{\frac{m}{2} - 1}{i} \binom{i}{j} \left[\left(\frac{a\alpha}{\beta}\right)^q \frac{\Delta^{\ell+q}}{q(\ell+q)} + \frac{1}{q} B_{\Delta}(\ell, q+1) - \left(\frac{\lambda}{\mu}\right)^{\ell} \lambda^q \left\{ \left(\frac{a\alpha}{\beta} \lambda\right)^q \frac{\Delta_1^{\ell+q}}{q(\ell+q)} + \frac{1}{q} B_{\Delta_1}(\ell, q+1) \right\} \right] \right\} ,$$

where the symbols have been defined previously.

4.4 Distribution of w

It was shown in paragraph (4.2) that the cumulative distribution of W can be expressed as a function of A, C, and D. To use these integrals it is necessary to restore the prime notation, ('). We shall associate the symbol w with the random variable W and denote the latter's cumulative distribution function by $F(w)$.

Consider the case $0 < w < 1$. The probability $P(W \leq w)$ was discussed in connection with (4.10) which gave

(4.43)
$$
\begin{aligned}
P(W \leq w) &= \int_0^\infty \int_{at'}^{t'} \int_0^{bt'+cu'} f\, dt'du'dv' \\
&+ \int_0^\infty \int_{t'}^\infty \int_{bt'+cu'}^\infty f\, dt'du'dv' \\
&+ \int_0^\infty \int_{t'}^\infty \int_0^{t'+u'} f\, dt'du'dv' \\
&+ \int_0^\infty \int_0^{t'} \int_{t'+u'}^\infty f\, dt'du'dv' \quad .
\end{aligned}
$$

Each of these integrals will be examined in turn.

We have

(4.44)
$$
\begin{aligned}
\int_0^\infty \int_{at'}^{t'} \int_0^{bt'+cu'} &= \int_0^\infty \int_0^{t'} \int_0^{bt'+cu'} - \int_0^\infty \int_0^{at'} \int_0^{bt'+cu} \\
&= D_{n,m}(\alpha,\beta,\gamma';\ 1,b,c) \\
&\quad - D_{n,m}(\alpha,\beta,\gamma';\ a,b,c) \quad ,
\end{aligned}
$$

where

$$a = \frac{1 - \frac{p}{2(p-1)}\, w}{1 + \frac{p}{2(p-1)}\, w} ,$$

$$b = \frac{\frac{p}{2(p-1)}\, w - 1}{\frac{p}{2(p-1)}\, w} ,$$

(4.45)

$$c = \frac{\frac{p}{2(p-1)}\, w + 1}{\frac{p}{2(p-1)}\, w} ,$$

$$\alpha = [1 - (p-1)\rho_1 + p\rho] ,$$

$$\beta = [1 + (p-1)\rho_1 - p\rho] ,$$

$$\gamma' = \frac{1}{p-1} (1 - \rho_1) ,$$

and $\rho_{\xi\eta} = \frac{\rho}{\rho_1}$.

The second integral of (4.43) is

$$\int_0^\infty \int_{t'}^\infty \int_{bt'+cu'}^\infty = \int_0^\infty \int_0^\infty \int_0^\infty - \int_0^\infty \int_0^t \int_0^\infty - \int_0^\infty \int_0^\infty \int_0^{bt'+cu'} + \int_0^\infty \int_0^{t'} \int_0^{bt'+cu'}$$

(4.46)

$$\begin{aligned} &= 1 - A_{n,m}(\alpha,\beta,\gamma';\ 1) \\ &\quad - C_{n,m}(\alpha,\beta,\gamma';\ b,c) \\ &\quad + D_{n,m}(\alpha,\beta,\gamma';\ 1,b,c) \end{aligned}$$

The next integral is

$$(4.47)\quad \int_0^\infty \int_{t'}^\infty \int_0^{t'+u'} = \int_0^\infty \int_0^\infty \int_0^{t'+u'} - \int_0^\infty \int_0^{t'} \int_0^{t'+u'}$$

$$= C_{n,m}(\alpha,\beta,\gamma';\ 1,1) - D_{n,m}(\alpha,\beta,\gamma';\ 1,1,1)\ .$$

Finally,

$$(4.48)\quad \int_0^\infty \int_0^{t'} \int_{t'+u'}^\infty = \int_0^\infty \int_0^{t'} \int_0^\infty - \int_0^\infty \int_0^{t'} \int_0^{t'+u'}$$

$$= A_{n,m}(\alpha,\beta,\gamma';\ 1) - D_{n,m}(\alpha,\beta,\gamma';\ 1,1,1)\ .$$

We combine this result with (4.44), (4.46) and (4.47) to obtain

<u>Theorem 4.1</u>. For $0 < w < 1$,

$$(4.49)\quad \begin{aligned} F(w) = 1 &+ C_{n,m}(\alpha,\beta,\gamma';\ 1,1) + 2D_{n,m}(\alpha,\beta,\gamma';1,b,c) \\ &- C_{n,m}(\alpha,\beta,\gamma';\ b,c) - 2D_{n,m}(\alpha,\beta,\gamma';\ 1,1,1) \\ &- D_{n,m}(\alpha,\beta,\gamma';\ a,b,c)\ . \end{aligned}$$

If we denote the right hand side of (4.49) by $F(\alpha,\beta,\gamma';\ w)$, $F(-|w|) = 1 - F(\beta,\alpha,\gamma',|w|)$. [The definitions of a, b, and c must be interpreted with $|w|$ in place of w when $w < 0$.]

