RANK CORRELATION IN A SINGLY TRUNCATED BIVARIATE NORMAL DISTRIBUTION

by

Merril Wayne Hume

Thesis submitted to the Graduate Faculty of the
Virginia Polytechnic Institute
in candidacy for the degree of
DOCTOR OF PHILOSOPHY
in
Statistics

APPROVED: M. A. Aitkin, Chairman

Boyd Harshbarger, Head
Department of Statistics

Donald R. Jensen
Department of Statistics

Clyde Y. Kramer
Department of Statistics

Emory Pace
Department of Mathematics

April, 1966
Blacksburg, Virginia

Table of Contents

I. Rank Correlation Coefficients

1.1 Introduction and summary

Considerable attention has been devoted to the rank correlation coefficients of Spearman and Kendall, denoted throughout the sequel as r_S and r_K, respectively. These coefficients are not new. Spearman first proposed his coefficient of rank correlation in 1904. Greiner (1909) and Esscher (1924) considered the coefficient r_K as a method of estimating correlation in a normal population; however, the coefficient was rediscovered by Kendall (1938) who considered it purely as a rank coefficient. These rank correlation coefficients were first proposed as measures of association between two groupings, requiring no assumptions on the parent distribution of the observations. In later work, Moran (1948) and Kendall (1962), among others, considered the distributions of r_S and r_K when the parent distribution is the bivariate normal. We shall be concerned with the moments and related properties of r_S and r_K when the underlying distribution is the singly truncated bivariate normal.

This study closely parallels previous considerations of these rank correlation coefficients arising from a non-truncated bivariate normal distribution. However, it is evident from the complications produced by truncation that analytical results are more difficult to obtain, and, as in the non-truncated case, few of the distributional problems of these coefficients can be fully investigated.

The problem that first arises in this study is the justification for the use of ranking procedures when the parent population has the singly truncated bivariate normal distribution. This question may be answered by considering the correlation between ranks and variate values in each marginal distribution. It is generally the case that if this correlation is high, conclusions drawn from the ranks are similar to those drawn from the variate values. An investigation of the correlation between ranks and variate values for the two marginal distributions of the singly truncated bivariate normal distribution is carried out in Chapter II.

The next problem with which we are concerned is the determination of the expectations of the rank correlation coefficients r_S and r_K for the truncated distribution. These results are obtained in terms of integrals which have been evaluated numerically for selected values of the correlation coefficient of the parent population, point of truncation and sample size. Determination of the variances of these coefficients proved to be a more difficult undertaking. Although analytical expressions for the variances of these coefficients have been obtained for the non-truncated case (in approximate form only for r_S), the complications encountered in the case of the truncated distribution made analytical expressions too difficult to obtain except in special cases. These investigations are presented in Chapter III.

As the variances of r_S and r_K proved to be analytically intractable, they were determined empirically with the use of an electronic computer. Monte Carlo methods were employed to provide estimates of the first four moments of the coefficients r_S and r_K. Although this study was conducted primarily to obtain accurate estimates of the variances, the higher moments indicate the effects of truncation on the distributional forms of r_S and r_K. With knowledge of the variances, a comparison of the coefficients of variation of r_S and r_K was carried out in an attempt to indicate which of the rank correlation coefficients best serves as an estimator of parent correlation. Also, the asymptotic procedures prescribed by Kendall (1962) are greatly enhanced with knowledge of the variances. These empirical investigations are presented in Chapter IV.

Chapter V unifies the major areas of this research with a brief discussion of the applications and drawbacks in the use of r_S and r_K in connection with the general theory of ranking procedures. Examples are given to illustrate the use of the tables developed in previous chapters.

1.2 The rank correlation coefficients of Spearman and Kendall

The rank correlation coefficients r_S and r_K are most commonly used as measures of association between two variables. When the underlying parent distribution of these variables is, in fact, the singly truncated bivariate normal, the rank correlation coefficients r_S and r_K not only provide measures of association between the truncated and non-truncated variables but also may be used as estimators of the underlying parent correlation ρ. Although these coefficients lack the usual desirable properties of estimators, there are two strong justifications for their use. First, they are computationally simple. Second, there are situations in which it is impossible or impractical to obtain quantitative measurements on one or both variables although rankings can be made.

The computation either of the rank correlation coefficients r_S and r_K is straightforward. There are a variety of computational forms available, many of which are given in Kendall (1962). The simplest of these is as follows:

From a set of n pairs of observations* $\{x_i, y_i\}$, $i=1, \ldots, n$, replace the observations by their respective ranks to form the n pairs $\{u_i, v_i\}$, $i=1, \ldots, n$, where the u's and v's are permutations of the integers 1 to n. Then Spearman's rank correlation coefficient, r_S, is the product-moment correlation coefficient of the u_i, v_i and may be computed from the sum of

*In situations where quantitative observations are not obtainable, ranks are used directly.

squared differences. We find

$$(1.1) \qquad r_S = 1 - \frac{6S}{n(n^2 - 1)}$$

where

$$S = \sum_{i=1}^{n} (u_i - v_i)^2 .$$

Kendall's coefficient of rank correlation, r_K, is most readily determined by arranging the u_i in natural order. That is, order the n paired rankings as $\{i, v_i\}$, $i=1, \ldots, n$. Now, count the number of $v_j > v_i$ for all $j > i$ and add the counts to obtain P; then

$$(1.2) \qquad r_K = \frac{4P}{n(n-1)} - 1 .$$

These calculations are demonstrated with the following example. Consider the rankings of a group of nine students, denoted by A, B, . . . , I, on the basis of their abilities in algebra and calculus. These data may be represented as

Student	A	B	C	D	E	F	G	H	I
Algebra	4	3	7	2	6	9	8	1	5
Calculus	5	7	3	1	9	6	8	2	4
d^2	1	16	16	1	9	9	0	1	1

where d^2 is the squared difference of the two rankings. For Spearman's coefficient we have

$$S = 1 + 16 + \ldots + 1$$
$$= 54,$$

hence, from (1.1),

$$r_S = 1 - \frac{6(54)}{9(80)}$$
$$= 0.55 .$$

For Kendall's coefficient, we may consider the algebra ranking in its natural order. Then, the rearranged calculus ranking is

2 1 7 5 4 9 3 8 6

and we have

$$P = 7 + 7 + 2 + 3 + 3 + 0 + 2 + 0$$
$$= 24$$

and, from (1.2),

$$r_K = \frac{4(24)}{9(8)} - 1$$
$$= 0.33 .$$

In the above numerical example, there is a considerable difference between the values of the two rank correlation coefficients r_S and r_K. In practice it is usually the case that, when neither coefficient is close to unity, $r_S \approx 1.5\, r_K$ (see Kendall (1962)), but this is by no means an invariable rule. This result is readily explained.

The scoring of r_K is the simplest possible and assigns a unit score for each pair of individuals, no matter how widely separated they are in the ranking. The scoring for r_S is more elaborate and gives a weight proportional to the square of the distance between individuals.

In this section we have merely introduced the rank correlation coefficients of Spearman and Kendall. Considerable attention has been given to these coefficients. Kendall (1962) presents a comprehensive study of these ranking procedures, incorporating an exhaustive survey of related works; hence, these results will not be reproduced here. However, pertinent results will be given as they arise in connection with the work of later chapters.

1.3 The singly truncated bivariate normal distribution

The singly truncated bivariate normal distribution arises in many situations. The following examples are given as illustrations:

> EXAMPLE 1 (examination selection) : A corporation or educational institution requires all applicants to take an admissions test. Applicants are admitted only if their admissions test scores exceed a "cutting score"; otherwise, they are rejected. At some later time an achievement test is administered to all those admitted.

EXAMPLE 2 (acceptance sampling): In analyses of acceptance sampling data, it may be of interest to correlate a physical characteristic such as size, weight, density, hardness, etc., with a performance characteristic such as life span, sales volume, operating cost, or other characteristic for which observations are available only on accepted items.

It is to be noted that the point of truncation may be preset, as in the acceptance sampling situation, or determined by the scores as in Example 1 if the institution has only a limited number of openings.

As shown in the previous examples, two measurements are taken on an individual--all that is known is the point, k, of truncation. With the assumption that the paired measurements from an individual follow the bivariate normal distribution, we consider the probability density

$$(1.3)\quad g(x, y;\rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)} \left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]\right\}.$$

Discarding all values outside the region $R \equiv \{(x, y): k \leq x < \infty,\ -\infty < y < \infty\}$, the remaining population has in R the probability density

$$g(x, y;\rho)/D \tag{1.4}$$

where

$$D = \iint_R g(x, y;\rho)dxdy.$$

Performing the integration, we find that

$$D = Q(\frac{k-\mu_1}{\sigma_1}) = Q(a) = \frac{1}{\sqrt{2\pi}}\int_a^\infty e^{-\frac{1}{2}t^2} dt$$

where

$$a = \frac{k-\mu_1}{\sigma_1} .$$

As is often the case in selection on the basis of examinations, the results of the admissions tests may be considered as the entire population and the distribution of these scores is truncated at a known point. This leads us to consider the case in which μ_1 and σ_1 are known (whence $a = (k-\mu_1)/\sigma_1$ is also known). The parameters μ_1 and σ_1 of the physical characteristic in acceptance sampling situations may be known from previous experience or assumed known when large samples are taken. Hence, we can standardize the marginal distribution of x in (1.3) with the point of truncation $a = (k - \mu_1)/\sigma_1$ being known. The marginal distribution of y in (1.3) may also be standardized without loss of generality. This result holds for the calculation of the correlation between ranks and variate values (Chapter II) as it is well known that the product-moment correlation coefficient is independent of scale and location. Clearly, since r_S and r_K are calculated from ranks, the moments of these coefficients (Chapters III and IV) are also independent of scale and location parameters. Hence, the underlying parent distribution which is considered, the singly truncated bivariate normal distribution, may be taken as having the probability density

$$(1.5) \qquad f(x, y;\rho) = \frac{1}{2\pi Q(a)\sqrt{1-\rho^2}} \exp\{\frac{-1}{2(1-\rho^2)}[x^2 - 2\rho xy + y^2]\}$$

where $a \leq x < \infty$, $-\infty < y < \infty$ and a is the known standardized point of truncation.

II. Correlation Between Ranks and Variate Values

2.1 Introduction

If N samples of size n are drawn from a population having a continuous distribution with finite variance**, and the sample values are ranked from 1 to n in each sample, the correlation between ranks and variate values may be calculated for the complete set of Nn observations. The limiting value of this correlation, as the number N of samples tends to infinity, is called the correlation between ranks and variate values in samples of size n, and is denoted by C_n. The limiting value of C_n as the sample size n tends to infinity, denoted by C, is non-negative, and for any continuous distribution

$$C_n = \left(\frac{n-1}{n+1}\right)^{\frac{1}{2}} C .$$

The correlation, C_n, between ranks and variate values in samples of size n was first derived by Stuart (1954). This derivation can also be found in Kendall (1962). Stuart showed that

$$(2.1) \qquad C_n = \left[\frac{12(n-1)}{\sigma^2(n+1)}\right]^{1/2} \left[\int_{-\infty}^{\infty} xF(x)dF(x) - \frac{1}{2}\mu\right]$$

** For distributions having no variance, see Stuart (1955).

and that

$$C = \lim_{n\to\infty} C_n$$

(2.2)

$$= \left(\frac{12}{\sigma^2}\right)^{\frac{1}{2}}\left[\int_{-\infty}^{\infty} xF(x)\,dF(x) - \frac{1}{2}\mu\right]$$

where F(x), μ and σ^2 are the distribution function, mean and variance, respectively. Denoting the expression in brackets in (2.2) by $^{\Delta}/4$, Stuart showed that

$$\Delta = 4\int_{-\infty}^{\infty} x\{F(x) - \frac{1}{2}\}\,dF(x)$$

$$= 2\int_{-\infty}^{\infty} F(x)\{1 - F(x)\}dx$$

(2.3)

$$= 2\int_{-\infty}^{\infty}\{\int_{-\infty}^{x}(x-y)\,dF(y)\}\,dF(x).$$

The last form, known as Gini's mean difference, is the most suitable for the work of this chapter. From (2.2) and (2.3), we have

(2.4)

$$C = \frac{\sqrt{3}\,\Delta}{2\sigma}\ .$$

For the normal distribution, $C = (3/\pi)^{\frac{1}{2}} = 0.9772$, and hence for a sample as small as $n = 25$, $C_n = 0.94$ for this distribution. By virtue of this fairly close relationship between ranks and variate values, one might expect that by replacing variate values by ranks and operating on the latter as if they were the primary variates, similar conclusions could be drawn with a large saving in computation.

Stuart (1954) uses the correlation between ranks and variate values in the normal distribution in an attempt to establish, in a unified manner, the asymptotic relative efficiencies against normal alternatives of several distribution-free test statistics which are essentially efficient estimators with variate values replaced by ranks. By relating the asymptotic relative efficiency directly to the magnitude of C, Stuart has shown that the test based on ranks will be powerful, in the sense of having high asymptotic relative efficiency, when the value of C is near unity. However, the relation between asymptotic relative efficiency and small sample power is not well established.

In this chapter, the correlation between ranks and variate values is examined for the two marginal distributions of the singly truncated bivariate normal distribution. This correlation has been evaluated numerically in each case for selected values of a, the point of truncation, and ρ. It is found that truncation has little effect on the correlation, except for very large values of a and ρ.

2.2 The truncated normal distribution

Consider the standardized normal distribution, truncated below at $x = a$. The density function is

$$f(x) = \frac{1}{Q(a)} Z(x)\,, \qquad a \leq x < \infty,$$

where

$$Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,, \quad Q(x) = \int_x^\infty Z(t)\, dt\,.$$

Then from (2.3)

$$\Delta(a) = \Delta = \frac{2}{Q^2(a)} \int_a^\infty \int_a^x (x-y)\, Z(x) Z(y)\, dy dx\,.$$

$$= \frac{1}{Q^2(a)\pi} \int_a^\infty \int_a^x (x-y) \exp\{-\frac{1}{2}(x^2+y^2)\}\, dy dx.$$

To evaluate this, set

$$x = a + r\cos\theta\,, \qquad y = a + r\sin\theta$$

then

$$\Delta(a) = \frac{1}{Q^2(a)\pi} \int_0^\infty \int_0^{\pi/4} r^2(\cos\theta - \sin\theta)$$

$$\exp(-\frac{1}{2}\{(a + r\cos\theta)^2 + (a + r\sin\theta)^2\})\, d\theta\, dr$$

$$= \frac{e^{-a^2}}{Q^2(a)\pi} \int_0^\infty r^2 e^{-\frac{1}{2}r^2} dr \int_0^{\pi/4} (\cos\theta - \sin\theta) e^{-ar(\cos\theta + \sin\theta)} d\theta$$

$$= \frac{e^{-a^2}}{Q^2(a)\pi} \int_0^\infty r^2 e^{-\frac{1}{2}r^2} \left[\frac{e^{-ar(\cos\theta + \sin\theta)}}{-ar} \right]_0^{\pi/4} dr$$

$$(2.5) \quad = \frac{e^{-a^2}}{Q^2(a)\pi a} \int_0^\infty r e^{-\frac{1}{2}r^2} \left[e^{-ar} - e^{-\sqrt{2}ar} \right] dr$$

$$= A(I_1 - I_2) .$$

The first integral is

$$I_1 = \int_0^\infty r e^{-\frac{1}{2}(r^2 + 2ar)} dr = e^{\frac{1}{2}a^2} \int_0^\infty r e^{-\frac{1}{2}(r+a)^2} dr$$

$$= e^{\frac{1}{2}a^2} \int_a^\infty (t-a) e^{-\frac{1}{2}t^2} dt$$

where $t = r + a$

$$= e^{\frac{1}{2}a^2} \{ e^{-\frac{1}{2}a^2} - a\sqrt{2\pi}\, Q(a) \}$$

$$(2.6) \quad = 1 - a R(a)$$

where $R(x) = Q(x)/Z(x)$ is Mills' ratio (Mills (1926)).