The case for $w \geq 1$ gives $F(w)$ in the same form as (4.43) with the following exception. The first integral on the right hand side of (4.43) must be replaced by

$$(4.50)\quad \int_0^\infty \int_0^{t'} \int_0^{bt'+cu'} f\ dt'du'dv' = D_{n,m}(\alpha,\beta,\gamma';\ 1,b,c)\ .$$

This case gives

$$F(w) = 1 + 2D_{n,m}(\alpha,\beta,-'; 1,b,c)$$
$$+ C_{n,m}(\alpha,\beta,\gamma'; 1,1) - C_{n,m}(\alpha,\beta,\gamma'; b,c)$$
$$- 2D_{n,m}(\alpha,\beta,\gamma ; 1,1,1) \quad .$$

The value of $F(-w)$, $w \geq 1$, is obtained by appropriate changes in the roles of α and β , as in the previous case.

4.5 Illustration

Computation by means of the functions A, C, and D will be illustrated next. Let us consider a numerical example in which $n = 4$, $p = 2$, and reliability coefficient is $\rho_1 = .9$. We take $\sigma = 1$ since this introduces no loss of generality. The only remaining parameter is the correlation corrected for attenuation which, for this example, we take to be $\rho_{\xi\eta} = .5$.

These assumptions establish

$$\alpha = 1 + (p-1)\rho_1 + p\rho_1\rho_{\xi\eta} = 2.8$$

(4.52) $$\beta = 1 + (p-1)\rho_1 - p\rho_1\rho_{\xi\eta} = 1$$

$$\gamma = 1 - \rho_1 = 0.1 \quad .$$

These values may be substituted into equation (4.27) to find

(4.53) $$A = 6\ B_{\left(\frac{2.8a}{1+2.8a}\right)}(2,2) \quad .$$

Evaluation of the incomplete Beta-function appearing on the right hand side is accomplished by reference to [14].

The functions C and D will be used to compute $P(W \geq 1)$. Interest in this problem stems from the consideration of W (or, alternately, the statistic w) as an estimate of the correlation coefficient $\rho_{\xi\eta}$. Of course, $|\rho_{\xi\eta}| \leq 1$ and the probability that its "estimate" shall exceed unity should be considered.

An upper bound for $P(W > 1)$ can be obtained in the general case through the use of C alone. We have

$$\begin{aligned} P(W > 1) &= P(T + U - V' > 0 \text{ and } 2U - V' < 0) \\ &\quad + P(T + U - V' < 0 \text{ and } 2U - V' > 0) \\ &\leq P(2U < V') + P(T + U < V') \quad . \end{aligned}$$

This may be written

$$P(W > 1) \leq 2 - P(V' < 2U) - P(V' < T + U)$$

or

$$P(W > 1) \leq 1 - C(\alpha,\beta,\gamma'; 02) - C(\alpha,\beta,\gamma; 1,1) \ .$$

Returning to the example, we find

$$\gamma' = \frac{\gamma}{p - 1} = \gamma \tag{4.54}$$

and

$$P(W \geq 1) = \int_0^\infty \int_0^{t'} \int_{2u}^{t'+u'} f \, dt'du'dv' + \int_0^\infty \int_{t'}^\infty \int_{t'+u'}^{2u'} f \, dt'du'dv' \quad , \tag{4.55}$$

as may be determined from Figure 4.1.

The right hand side of the last equation may be written

$$\int_0^\infty \int_0^{t'} \int_0^{t'+u'} - \int_0^\infty \int_0^{t'} \int_0^{2u} + \int_0^\infty \int_0^\infty \int_0^{2u'} - \int_0^\infty \int_0^{t'} \int_0^{2u'}$$

$$- \int_0^\infty \int_0^\infty \int_0^{t'+u'} + \int_0^\infty \int_0^{t'} \int_0^{t'+u'} \tag{4.56}$$

$$= 2[D(1,1,1) - D(1,0,2)] + C(0,2) - C(1,1)$$

where the subscripts n and m together with the first three arguments of each function have been deleted to simplify the notation. Figure 4.1 indicates that the region of integration for $D(1,1,1)$ includes that for $D(1,0,2)$. Hence, we must have $D(1,1,1) \geq D(1,0,2)$. The relative values of $C(1,1)$ and $C(0,2)$ cannot be determined from the figure.

It is seen from (4.56) that we shall encounter two cases: $a = b = c = 1$ and $a = 1$, $b = 0$, $c = 2$. Table 4.1 contains pertinent information for each of the cases.

Completed work sheets for the computing $P(W \geq 1)$ are appended to this chapter as Tables 4.2, ..., 4.12. The tables are related and must be used with equations (4.33), (4.42), and (4.56) to determine the desired probability.

$C(1,1)$ is obtained by combining items labeled 5 in Table 4.4 with items labeled 3 in Table 4.3. When this is done and the result multiplied by

Table 4.1

Ancillary Data

Function	Case: $a = b = c = 1$	Case: $a = 1,\ b = 0,\ c = 2$
$\frac{a\alpha}{\beta}$	2.8	2.8
$\frac{\gamma}{\gamma + c\beta} = \lambda$	0.090909	0.047619
$\frac{\gamma + b\alpha}{\gamma + c\beta} = \mu$	2.636363	0.047619
$\frac{\lambda}{\mu}$	0.034482	1.0
Δ	0.26316	0.26316
Δ_1	0.797101	0.882353

$\Gamma(n+\frac{m}{2}) \Big/ \left\{[\Gamma(\frac{n}{2})]^2\ \Gamma(\frac{m}{2})\right\} = 840$, we have

C(1,1) = 840(0.001190) = 0.999600 .

Similarly, using items 2 in Table 4.5, we have

C(0,2) = 840(0.001157) = 0.971880 .

Had the interest been in finding an upper bound, we have: $P(W > 1) \leq 2 - C(0,2) - C(1,1) = .0285$ (from the paragraph which precedes (4.54)).

The quantity $\Delta = \beta/(\beta + a\alpha)$ is not a function of b and c . Thus, after examining equation (4.42), we conclude that parts of D(1,1,1) and D(1,0,2) will cancel when we take the difference D(1,1,1) - D(1,0,2) . Computation of the part which cancels is unnecessary but has been completed for illustrative purposes. Where values of the incomplete Beta-function were required, linear interpolation in [14] was used.

Items 5 in Table 4.7 and 3 in Table 4.10 are combined with items 3 in Table 4.3 to yield

D(1,1,1) = 840(0.000577 - 0) = .467880 .

Similarly, using Tables 4.7 (again) and 4.12 with Table 4.3, we have

D(1,0,2) = 840(0.0000557 - 0.000028) = 0.444360 .

The computation of $P(W > 1)$ is complete except for substituting our results in equation (4.56). This gives

$$(4.57) \quad P(W > 1) = 2(0.467880 - 0.444360) + 0.971880 - 9.999600 ,$$

$$= 0.01932 .$$

Table 4.2

General Data (1)

i	Item	j = 0	j = 1	j = 2	j = 3
0	1	2			
	2	2			
	3	3			
	4	4			
	5	8			
1	1	3	2		
	2	2	3		
	3	3	4		
	4	5	5		
	5	10	15		
2	1	4	3	2	
	2	2	3	4	
	3	3	4	5	
	4	6	6	6	
	5	12	18	24	
3	1	5	4	3	2
	2	2	3	4	5
	3	3	4	5	6
	4	7	7	7	7
	5	14	21	28	35

Item 1 is $2 + i - j$.
Item 2 is $2 + j$.
Item 3 is $3 + j$.
Item 4 is $4 + i$.
Item 5 is $(2 + j)(4 + i)$.

Table 4.3

General Data (2)

i	Item	j = 0	j = 1	j = 2	j = 3
0	1	1			
	2	1			
	3	1			
	4	24			
1	1	3	3		
	2	1	1		
	3	-3	-3		
	4	60	60		
2	1	3	3	3	
	2	1	2	1	
	3	3	6	3	
	4	120	180	120	
3	1	1	1	1	1
	2	1	3	3	1
	3	-1	-3	-3	-1
	4	210	420	420	210

Item 1 is $\binom{\frac{m}{2} - 1}{i}$.

Item 2 is $\binom{i}{j}$.

Item 3 is $\binom{\frac{m}{2} - 1}{i}\binom{i}{j}$.

Item 4 is $[B_1(\ell, \ell + q)]^{-1}$.