The second integral is

$$I_2 = \int_0^{\infty} r\, e^{-\frac{1}{2}(r^2 + 2\sqrt{2}\, ar)} dr = e^{a^2} \int_0^{\infty} r\, e^{-\frac{1}{2}(r+\sqrt{2}a)^2} dr$$

$$= e^{a^2} \int_{a\sqrt{2}}^{\infty} (t - a\sqrt{2}) e^{-\frac{1}{2}t^2} dt$$

where $t = r + \sqrt{2}\, a$

$$= e^{a^2}(e^{-a^2} - a\sqrt{2}\sqrt{2\pi}\, Q(a\sqrt{2}))$$

(2.7) $\quad = 1 - a\sqrt{2}\, R(a\sqrt{2})$.

Hence from (2.5), (2.6) and (2.7)

$$\Delta(a) = \frac{e^{-a^2}}{Q^2(a)\pi} \{\sqrt{2}\, R(a\sqrt{2}) - R(a)\}$$

(2.8) $\quad = \dfrac{2}{R^2(a)} \{\sqrt{2}\, R(a\sqrt{2}) - R(a)\}$.

Now, from (2.4),

(2.9) $\quad C(a) = C = \dfrac{\sqrt{3}\, \Delta(a)}{2\sigma(a)}$

where $\sigma^2(a)$ is the variance of the truncated distribution. It can be shown (see, for example, Aitkin (1964)) that

$$(2.10) \qquad \sigma^2(a) = \frac{R^2(a) + aR(a) - 1}{R^2(a)} .$$

From (2.8), (2.9) and (2.10) we find

$$(2.11) \qquad C(a) = \frac{\sqrt{3}\,\{\sqrt{2}\,R(a\sqrt{2}) - R(a)\}}{R(a)\sqrt{R^2(a) + aR(a) - 1}}$$

$$(2.12) \qquad = \frac{\sqrt{3}\,\{\frac{1}{\sqrt{\pi}} Q(a\sqrt{2}) - Z(a)Q(a)\}}{Q(a)\sqrt{Q^2(a) + aZ(a)Q(a) - Z^2(a)}}$$

It is seen from (2.12) that as $a \to -\infty$, $C(a) \to \sqrt{3/\pi} = 0.9772$ as it must, since this is the value for the non-truncated normal distribution.

The Laplace continued fraction expansion of R(a) (see, for example, Sheppard (1939)),

$$R(a) = \frac{1}{a+}\ \frac{1}{a+}\ \frac{2}{a+}\ \frac{3}{a+} \cdot \cdot \cdot, \quad a > 0,$$

may be used to find the limiting value of C(a) as a tends to infinity. Substituting the third convergent of R(a),

$$R_3(a) = \frac{1}{a+}\,\frac{1}{a+}\,\frac{2}{a}$$

$$= \frac{a^2 + 2}{a^3 + 3a}$$

into (2.11), we have for large a

$$C(a) \approx \frac{\sqrt{3}\left\{\dfrac{2a^2 + 2}{2a^3 + 3a} - \dfrac{a^2 + 2}{a^3 + 3a}\right\}}{\dfrac{a^2 + 2}{a^3 + 3a}\sqrt{\left(\dfrac{a^2 + 2}{a^3 + 3a}\right)^2 + \dfrac{a^2 + 2}{a^2 + 3} - 1}} \; .$$

Simplifying,

$$C(a) \approx \frac{\sqrt{3}\, a^3 (a^2 + 3)}{(2a^2 + 3)(a^2 + 2)\sqrt{a^2 + 4}}$$

and, hence

$$\lim_{a \to \infty} C(a) = \sqrt{3}/2 = 0.8660 \; .$$

However, the approach to this limit is very slow, as shown in the table of C(a) in Section (2.4).

2.3 The non-truncated marginal distribution

For the other marginal distribution of the singly truncated bivariate normal distribution, the density function is

$$f(x) = \frac{1}{2\pi Q(a)\sqrt{1-\rho^2}} \int_a^\infty \exp\{-\frac{1}{2(1-\rho^2)}[x^2 - 2\rho xu + u^2]\} du$$

$$= \frac{1}{Q(a)\sqrt{1-\rho^2}} Z(x) \int_a^\infty Z(\frac{u-\rho x}{\sqrt{1-\rho^2}}) \, du,$$

$$-\infty < x < \infty .$$

Then, from (2.3), for this distribution,

$$\Delta(a, \rho) = \Delta = 2\int_{-\infty}^{\infty} \int_{-\infty}^{x} (x-y) dF(y) dF(x)$$

(2.13)
$$= \frac{2}{Q^2(a)(1-\rho^2)} \int_a^\infty Z(u)du \int_a^\infty Z(t)dt \int_{-\infty}^\infty dx$$

$$\int_{-\infty}^{x} (x-y) Z(\frac{x-\rho u}{\sqrt{1-\rho^2}}) \, Z(\frac{y-\rho t}{\sqrt{1-\rho^2}}) dy .$$

Making the transformation

$$x = \rho u + \sqrt{1-\rho^2}\frac{(v+w)}{\sqrt{2}}, \quad y = \rho t + \sqrt{1-\rho^2}\frac{(v-w)}{\sqrt{2}} ,$$

we have

(2.14) $$\Delta(a,\rho) = \frac{2}{Q^2(a)} \int_a^\infty Z(u)du \int_a^\infty Z(t)dt \int_{-\infty}^\infty dv \int_{\frac{-\rho(u-t)}{\sqrt{2(1-\rho^2)}}}^\infty$$

$$\{\rho(u-t) + \sqrt{2(1-\rho^2)}w\} Z(\frac{v+w}{\sqrt{2}}) Z(\frac{v-w}{\sqrt{2}})\ dw.$$

Now the (v, w) integral is

$$I = \int_{-\infty}^\infty dv \int_{\frac{\rho(t-u)}{\sqrt{2(1-\rho^2)}}}^\infty \{\rho(u-t) + \sqrt{2(1-\rho^2)}w\} Z(v)Z(w)dw$$

$$= \int_{\frac{\rho(t-u)}{\sqrt{2(1-\rho^2)}}}^\infty \{\rho(u-t) + \sqrt{2(1-\rho^2)}w\}\ Z(w)\ dw$$

$$= \rho(u-t)Q\{\frac{\rho(t-u)}{\sqrt{2(1-\rho^2)}}\} + \sqrt{2(1-\rho^2)}\ Z\{\frac{\rho(t-u)}{\sqrt{2(1-\rho^2)}}\}$$

and, hence

(2.15) $$\Delta(a,\rho) = \frac{2}{Q^2(a)} \int_a^\infty \int_a^\infty Z(u)Z(t)[\sqrt{2(1-\rho^2)} Z\{\frac{\rho(t-u)}{\sqrt{2(1-\rho^2)}}\} -$$

$$\rho(t-u)Q\{\frac{\rho(t-u)}{\sqrt{2(1-\rho^2)}}\}]\ du\,dt.$$

(2.16) $$= \frac{2}{Q^2(a)} [I_1 - I_2]\ .$$

To proceed further, we evaluate the first integral in two ways. We have

$$I_1 = \sqrt{2(1-\rho^2)} \int_a^\infty \int_a^\infty Z(u)Z(t)\, Z\left(\frac{\rho(t-u)}{\sqrt{2(1-\rho^2)}}\right) du\, dt$$

$$= \frac{\sqrt{1-\rho^2}}{2\pi^{3/2}} \int_a^\infty \int_a^\infty \exp\left(-\frac{1}{2}\left[u^2 + t^2 + \frac{\rho^2(t-u)^2}{2(1-\rho^2)}\right]\right) du\, dt$$