Table 4.4

Computation - C(1,1)

i	Item	j = 0	j = 1	j = 2	j = 3
0	1	0.001189			
	2	0.008264			
	3	0.000010			
	4	0.999990			
	5	0.041666			
1	1	0.000004	0.001189		
	2	0.008264	0.000751		
	3	0	0		
	4	1.000000	1.000000		
	5	0.016667	0.016667		
2	1	0	0.000004	0.001189	
	2	0.008264	0.000751	0.000068	
	3	0	0	0	
	4	1.000000	1.000000	1.000000	
	5	0.008333	0.005555	0.008333	
3	1	0	0	0.000004	0.001189
	2	0.008264	0.000751	0.000068	0.000006
	3	0	0	0	0
	4	1.000000	1.000000	1.000000	1.000000
	5	0.004761	0.002380	0.002380	0.004761

Item 1 is $\left(\frac{\gamma}{\gamma + b\alpha}\right)^{\ell}$. Item 2 is $\left(\frac{\gamma}{\gamma + c\beta}\right)^{q}$.

Item 3 is $\left[\left(\frac{\gamma}{\gamma + b\alpha}\right)^{\ell}\left(\frac{\gamma}{\gamma + c\beta}\right)^{q}\right]$. Item 4 is $\left[1 - \left(\frac{\gamma}{\gamma + b\alpha}\right)^{\ell}\left(\frac{\gamma}{\gamma + c\beta}\right)^{q}\right]$.

Item 5 is $\frac{\Gamma(\ell)\,\Gamma(q)}{\Gamma(\ell + q)}\left[1 - \left(\frac{\gamma}{\gamma + b\alpha}\right)^{\ell}\left(\frac{\gamma}{\gamma + c\beta}\right)^{q}\right]$.

Table 4.5

Computation - C(0,2)

i	Item	j = 0	j = 1	j = 2	j = 3
0	1	0.997732			
	2	0.041572			
1	1	0.997732	0.999892		
	2	0.016629	0.016665		
2	1	0.997732	0.999892	0.999995	
	2	0.008314	0.005554	0.008333	
3	1	0.997732	0.999892	0.999995	1.000000
	2	0.004751	0.002381	0.002381	0.004761

Item 1 is $1 - \left(\frac{\gamma}{\gamma + b\alpha}\right)^q$.

Item 2 is $\frac{\Gamma(\ell)\,\Gamma(q)}{\Gamma(\ell + q + 1)}\left[1 - \left(\frac{\gamma}{\gamma + b\alpha}\right)^q\right]$.

Table 4.6

Computation - D(1,1,1) and D(1,0,2)
(1)

i	Item	j = 0	j = 1	j = 2	j = 3
0	1	7.840000			
	2	0.004796			
	3	0.037600			
	4	0.004700			
1	1	7.840000	21.952000		
	2	0.001262	0.001262		
	3	0.009894	0.027703		
	4	0.000989	0.001846		
2	1	7.840000	21.952000	61.465600	
	2	0.000332	0.000332	0.000332	
	3	0.002603	0.007288	0.020407	
	4	0.000216	0.000404	0.000849	
3	1	7.840000	21.952000	61.465600	172.103680
	2	0.000087	0.000087	0.000087	0.000087
	3	0.000682	0.001910	0.005348	0.014973
	4	0.000048	0.000091	0.000191	0.000427

Item 1 is $\left(\frac{a\alpha}{\beta}\right)^{2+j}$.

Item 2 is Δ^{4+i}

Item 3 is $\left(\frac{a\alpha}{\beta}\right)^{2+j} \Delta^{4+i}$.

Item 4 is $\left(\frac{a\alpha}{\beta}\right)^{2+j} \frac{\Delta^{4+i}}{(2+j)(4+i)}$.

Table 4.7

Computation - D(1,1,1) and D(1,0,2)

(2)

i	Item	j = 0	j = 1	j = 2	j = 3
0	1	0.022094			
	2	0.011047			
	3	0.015747			
1	1	0.003931	0.019746		
	2	0.001965	0.006582		
	3	0.002954	0.008428		
2	1	0.000750	0.003181	0.016556	
	2	0.000375	0.001060	0.004139	
	3	0.000591	0.001464	0.004988	
3	1	0.000150	0.000596	0.002507	0.013980
	2	0.000075	0.000165	0.000627	0.002796
	3	0.000123	0.000255	0.000818	0.003223

Item 1 is $B_{\Delta}(2+i-j,\ 3+j)$.

Item 2 is $\frac{1}{2+j} B_{\Delta}(2+i-j,\ 3+j)$.

Item 3 is $\left(\frac{a\alpha}{\beta}\right)^{2+j} \frac{\Delta^{4+i}}{(2+j)(4+i)} + \frac{1}{2+j} B_{\Delta}(2+i-j,\ 3+j)$.

Table 4.8

Computation - D(1,1,1)
(1)

i	Item	j = 0	j = 1	j = 2	j = 3
0	1				
	2	0.064793			
	3	0.403695			
	4	0.026157			
	5	0.003269			
1	1				
	2	0.064793	0.016493		
	3	0.321786	0.321786		
	4	0.020849	0.005307		
	5	0.002085	0.000353		
2	1				
	2	0.064793	0.016493	0.004198	
	3	0.256496	0.256496	0.256496	
	4	0.016619	0.004230	0.001077	
	5	0.001384	0.000235	0.000044	
3	1				
	2	0.064793	0.016493	0.004198	0.001069
	3	0.204453	0.204453	0.204453	0.204453
	4	0.013247	0.003372	0.000858	0.000219
	5	0.000946	0.000160	0.000030	0.000006

Item 1: See Item 2, Table 4.4 (λ^q) .

Item 2 is $(\frac{a\alpha}{\beta}\lambda)^q$. Item 3 is $\Delta_1^{\ell+q}$.

Item 4 is $(\frac{a\alpha}{\beta}\lambda)^q \Delta_1^{\ell+q}$. Item 5 is $(\frac{a\alpha}{\beta}\lambda)^q \frac{\Delta_1^{\ell+q}}{q(\ell+q)}$.

Table 4.9

Computation - D(1,1,1)
(2)

i	Item	j = 0	j = 1	j = 2	j = 3
0	1	0.971641			
	2	0.485820			
	3	0.489089			
	4	0.004042			
1	1	0.939770	0.992884		
	2	0.469885	0.330961		
	3	0.471970	0.331314		
	4	0.003900	0.000249		
2	1	0.897383	0.982096	0.998278	
	2	0.448691	0.327365	0.249569	
	3	0.450075	0.327600	0.249613	
	4	0.003719	0.000246	0.000017	
3	1	0.846859	0.964899	0.994981	0.999594
	2	0.423429	0.321633	0.248745	0.199918
	3	0.424375	0.321793	0.248775	0.199924
	4	0.003705	0.000242	0.000017	0.000001

Item 1 is $B_{\Delta_1}(\ell, q+1)$. Item 2 is $\frac{1}{q} B_{\Delta_1}(\ell, q+1)$.

Item 3 is $(\frac{a\alpha}{\beta} \lambda)^q \frac{\Delta_1^{\ell+q}}{q(\ell+q)} + \frac{1}{q} B_{\Delta_1}(\ell, q+1)$.

Item 4 is $\lambda^q \left[(\frac{a\alpha}{\beta} \lambda)^q \frac{\Delta_1^{\ell+q}}{q(\ell+q)} + \frac{1}{q} B_{\Delta_1}(\ell, q+1) \right]$.

Table 4.10

Computation - D(1,1,1)
(3)

i	Item	j = 0	j = 1	j = 2	j = 3
0	1	0.001189			
	2	0.000005			
	3	0			
1	1	0.000041	0.001189		
	2	0	0		
	3	0	0		
2	1	0.000001	0.000041	0.001189	
	2	0	0	0	
	3	0	0	0	
3	1	0	0.000001	0.000041	0.001189
	2	0	0	0	0
	3	0	0	0	0

Item 1 is $(\lambda/\mu)^{\ell}$.

Item 2 is λ^{q} .

Item 3 is $(\lambda/\mu)^{\ell} \lambda^{q}\left[\left(\frac{a\alpha}{\beta}\lambda\right)^{q} \frac{\Delta_1^{\ell+q}}{q(\ell+q)} + \frac{1}{q} B_{\Delta_1}(\ell, q+1)\right]$.

Table 4.11

Computation D(1,0,2)
(1)

i	Item	j = 0	j = 1	j = 2	j = 3
0	1	0.002268			
	2	0.017778			
	3	0.606134			
	4	0.010776			
	5	0.001347			
1	1	0.002268	0.000108		
	2	0.017778	0.002370		
	3	0.534824	0.534824		
	4	0.009508	0.001268		
	5	0.000951	0.000084		
2	1	0.002268	0.000108	0.000005	
	2	0.017778	0.002370	0.000316	
	3	0.471904	0.471904	0.471904	
	4	0.008390	0.001118	0.000149	
	5	0.000699	0.000062	0.000006	
3	1	0.002268	0.000108	0.000005	0
	2	0.017778	0.002370	0.000316	0.000042
	3	0.416386	0.416386	0.416386	0.416386
	4	0.007403	0.000987	0.000132	0.000017
	5	0.000528	0.000047	0.000004	0

Item 1 is λ^{2+j} .　　Item 2 is $(\frac{a\alpha}{\beta}\lambda)^{2+j}$.

Item 3 is Δ_1^{4+i} .　　Item 4 is $(\frac{a\alpha}{\beta}\lambda)^{2+j}\ \Delta_1^{4+i}$.

Item 5 is $(\frac{a\alpha}{\beta}\lambda)^{2+j}\ \frac{\Delta_1^{4+i}}{(2+j)(4+i)}$.