$$= \frac{\sqrt{1-\rho^2}}{2\pi^{3/2}} \int_a^\infty \int_a^\infty \exp\left(-\frac{2-\rho^2}{4(1-\rho^2)}\left[u^2 - \frac{2\rho^2 ut}{2-\rho^2} + t^2\right]\right) du\, dt$$

$$= \frac{1-\rho^2}{\sqrt{\pi}} \; \frac{1}{2\pi\sigma^2\sqrt{1-\tau^2}} \int_a^\infty \int_a^\infty \exp\left(\frac{-1}{2(1-\tau^2)\sigma^2}\left[u^2 - 2\tau ut + t^2\right]\right) du\, dt$$

where

$$\tau = \frac{\rho^2}{2-\rho^2}, \qquad \sigma^2 = \frac{2-\rho^2}{2},$$

and thus

$$(2.17) \quad I_1 = \frac{1-\rho^2}{\sqrt{\pi}} L\left(\frac{a\sqrt{2}}{\sqrt{2-\rho^2}},\ \frac{a\sqrt{2}}{\sqrt{2-\rho^2}};\ \frac{\rho^2}{2-\rho^2}\right),$$

where $L(h, k;\rho)$ is the volume under the standardized bivariate normal surface with correlation ρ from (h, k) to (∞, ∞). This function is tabulated by Pearson (1952) and others.

Alternatively, from (2. 15),

$$I_1 = \sqrt{2(1-\rho^2)} \int_a^\infty \int_a^\infty Z(u)Z(t)Z\left(\frac{\rho(t-u)}{\sqrt{2(1-\rho^2)}}\right) du dt .$$

Setting

$$x = \frac{t-u}{\sqrt{2}} \quad , \qquad y = \frac{t+u}{\sqrt{2}} \quad ,$$

we have

$$I_1 = \sqrt{2(1-\rho^2)} \int_{a\sqrt{2}}^\infty Z(y) \int_{a\sqrt{2}-y}^{y-a\sqrt{2}} Z(x)Z\left(\frac{\rho x}{\sqrt{1-\rho^2}}\right) dx\, dy$$

$$= \frac{\sqrt{1-\rho^2}}{\sqrt{\pi}} \int_{a\sqrt{2}}^\infty Z(y) \int_{a\sqrt{2}-y}^{y-a\sqrt{2}} Z\left(\frac{x}{\sqrt{1-\rho^2}}\right) dx\, dy$$

$$= \frac{1-\rho^2}{\sqrt{\pi}} \int_{a\sqrt{2}}^\infty Z(y) \left\{ 2P \frac{(y-a\sqrt{2})}{\sqrt{1-\rho^2}} - 1 \right\} dy$$

where $P(x) = 1 - Q(x)$

$$(2.18) \qquad = \frac{1-\rho^2}{\sqrt{\pi}} [2 \int_{a/2}^{\infty} Z(y)\, P(\frac{y - a/2}{\sqrt{1-\rho^2}})\, dy - Q(a/2)] .$$

To evaluate the integral in (2. 18), consider the integral defined by

$$(2.19) \qquad J(a, b, c) = \int_a^{\infty} Z(x)\, P(bx + c)\, dx$$

$$= \int_a^{\infty} Z(x) \int_{-\infty}^{bx+c} Z(y)\, dydx .$$

Setting $v = x, \quad w = bx+c-y$,

we have

$$J(a, b, c) = \frac{1}{2\pi} \int_a^{\infty} \int_0^{\infty} \exp(-\frac{1}{2}\{v^2 + (bv+c-w)^2\}) dwdv$$

$$= \frac{1}{2\pi} \int_a^{\infty} \int_0^{\infty} \exp(-\frac{1+b^2}{2}\{v^2 - \frac{2bv(w-c)}{1+b^2} +$$

$$\frac{1}{1+b^2}(w-c)^2\})\, dwdv$$

$$= \frac{1}{2\pi}\sqrt{1+b^2}\int_a^{\infty}\int_{\frac{-c}{\sqrt{1+b^2}}}^{\infty}\exp\left(-\frac{1+b^2}{2}\left\{x^2 - \frac{2b}{\sqrt{1+b^2}}xy + y^2\right\}\right) dydx$$

hence

(2.20) $$J(a, b, c) = L\left(a, \frac{-c}{\sqrt{1+b^2}}; \frac{b}{\sqrt{1+b^2}}\right).$$

From (2.18), (2.19) and (2.20), we have

(2.21) $$I_1 = \frac{1-\rho^2}{\sqrt{\pi}}\left[2L\left(a\sqrt{2}, \frac{a\sqrt{2}}{\sqrt{2-\rho^2}}; \frac{1}{\sqrt{2-\rho^2}}\right) - Q(a\sqrt{2})\right].$$

Equating (2.17) and (2.21), and rewriting the variables, we have

(2.22) $$L(a, a; 2\theta^2 - 1) = 2L\left(\frac{a}{\theta}, a; \theta\right) - Q\left(\frac{a}{\theta}\right)$$

where

$$\frac{1}{\sqrt{2}} \leq \theta \leq 1 \quad \text{and} \quad -\infty < a < \infty.$$

Consider now the second integral in (2.16); we have from (2.15)

$$I_2 = \rho \int_a^\infty \int_a^\infty (t-u)Z(u)Z(t)\, Q\{\frac{\rho(t-u)}{\sqrt{2(1-\rho^2)}}\}\, dudt .$$

Making the transformation

$$x = \frac{t-u}{\sqrt{2}}, \qquad y = \frac{t+u}{\sqrt{2}} ,$$

we have

$$I_2 = \sqrt{2}\rho \int_{a\sqrt{2}}^\infty Z(y) \int_{a\sqrt{2}-y}^{y-a\sqrt{2}} x\, Z(x)\, Q(\frac{\rho x}{\sqrt{1-\rho^2}})\, dxdy .$$

The x-integral is

$$I = \int_{a\sqrt{2}-y}^{y-a\sqrt{2}} x\, Z(x)\, Q(\frac{\rho x}{\sqrt{1-\rho^2}})\, dx$$

$$= - Z(x)\, Q(\frac{\rho x}{\sqrt{1-\rho^2}})\Big|_{a\sqrt{2}-y}^{y-a\sqrt{2}} - \frac{\rho}{\sqrt{1-\rho^2}} \int_{a\sqrt{2}-y}^{y-a\sqrt{2}} Z(x) Z(\frac{\rho x}{\sqrt{1-\rho^2}})\, dx$$

$$= Z(y-a\sqrt{2})[2P\{\frac{\rho(y-a\sqrt{2})}{\sqrt{1-\rho^2}}\} - 1] - \frac{\rho}{\sqrt{2\pi}}[2P(\frac{y-a\sqrt{2}}{\sqrt{1-\rho^2}}) - 1] .$$

Hence

$$I_2 = \sqrt{2}\rho \int_{a/2}^{\infty} Z(y)\{Z(y-a\sqrt{2})[2P(\frac{\rho(y-a/2)}{\sqrt{1-\rho^2}}) - 1]$$

$$- \frac{\rho}{\sqrt{2\pi}}[2P(\frac{y-a\sqrt{2}}{\sqrt{1-\rho^2}}) - 1]\}\, dy$$

(2.23) $$= \sqrt{2}\rho\{2I_3 - I_4 - \frac{\rho}{\sqrt{2\pi}}(2I_5 - I_6)\}$$

where

$$I_3 = \int_{a\sqrt{2}}^{\infty} Z(y)Z(y-a\sqrt{2})\, P\{\frac{\rho(y-a\sqrt{2})}{\sqrt{1-\rho^2}}\}\, dy$$

$$= \frac{1}{\sqrt{2}}\, Z(a) \int_{a}^{\infty} Z(s)\, P\{\frac{\rho(s-a)}{\sqrt{2(1-\rho^2)}}\}\, ds$$

where $s = \sqrt{2}\, y - a$

$$= \frac{1}{\sqrt{2}}\, Z(a)L(a, \frac{a\rho}{\sqrt{2-\rho^2}}\, ; \frac{\rho}{\sqrt{2-\rho^2}}), \qquad \text{from (2.20)},$$

$$I_4 = \int_{a\sqrt{2}}^{\infty} Z(y)Z(y-a\sqrt{2})\, dy$$

$$= \frac{1}{\sqrt{2}} Z(a)\, Q(a)$$

$$I_5 = \int_{a\sqrt{2}}^{\infty} Z(y)\, P\left(\frac{y-a\sqrt{2}}{\sqrt{1-\rho^2}}\right) dy$$

$$= \quad L\left(a\sqrt{2}\,,\ \frac{a\sqrt{2}}{\sqrt{2-\rho^2}}\,;\ \frac{1}{\sqrt{2-\rho^2}}\right), \quad \text{from (2. 20)}\,,$$

$$I_6 = \int_{a\sqrt{2}}^{\infty} Z(y)dy$$

$$= \quad Q(a\sqrt{2})\ .$$

Collecting the terms of (2. 23), we have

$$I_2 = \sqrt{2}\rho\Big[\sqrt{2}\ Z(a)\ L\left(a,\ \frac{a\rho}{\sqrt{2-\rho^2}}\,;\ \frac{\rho}{\sqrt{2-\rho^2}}\right) - \frac{1}{\sqrt{2}} Z(a)Q(a)$$

$$- \frac{\rho}{\sqrt{2\pi}}\ \Big\{2L\left(a\sqrt{2},\ \frac{a\sqrt{2}}{\sqrt{2-\rho^2}}\,;\ \frac{1}{\sqrt{2-\rho^2}}\right) - Q(a\sqrt{2})\Big\}\Big].$$

The second bracket of this expression may be simplified using (2. 22) to give

(2. 24) $$I_2 = 2\rho Z(a)\,L\left(a,\ \frac{a\rho}{\sqrt{2-\rho^2}}\,;\ \frac{\rho}{\sqrt{2-\rho^2}}\right) - \rho Z(a)\ Q(a)$$

$$- \frac{\rho^2}{\sqrt{\pi}}\, L\left(\frac{a\sqrt{2}}{\sqrt{2-\rho^2}}\,,\ \frac{a\sqrt{2}}{\sqrt{2-\rho^2}}\,;\ \frac{\rho^2}{2-\rho^2}\right)\ .$$

Finally, from (2.16), (2.17) and (2.24), we have

$$\Delta(a, \rho) = \frac{2}{Q^2(a)} \{ \frac{1}{\sqrt{\pi}} L(\frac{a\sqrt{2}}{\sqrt{2-\rho^2}}, \frac{a\sqrt{2}}{\sqrt{2-\rho^2}}; \frac{\rho^2}{2-\rho^2})$$

$$- \rho Z(a) [2L(a, \frac{a\rho}{\sqrt{2-\rho^2}}; \frac{\rho}{\sqrt{2-\rho^2}}) - Q(a)]\}$$

for all a and ρ, with $-\infty < a < \infty$, $|\rho| \leq 1$.

To find C(a, ρ) we have, from (2.4),

$$C(a, \rho) = \frac{\sqrt{3}\, \Delta(a, \rho)}{2\sigma(a, \rho)}$$

where $\sigma^2(a, \rho)$ is the variance of the distribution. It can be shown (see, for example, Aitkin (1964)) that

$$\sigma^2(a, \rho) = \frac{R^2(a) + \rho^2(aR(a) - 1)}{R^2(a)}$$

$$= \frac{Q^2(a) + \rho^2(aZ(a)Q(a) - Z^2(a))}{Q^2(a)} .$$

Thus

$$(2.25) \qquad C(a, \rho) = \frac{\sqrt{3}\{\frac{1}{\sqrt{\pi}} f_1(a, \rho) - \rho Z(a) f_2(a, \rho)\}}{Q(a)\sqrt{Q^2(a) + \rho^2(aZ(a)Q(a) - Z^2(a))}}$$

where

$$f_1(a, \rho) = L(\frac{a\sqrt{2}}{\sqrt{2-\rho^2}}, \frac{a\sqrt{2}}{\sqrt{2-\rho^2}}; \frac{\rho^2}{2-\rho^2})$$

and

$$f_2(a, \rho) = 2L(a, \frac{a\rho}{\sqrt{2-\rho^2}}; \frac{\rho}{\sqrt{2-\rho^2}}) - Q(a) .$$

Thus $C(a, \rho)$ can be evaluated readily if a routine for calculating $L(h, k;\rho)$ is available. A tabulation for selected values of a and ρ may be found in Section (2.4). We note here some special cases.

For all ρ, we have

$$(2.26) \qquad C(a, -\rho) = C(a, \rho).$$

This result follows since, from (2.25),

$$C(a, -\rho) = \frac{\sqrt{3}\{\frac{1}{\sqrt{\pi}} f_1(a, -\rho) + \rho Z(a) f_2(a, -\rho)\}}{Q(a)\sqrt{Q^2(a) + \rho^2(aZ(a)Q(a) - Z^2(a))}}$$

where

$$f_1(a, -\rho) = f_1(a, \rho)$$

since it is a function of ρ^2 alone, while

$$f_2(a, -\rho) = 2L(a, \frac{-a\rho}{\sqrt{2-\rho^2}}; \frac{-\rho}{\sqrt{2-\rho^2}}) - Q(a)$$

$$= 2[Q(a) - L(a, \frac{a\rho}{\sqrt{2-\rho^2}}; \frac{\rho}{\sqrt{2-\rho^2}})] - Q(a)$$

$$= -f_2(a, \rho)$$

and hence

$$C(a, -\rho) = C(a, \rho).$$

For $a = -\infty$, we have

(2.27) $$C(-\infty, \rho) = \sqrt{3}/\pi \quad .$$

For $\rho = 0$, we have

(2.28) $$C(a, 0) = \sqrt{3}/\pi \ .$$

Both (2.27) and (2.28) follow immediately since the resulting $f(x)$ in both cases is that of the normal density function for which it is known that $C = \sqrt{3}/\pi$.

For $\rho = 1$ (or similarly $\rho = -1$, by (2.26)), we have

$$C(a, 1) = \frac{\sqrt{3}\{\frac{1}{\sqrt{\pi}} L(a/2, a/2; 1) - Z(a)[2L(a, a; 1) - Q(a)]\}}{Q(a)\sqrt{Q^2(a) + a\, Z(a)\, Q(a) - Z^2(a)}}$$

$$(2.29) \qquad = \frac{\sqrt{3}\{\frac{1}{\sqrt{\pi}} Q(a/2) - Z(a)\, Q(a)\}}{Q(a)\sqrt{Q^2(a) + aZ(a)Q(a) - Z^2(a)}}$$

and this is of the same form as (2.12). This result follows since for $\rho = \pm 1$ the bivariate normal distribution degenerates and in this case both marginal distributions have the same (truncated) form.

For $a = 0$, we have

$$C(0, \rho) = \frac{\sqrt{3}\{\frac{1}{\sqrt{\pi}} L(0, 0; \frac{\rho^2}{2-\rho^2}) \frac{-\rho}{\sqrt{2\pi}}[2L(0, 0; \frac{\rho}{\sqrt{2-\rho^2}}) - \frac{1}{2}]\}}{\frac{1}{2}\sqrt{\frac{1}{4} - \frac{\rho^2}{2\pi}}} .$$

Now

$$L(0, 0, \rho) = \frac{1}{2}\left(1 - \frac{1}{\pi}\cos^{-1}\rho\right)$$

(Sheppard (1898)) and hence

$$(2.30) \quad C(0, \rho) = \sqrt{\frac{6}{\pi - 2\rho^2}}\left[\sqrt{2} - \rho - \frac{1}{\pi}\left(\sqrt{2}\cos^{-1}\frac{\rho^2}{2-\rho^2} - 2\rho\cos^{-1}\frac{\rho}{\sqrt{2-\rho^2}}\right)\right] .$$

Setting $\rho = 0$, we have

$$C(0,0) = \sqrt{\frac{6}{\pi}}(\sqrt{2} - \frac{\sqrt{2}}{\pi}\cos^{-1} 0)$$

$$= \sqrt{3}/\pi$$

as required from (2. 28). Setting $\rho = \sqrt{2/3} = 0.8155$, we have

$$C(0, \sqrt{2/3}) = \sqrt{\frac{6}{\pi - 4/3}}[\sqrt{2} - \sqrt{2/3} - \frac{1}{\pi}(\sqrt{2}\cos^{-1}\frac{1}{2} - 2\sqrt{2/3}\cos^{-1}\frac{1}{2})]$$

$$= \frac{4 - \sqrt{3}}{\sqrt{3\pi - 4}}$$

(2. 31) $\quad = 0.9737$.