Table 4.12

Computation $D(1,0,2)$
(2)

i	Item	j = 0	j = 1	j = 2	j = 3
0	1	0.082836			
	2	0.041418			
	3	0.042765			
	4	0.000097			
1	1	0.032880	0.049946		
	2	0.016440	0.016652		
	3	0.017391	0.016736		
	4	0.000039	0.000002		
2	1	0.016254	0.016627	0.033328	
	2	0.008127	0.005542	0.008332	
	3	0.008826	0.005604	0.008338	
	4	0.000020	0.000001	0	
3	1	0.009147	0.004948	0.009250	0.023809
	2	0.004573	0.001649	0.002312	0.004761
	3	0.005101	0.001696	0.002316	0.004761
	4	0.000012	0	0	0

Item 1 is $B_{\Delta_1}(2+i-j,\ 3+j)$.

Item 2 is $\frac{1}{2+j} B_{\Delta_1}(2+i-j,\ 3+j)$.

Item 3 is $\left(\frac{a\alpha}{\beta}\lambda\right)^{2+j} \frac{\Delta_1^{4+i}}{(2+j)(4+i)} + \frac{1}{2+j} B_{\Delta_1}(2+i-j,\ 3+j)$.

Item 4 is $\lambda^{2+j}\left\{\left(\frac{a\alpha}{\beta}\lambda\right)^{2+j} \frac{\Delta_1^{4+i}}{(2+j)(4+i)} + \frac{1}{2+j} B_{\Delta_1}(2+i-j,\ 3+j)\right\}$.

CHAPTER V

SUMMARY AND SOME ALLIED PROBLEMS

5.1 Basic assumptions and techniques

The foregoing chapters have been based upon the general concepts of test reliability and correlations corrected for attenuation arising in the field of statistics in psychology. Basic to the discussion has been the assumption that the observable variates are jointly distributed according to the multivariate normal distribution in which the means are zero and the covariance matrix $\tilde{\Sigma}$ has a specified pattern as well as being positive definite. The pattern of the covariance matrix is determined by the assumed structures of the variates. But the structures cannot be deduced from the covariance matrix. In order that $\tilde{\Sigma}$ be consistent with the assumed structures, defined just prior to (1.6), we must have positive reliability coefficients ρ_1 and ρ_2 together with $|\rho/\sqrt{\rho_1\rho_2}| < 1$. During much of the analysis, the condition $\rho_1 = \rho_2$ obtained, and with it, the conditions $\rho_1 > 0$, $|\rho/\rho_1| < 1$.

Multivariate analysis techniques, e.g. [1], [3], and [15], were employed throughout. Estimates of covariance matrix elements were derived by the method of maximum likelihood, without reference to restrictions $\rho_1 > 0$ and $|\rho/\rho_1| < 1$.

It was proposed that the resulting estimates be used when it is known that the two inequalities are satisfied. Consequent anomalies were pointed out.

Canonical correlations were established by the methods of [15]. It was found that the canonical variates derived from $\tilde{\Sigma}$ (or Σ) are independent of the specific values of the parameters, these variates being determined when the sign of ρ is known. This finding gave rise to the terms quasi-canonical correlations and quasi-canonical variates. Methods of simple correlation analysis were found applicable to this quasi-canonical correlation.

The asymptotic distribution of estimators $\hat{\sigma}^2$, $\hat{\rho}_1$, $\hat{\rho}$, and w were derived by the methods of [3] and [4]. The exact distribution of w was found by standard techniques which reduced the sample density to the joint distribution of a set of independent statistics whose distributions are known. From that point in the development, derivation of cumulative probabilities for w became a three dimensional calculus problem.

5.2 Summary of results

Consequences of the assumed structures are set forth in Chapter I. The covariance matrix, $\tilde{\Sigma}$, for the observable variates was established there. An orthogonal matrix A yielded transformed variates whose covariance matrix became the basis for a simple test of the hypothesis $\rho_1 = \rho_2$

versus $\rho_1 \neq \rho_2$. The covariance matrix for the transformed variates is diagonal when $\rho_1 = \rho_2$.

Canonical correlations associated with two models, one in which $\rho_1 \neq \rho_2$ and one in which $\rho_1 = \rho_2$, were derived and the previously mentioned results are given in Chapters II and III.

The maximum likelihood estimates, denoted by $\hat{\sigma}^2$, $\hat{\rho}_1$ and $\hat{\rho}$, are determined in Chapter III. These estimates are expressed in terms of independent $\frac{1}{2}\chi^2$ variates t, u, and v defined by (3.31). The asymptotic distributions of $\hat{\sigma}^2$, $\hat{\rho}_1$ $\hat{\rho}$, and w were found to be normal with means and variances specified in paragraph 3.5. Anomalies of the estimators $\hat{\rho}_1$ and w are alluded to in paragraph 5.1, preceding.

The distribution of the statistic w was derived in terms of integrals denoted by A, C, and D. Values of these functions are given as linear combinations of the complete and incomplete Beta-functions. Calculations using these results are illustrated for a sample of size 4 in which the number p of variates in each set is 2 and values for ρ_1 and $\rho_{\xi\eta}$ are specified.