2. 4 Computations and conclusions

Selected values of C(a) defined by (2. 11) and (2. 12) and of $C(a, \rho)$ defined by (2. 25) were evaluated numerically on an IBM 7040 electronic computer. Details of these programs appear in Appendix B . The results of these calculations appear below.

Table 3. 1 Values of C(a)

a	$-\infty$	-4	-3	-2	-1. 86	-1	0
C(a)	0. 9772	0. 9773	0. 9787	0. 9826	0. 9827	0. 9751	0. 9496
1	2	3	4	5	6	7	8
0. 9227	0. 9035	0. 8913	0. 8837	0. 8788	0. 8757	0. 8735	0. 8719
9	10	11	12	13	14	15	25
0. 8708	0. 8700	0. 8693	0. 8688	0. 8685	0. 8681	0. 8679	0. 8667
			50	∞			
			0. 8662	0. 8660			

It will be observed that C(a) has a maximum at about a = -1.86 and then decreases rapidly with increasing a. This maximum was found numerically and could not be determined analytically. For large a, the approach to the limit $\sqrt{3}/2$ is very slow. For all practical purposes it is apparent that truncation has little effect on C(a).

Table 3.2 Values of C(a, ρ)

ρ \ a	-∞	-3	-2	-1	0
0.00	0.9772	0.9772	0.9772	0.9772	0.9772
0.50	0.9772	0.9773	0.9775	0.9772	0.9770
0.70	0.9772	0.9776	0.9783	0.9772	0.9760
0.90	0.9772	0.9782	0.9805	0.9767	0.9695
0.92	0.9772	0.9783	0.9809	0.9766	0.9677
0.94	0.9772	0.9784	0.9813	0.9764	0.9653
0.96	0.9772	0.9785	0.9817	0.9761	0.9620
0.98	0.9772	0.9786	0.9821	0.9758	0.9573
1.00	0.9772	0.9787	0.9826	0.9751	0.9496

ρ \ a	1	2	3	4	5
0.00	0.9772	0.9772	0.9772	0.9772	0.9772
0.50	0.9770	0.9771	0.9772	0.9772	0.9772
0.70	0.9762	0.9766	0.9769	0.9770	0.9772
0.90	0.9684	0.9709	0.9733	0.9749	0.9758
0.92	0.9656	0.9684	0.9716	0.9738	0.9751
0.94	0.9615	0.9644	0.9685	0.9717	0.9737
0.96	0.9553	0.9574	0.9626	0.9672	0.9705
0.98	0.9448	0.9435	0.9485	0.9550	0.9608
1.00	0.9227	0.9035	0.8913	0.8837	0.8788

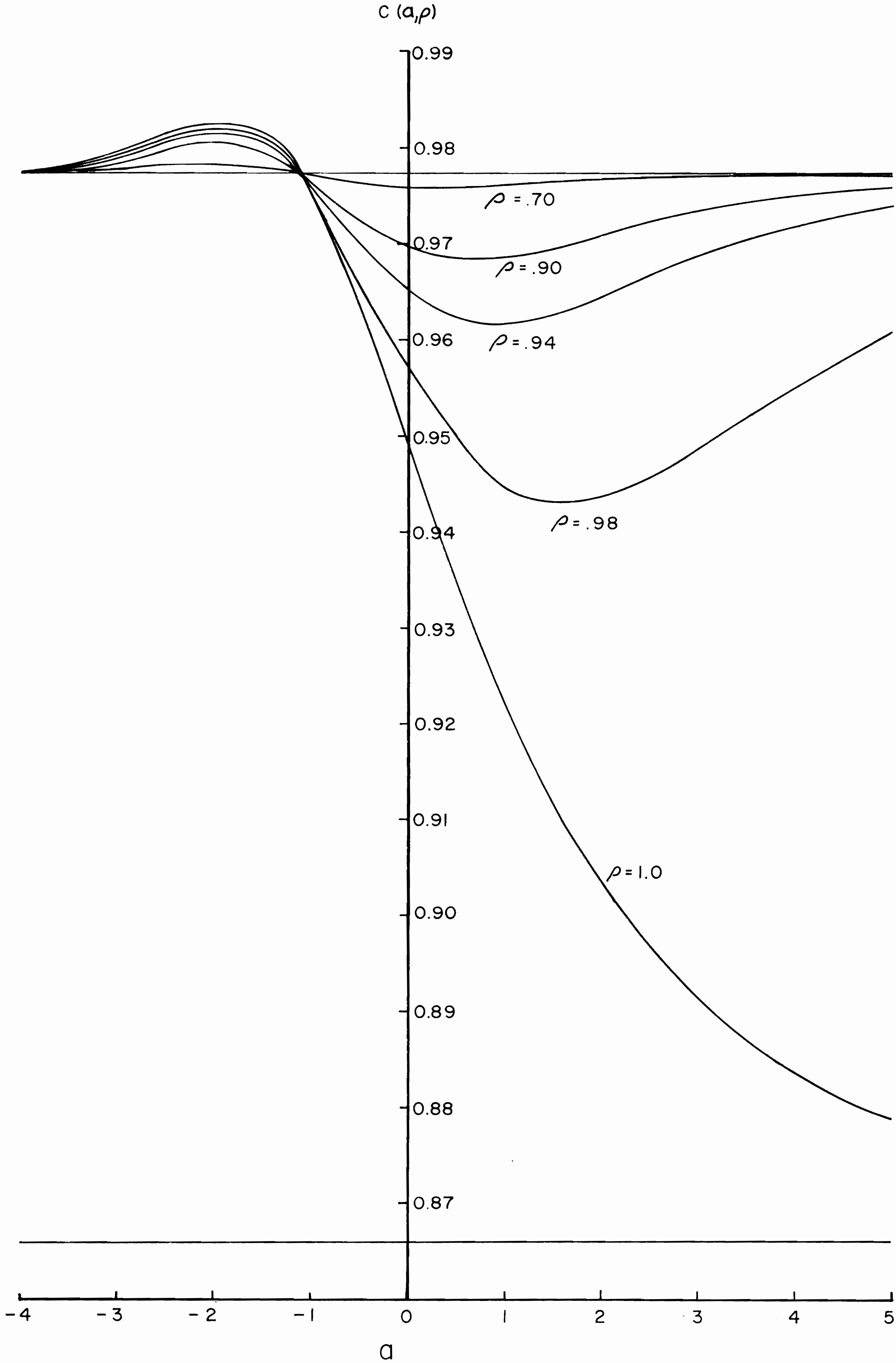
It will be noted that $C(a, \rho)$ also has a maximum at about $a = -2$, and a minimum between $a = 0$ and $a = 2$, depending on the value of ρ. However, the value of $C(a, \rho)$ never falls lower than 0.96 unless $\rho > 0.94$, whatever the value of a. Hence, for all practical purposes, truncation has little effect on $C(a, \rho)$.

Clearly for both marginal distributions of the singly truncated bivariate normal distribution, the correlation between ranks and variate values is little affected by truncation. Therefore, as in the non-truncated distribution, variate values, when available, may be replaced by ranks in the singly truncated bivariate normal distribution with a consequent saving in computation and little loss of efficiency.