It is true that the cumulative distribution F(w) , given by (4.49), is a function of α, β, and γ and hence ρ_1 and $\rho_{\xi\eta}$. Similarly these parameters are present in the asymptotic distributions to which reference has been made in this chapter. Hence, the exactitude of these distributions is limited to cases in which one of the two parameters is known.

An upper bound for $P(w \geq 1)$ is given prior to (4.54) in terms of the function C. The consistency property was established for w; so that, if $\rho_{\xi\eta} < 1$, $P(w > 1)$ tends to zero with increasing values of n (the sample size).

5.3 Extensions of the theory

This paper sheds some light upon distributional problems connected with reliability coefficients and correlations corrected for attenuation. But assumptions which lead to the covariance matrices $\tilde{\Sigma}$ and Σ limit the usefulness of the results to a relatively small class of problems. It may be possible, for example, to relax the conditions of equal variances and $\rho_1 = \rho_2$ to derive results which are applicable to a much larger class of problems.

Even more radical changes in the model may be amenable to mathematical treatment similar to the present research. Professor Hotelling has suggested the following example: Let $\xi, \eta, \varepsilon_1, \ldots, \varepsilon_{2p}$ be defined as in paragraph 1.3 and let the standardized variables associated with ξ and η be ξ' and η', respectively. Consider random variables $X_1, \ldots, X_{2p}$ with the following structures:

$$X_i = \begin{cases} \alpha_i \xi' + \varepsilon_i & ; \quad i \leq p \\ \alpha_i \eta' + \varepsilon_i & ; \quad i > p , \end{cases}$$

where the coefficients α_i are positive constants depending upon i. Then $\sigma_{\xi'\eta'} = \rho_{\xi'\eta'}$, where standard notation is used. Let $\varepsilon_1, \ldots, \varepsilon_p$ have a common variance σ^2 and $\varepsilon_{p+1}, \ldots, \varepsilon_{2p}$ have a common variance σ'^2.

The constants α_i introduce a variety of relations between elements of the covariance matrix for $X_1, \ldots, X_{2p}$. We simplify notation by putting $X_i = Y_i$ and $\alpha_i = \beta_i$ whenever $i > p$ and leave the remaining notation unchanged.

We have, for example,

$$\sigma^2_{X_i} = \alpha_i^2 + \sigma^2 \quad ,$$

$$\sigma_{Y_i} = \beta_i^2 + \sigma'^2 \quad ,$$

and, when $i \neq j$,

$$\sigma_{X_i X_j} = \alpha_i \alpha_j \quad ,$$

$$\sigma_{Y_i Y_j} = \beta_i \beta_j \quad ,$$

$$\sigma_{X_i Y_j} = \alpha_i \beta_j \rho_{\xi'\eta'} \quad .$$

Among other relations which can be established, when $p \geq 3$, $i \neq j \neq k$, and $\ell \neq m \neq q$, we have

$$\rho_{X_i Y_j} = \frac{\alpha_i \beta_j \rho_{\xi'\eta'}}{\sqrt{(\alpha_i^2 + \sigma^2)(\beta_i + \sigma'^2)}} \quad ,$$

$$\frac{\rho_{X_i X_j}\,\rho_{X_j X_k}}{\rho_{X_j X_k}} = \frac{\alpha_i^2}{\alpha_i^2 + \sigma^2} = \frac{\rho_{X_i X_j}\,\rho_{X_i Y_k}}{\rho_{X_i Y_k}} \quad ,$$

$$\frac{\rho_{Y_iY_j}\,\rho_{Y_iY_k}}{\rho_{X_jY_k}} = \frac{\beta_i^2}{\beta_i^2 + \sigma'^2} = \frac{\rho_{X_iY_j}\,\rho_{Y_jY_k}}{\rho_{X_iY_k}} .$$

Thus, for example,

$$\rho_{X_iY_\ell} = \rho_{\xi'\eta'}\sqrt{\left[\frac{\rho_{X_iX_j}\,\rho_{X_iX_k}}{\rho_{X_jX_k}}\right]\left[\frac{\rho_{Y_\ell Y_m}\,\rho_{Y_\ell Y_q}}{\rho_{Y_mY_q}}\right]}$$

$$= \rho_{\xi'\eta'}\sqrt{\left[\frac{\rho_{X_iX_j}\,\rho_{X_iY_k}}{\rho_{X_jY_k}}\right]\left[\frac{\rho_{Y_\ell Y_m}\,\rho_{Y_\ell Y_q}}{\rho_{Y_mY_q}}\right]}$$

again with $p \geq 3$, $i \neq j \neq k$, and $\ell \neq m \neq q$.

Even when $p = 3$, it is apparent that $\rho_{\xi'\eta'}$ may be expressed in several forms, each involving elements of the correlation matrix. Further consequences of the structures have not been investigated. But the model, which may be suitable for some practical problems, is seen to be much more complicated than those to which Chapters I through IV refer. Models in those chapters may be regarded as special cases of the above.

5.4 Additional problems

In addition to unsolved problems mentioned in the introduction and at the end of Chapter III, there are others directly connected with the present study.

Consider, for example, the upper bound for $P(w > 1)$ as given in paragraph 4.5. Computation of this bound is straightforward but lengthy if n is large. It may be possible to sharpen this result or to establish a servicable upper bound which is more easily computed. Professor W. Hoeffding has suggested an approach to the latter alternative. We can take the Tchebycheff inequality given by [9, page 42] and apply it to the inequality which precedes (4.54. This gives

$$P(W > 1) \leq P(2U - V' < 0) + P(T + U - V' < 0)$$
$$\leq E\{e^{-\lambda(2U - V')}\} + E\{e^{-\mu(T + U - V')}\}$$

for all $\lambda, \mu > 0$. Selection of values for λ and μ can be based upon minimizing the indicated expectations.

In connection with the observation that both the exact and the asymptotic distributions of w contain nuisance parameters, no general methods have been found for testing hypotheses concerning $\rho_{\xi\eta}$ and determining confidence intervals when the reliability coefficients are unknown. Such methods would be useful. Professor S. N. Roy suggests the possibility that, in view of the findings of paragraph 3.6 concerning canonical correlations, the function ζ^* may be a servicable substitute for $\rho_{\xi\eta}$. ζ^* is given by (2.18) and its value is approximately $\rho/\sqrt{\rho_1\rho_2}$ when the reliability coefficients have values close to unity.

Another advantage in replacing the usual definition of correlation corrected for attenuation would be that, under the very reasonable assumption that the population dispersion matrix $\tilde{\Sigma}$ of the 2p variates must be positive definite, we would have $-1 < \zeta^* < 1$. This means that we would not encounter the possibility of this correlation being greater than unity in absolute value, which plagues the usual definition in terms of $\rho/\sqrt{\rho_1 \rho_2}$. It is easy to see that with the usual definition it is hard to suggest such a "natural" way to get around the difficulty. The function ζ^* is the correlation between sums of the first and second sets of p variates each, and well known distributions can be used in connection with ζ^* for the usual statistical purposes. Suppose, for example, that two equivalent forms are available for each of two tests, that the reliability coefficients are unknown, and that otherwise the present results apply. Possibly, the purposes of correction for attenuation can be accomplished through the use of correlations between test totals.

BIBLIOGRAPHY

[1] Anderson, T. W. An Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons, Inc., 1958.

[2] Browne, E. T. Introduction to the Theory of Determinants and Matrices. Chapel Hill: University of North Carolina Press, 1958.

[3] Cramér, Harald. Mathematical Methods in Statistics. Princeton: Princeton University Press, 1956.

[4] Hoeffding, Wassily, and Herbert Robbins. "The Central Limit Theorem for Dependent Variables," Duke Mathematical Journal, Vol. 3 (1948), 773-780.

[5] Hotelling, Harold. "Relations between Two Sets of Variates," Biometrika, Vol. XXVIII (1936), 321-377.

[6] --------"New Light on the Correlation Coefficient and its Transforms," Journal of the Royal Statistical Society, Series B, Vol. XV (1953), 193-232.

[7] Johnson, P. O. Statistical Methods in Research. New York: Prentice-Hall, Inc., 1949.

[8] Kelley, T. L. Fundamentals of Statistics. Cambridge: Harvard University Press, 1947.

[9] Kolmogorov, A. N. Foundations of the Theory of Probability. New York: Chelsea Publishing Company, 1956.

[10] Kuder, G. F., and M. W. Richardson. "The Theory of the Estimation of Test Reliability," Psychometrika, Vol. 2 (1937).

[11] Loevinger, Jane. "A Systematic Approach to the Construction and Evaluation of Tests of Ability," Psychological Monographs, Vol. 64 (1947).

[12] McNemar, Quinn. Psychological Statistics. New York: John Wiley and Sons, Inc., 1955.

[13] Mood, A. M. Introduction to the Theory of Statistics. New York: McGraw-Hill Book Co., Inc., 1950.

[14] Pearson, Karl. Tables of the Incomplete Beta-Function. Cambridge, England: The University Press, 1948.

[15] Roy, S. N. Some Aspects of Multivariate Analysis . New York: John Wiley and Sons, Inc.; and Indian Statistical Institute, Calcutta, 1957.

[16] -------and A. E. Sarhan. "On Inverting a Class of Patterned Matrices," Biometrika, Vol. 43 (1956), 227-231.

[17] Spearman, C. "The Proof and Measurement of Association between Two Things," American Journal of Psychology, Vol. 15 (1904), 72-101.

[18] -------"Correlation Calculated from Faulty Data," British Journal of Psychology, Vol. 3 (1909), 271-295.