A few contours of $C(a, \rho)$ are plotted in Figure 3.1. The upper horizontal line is $C(a, \rho) = \sqrt{3}/\pi$, the lower, $C(a, \rho) = \sqrt{3}/2$. Note $C(a, 1) = C(a)$.

Figure 3.1

C(a, ρ) as a function of a and ρ



III. Moments of the Rank Correlation Coefficients

3.1 Introduction

Considerable attention has been devoted to the rank correlation coefficients r_S and r_K in the case when the parent distribution is the bivariate normal. In this chapter the moments of r_S and r_K are considered for the singly truncated bivariate normal distribution. The expectations of r_S and r_K are obtained in terms of integrals which are evaluated numerically for selected values of parent correlation, ρ, point of truncation, a, and sample size, n. The variances prove to be intractable, except for the cases $\rho = 0$ or ± 1. Empirical estimation of the variances is considered in the next chapter.

3.2 The expectation of Spearman's r_S

Let $(x_1, y_1), \ldots, (x_n, y_n)$ be a sample of n pairs of values from a singly truncated bivariate normal population with probability density

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}\ Q(a)} \exp\{ -\frac{1}{2(1-\rho^2)} [x^2 - 2\rho xy + y^2]\}$$

where

$$a \leq x < \infty, \quad -\infty < y < \infty .$$

To obtain Spearman's r_S we replace $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ by their respective ranks and calculate the product moment correlation of the ranks as if they were variate values.

Let $p(x_i)$, $p(y_i)$ be the ranks of x_i and y_i. Using the technique applied by Moran (1948) for finding $E(r_S | a=-\infty)$, we define a function $H(t)$ such that

$$H(t) = 0 \quad \text{for } t \leq 0$$
$$= 1 \quad \text{for } t > 0,$$

then

$$p(x_i) - 1 = \sum_{j=1}^{n} H(x_j - x_i).$$

Spearman's r_S will be the correlation coefficient of the numbers $p(x_i)$ and $p(y_i)$, and hence also of $\{p(x_i) - 1\}$ and $\{p(y_i) - 1\}$.

Let

$$S = \sum_{i=1}^{n} \{p(x_i) - 1\}\{p(y_i) - 1\}.$$

It can be verified that

$$\sum_{i=1}^{n} \{p(x_i) - 1\} = \frac{n}{2}(n-1)$$

and

$$\sum_{i=1}^{n}\{p(x_i) - \bar{p}\}^2 = \frac{n}{12}(n^2-1)$$

where $\bar{p}$ is the mean of the ranks $p(x_i)$.

It follows that

(3.1) $$r_S = \frac{S - \frac{n}{4}(n-1)^2}{\frac{n}{12}(n^2-1)},$$

and so to find $E(r_S)$ it is enough to find $E(S)$. Now

(3.2) $$S = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} H(x_j-x_i)H(y_k-y_i) .$$

The terms in this expansion for which i=j or i=k are zero. There remain only two cases to consider.

3.2.1 First case

If $i \neq j = k$, we require the joint distribution of $x_j - x_i$ and $y_j - y_i$, for then

$$E(S) = n(n-1)E\{H(x_j-x_i)H(y_j-y_i)\}$$

(3.3) $$= n(n-1)Pr(x_j-x_i > 0,\ y_j-y_i > 0)$$

from the definition of H(t). Now the joint distribution of two independent pairs of values (x_1, y_1) and (x_2, y_2) is

$$f(x_1, x_2, y_1, y_2) = \frac{1}{4\pi^2 Q^2(a)(1-\rho^2)} \exp\{\frac{-1}{2(1-\rho^2)}$$

$$[x_1^2 + x_2^2 - 2\rho(x_1y_1 + x_2y_2) + y_1^2 + y_2^2]\}$$

where

$$a \leq x_1, \ x_2 < \infty, \quad -\infty < y_1, \ y_2 < \infty.$$

The required probability then becomes

$$R(a, \rho) = Pr(x_2-x_1 > 0, \ y_2-y_1 > 0)$$

$$= \int_a^\infty dx_2 \int_a^{x_2} dx_1 \int_{-\infty}^\infty dy_2 \int_{-\infty}^{y_2} f(x_1, x_2, y_1, y_2) dy_1$$

$$= \frac{1}{Q^2(a)(1-\rho^2)} \int_a^\infty Z(x_2)dx_2 \int_a^{x_2} Z(x_1)dx_1 \int_{-\infty}^\infty Z(\frac{y_2-\rho x_2}{\sqrt{1-\rho^2}})dy_2$$

$$\int_{-\infty}^{y_2} Z(\frac{y_1-\rho x_1}{\sqrt{1-\rho^2}})dy_1$$

$$= \frac{1}{Q^2(a)\sqrt{1-\rho^2}} \int_a^\infty Z(x_2)dx_2 \int_a^{x_2} Z(x_1)dx_1 \int_{-\infty}^\infty Z(\frac{y_2-\rho x_2}{\sqrt{1-\rho^2}})$$

$$P(\frac{y_2-\rho x_1}{\sqrt{1-\rho^2}})dy_2$$

$$= \frac{1}{Q^2(a)} \int_a^\infty Z(x_2)dx_2 \int_a^{x_2} Z(x_1)dx_1 \int_{-\infty}^\infty Z(w)P(w + \frac{\rho(x_2-x_1)}{\sqrt{1-\rho^2}})dw$$

where $w = \frac{y_2 - \rho x_2}{\sqrt{1-\rho^2}}$.

Then

$$R(a, \rho) = \frac{1}{Q^2(a)} \int_a^\infty Z(x_2)dx_2 \int_a^{x_2} Z(x_1)\ P(\frac{\rho(x_2 - x_1)}{\sqrt{2(1-\rho^2)}})\ dx_1$$

from (2.19) and (2.20). Set

$$u = \frac{x_2 - x_1}{\sqrt{2}}, \qquad v = \frac{x_2 + x_1}{\sqrt{2}}$$

then

$$R(a, \rho) = \frac{1}{Q^2(a)} \int_0^\infty Z(u)\ P(\frac{\rho u}{\sqrt{1-\rho^2}})du \int_{u+a/2}^\infty Z(v)dv$$

(3.4) $$= \frac{1}{Q^2(a)} \int_0^\infty Z(u)P(\frac{\rho u}{\sqrt{1-\rho^2}})\ Q\ (u + a/2)\ du\ .$$

This can be evaluated on an electronic computer with little difficulty. We note now the following special cases.

For all ρ, we have

$$R(a, -\rho) = \frac{1}{Q^2(a)} \int_0^\infty Z(u) P(\frac{-\rho u}{\sqrt{1-\rho^2}}) Q(u + a\sqrt{2})\, du$$

$$= \frac{1}{Q^2(a)} \int_0^\infty Z(u)\, Q(u + a\sqrt{2})\, du \; - \; R(a, \rho).$$

Differentiating the above integral under the integral sign with respect to a and integrating twice, we find that

$$(3.5) \qquad \int_0^\infty Z(u)\, Q(u + a\sqrt{2}) du = \frac{1}{2}\, Q^2(a)\,.$$

Hence

$$(3.6) \qquad R(a, -\rho) = \frac{1}{2} - R(a, \rho)\,.$$

When $a = -\infty$, we have

$$R(-\infty, \rho) = \int_0^\infty Z(u)\; P\,(\frac{\rho u}{\sqrt{1-\rho^2}})\, du$$

$$= L(0, 0; \rho)$$

$$(3.7) \qquad = \frac{1}{2}[1 - \frac{1}{\pi} \cos^{-1} \rho]\,.$$

This is the appropriate value for the non-truncated distribution, as required.

When $\rho = 0$, we have

$$R(a, 0) = \frac{1}{2Q^2(a)} \int_0^\infty Z(u)Q(u + a\sqrt{2})\, du$$

$$(3.8) \qquad = \frac{1}{4}$$

from (3.5).

When $a = 0$, we have

$$R(0, \rho) = 4 \int_0^\infty Z(u)\, P(\frac{\rho u}{\sqrt{1-\rho^2}})\, Q(u) du.$$

Now

$$\frac{dR(0,\rho)}{d\rho} = \frac{4}{(1-\rho^2)^{3/2}} \int_0^\infty u\, Z(u)\, Z(\frac{\rho u}{\sqrt{1-\rho^2}})\, Q(u) du$$

$$= \frac{4}{\sqrt{2\pi}(1-\rho^2)^{3/2}} \int_0^\infty u Z(\frac{u}{\sqrt{1-\rho^2}}) Q(u)\, du$$

$$= \frac{4}{\sqrt{2\pi}(1-\rho^2)} \int_0^\infty v Z(v)\, Q(\sqrt{1-\rho^2}\, v)\, dv$$

where $v = \frac{u}{\sqrt{1-\rho^2}}$.

Then

$$\frac{dR(0,\rho)}{d\rho} = \frac{4}{\sqrt{2\pi(1-\rho^2)}} \{[-Z(v)Q(\sqrt{1-\rho^2}\, v)]_0^{\infty} - \sqrt{1-\rho^2}\int_0^{\infty} Z(v)$$

$$Z(\sqrt{1-\rho^2}\, v)dv\}$$

$$= \frac{1}{\pi}\{\frac{1}{\sqrt{1-\rho^2}} - \frac{1}{\sqrt{2-\rho^2}}\}$$

$$R(0,\rho) = \frac{1}{\pi}\{\sin^{-1}\rho - \sin^{-1}\frac{\rho}{\sqrt{2}}\} + C$$

where C is a constant independent of ρ. From (3.8) we have

$$R(0,0) = \frac{1}{4} = C$$

so that

(3.9) $$R(0,\rho) = \frac{1}{4} + \frac{1}{\pi}\{\sin^{-1}\rho - \sin^{-1}\frac{\rho}{\sqrt{2}}\} .$$

Note also, from (3.4) and (3.7), that since $Q(u + a\sqrt{2}) \leq 1$, we have

(3.10) $$R(a,\rho) \leq R(-\infty,\rho)$$

with equality if and only if $a = -\infty$.

3.2.2 Second case

If $i \neq j \neq k$, we require the joint distribution of $x_j - x_i$ and $y_k - y_i$, for then

$$E(S) = n(n-1)(n-2)E\{H(x_j - x_i)H(y_k - y_i)\}$$

$$(3.11) \qquad = n(n-1)(n-2)\Pr(x_j - x_i > 0,\ y_k - y_i > 0).$$

Now the joint distribution of the values (x_1, y_1), (x_2, y_2) and (x_3, y_3) is

$$f(x_1, x_2, x_3, y_1, y_2, y_3) = \frac{1}{(2\pi)^3 Q^3(a)(1-\rho^2)^{3/2}} \exp\{\frac{-1}{2(1-\rho^2)}$$

$$[x_1^2 + x_2^2 + x_3^2 - 2\rho(x_1y_1 + x_2y_2 + x_3y_3) + y_1^2 + y_2^2 + y_3^2]\}$$

where

$$a \leq x_1, x_2, x_3 < \infty, \quad -\infty < y_1, y_2, y_3 < \infty.$$

Then

$$R^1(a, \rho) = \Pr(x_2 - x_1 > 0,\ y_3 - y_1 > 0)$$

$$= \int_a^\infty dx_3 \int_a^\infty dx_1 \int_{x_1}^\infty dx_2 \int_{-\infty}^\infty dy_2 \int_{-\infty}^\infty dy_1 \int_{y_1}^\infty$$

$$f(x_1, x_2, x_3, y_1, y_2, y_3)dy_3$$

$$= \frac{1}{Q^3(a)(1-\rho^2)^{3/2}} \int_a^\infty Z(x_3)dx_3 \int_a^\infty Z(x_1)dx_1 \int_{x_1}^\infty Z(x_2)dx_2$$

$$\int_{-\infty}^\infty Z(\frac{y_2-\rho x_2}{\sqrt{1-\rho^2}})dy_2 \int_{-\infty}^\infty Z(\frac{y_1-\rho x_1}{\sqrt{1-\rho^2}})dy_1 \int_{y_1}^\infty Z(\frac{y_3-\rho x_3}{\sqrt{1-\rho^2}})dy_3$$

$$= \frac{1}{Q^3(a)} \int_a^\infty Z(x_3)dx_3 \int_a^\infty Z(x_1)dx_1 \int_{x_1}^\infty Z(x_2)P\{\frac{\rho(x_3-x_1)}{\sqrt{2(1-\rho^2)}}\}dx_2$$

(3.12) $$= \frac{1}{Q^3(a)} \int_a^\infty Z(x_3)dx_3 \int_a^\infty Z(x_1)Q(x_1)P(\frac{\rho(x_3-x_1)}{\sqrt{2(1-\rho^2)}})dx_1 .$$

This double integral can be evaluated on an electronic computer. $R^1(a, \rho)$ can also be expressed as a single integral involving $L(h, k;\rho)$ as tabulated by Pearson (1952). We have

$$R^1(a, \rho) = \frac{1}{Q^3(a)} \int_a^\infty Z(x_1)Q(x_1)\int_a^\infty Z(x_3)P\{\frac{\rho(x_3-x_1)}{\sqrt{2(1-\rho^2)}}\}dx_3dx_1$$

$$(3.13) \qquad = \frac{1}{Q^3(a)} \int_a^\infty Z(x_1)Q(x_1)L\left(a, \frac{\rho x_1}{\sqrt{2-\rho^2}}; \frac{\rho}{\sqrt{2-\rho^2}}\right)dx_1$$

from (2.19) and (2.20). This is easier to evaluate numerically than (3.12) if the values of L are readily available. It can easily be seen that

$$R^1(a, 0) = \frac{1}{2Q^2(a)} \int_a^\infty Z(x_1)Q(x_1)dx_1$$

$$(3.14) \qquad = \frac{1}{4}$$

while

$$R^1(a, -\rho) = \frac{1}{Q^3(a)} \int_a^\infty Z(x_1)Q(x_1)\left[Q(a) - L\left(a, \frac{\rho x_1}{\sqrt{2-\rho^2}}; \frac{\rho}{\sqrt{2-\rho^2}}\right)\right]dx_1$$

$$(3.15) \qquad = \frac{1}{2} - R^1(a, \rho).$$

3.2.3 Expected value of r_S

From (3.1), we have

$$E(r_S) = \frac{12}{n(n^2-1)}\left\{E(S) - \frac{1}{4}n(n-1)^2\right\}$$

and substituting from (3. 3) and (3. 11), we have

$$E(r_S) = \frac{12}{n(n^2-1)} \{n(n-1)R(a,\rho) + n(n-1)(n-2)R^1(a,\rho) - \frac{1}{4}n(n-1)^2\}$$

(3.16) $$= \frac{12}{n+1} \{R(a,\rho) + (n-2)R^1(a,\rho) - \frac{n-1}{4}\} .$$

Writing $E(r_S|a,\rho)$ for $E(r_S)$, we have from (3. 6) and (3. 15),

(3.17) $$E(r_S|a,-\rho) = -E(r_S|a,\rho) .$$

The expected value $E(r_S)$ has been evaluated numerically on an IBM 7040 electronic computer for

ρ: 0.05(0.05) 0.95

a: $-\infty$, -2.0(0.2) 3.0

n: 5(5)25, 50, 100, ∞ .

These tabulations appear in Appendix A and the details of evaluation may be found in Appendix B.

3.3 Variance of Spearman's r_S

It is clear from (3. 1) that the variance of r_S may be determined from the variance of S and hence from $E(S^2)$. Now

$$(3.18) \qquad S^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} \sum_{r=1}^{n} H(x_i - x_j)H(y_i - y_k)H(x_p - x_q)H(y_p - y_r).$$

The terms in this expansion for which $i=j$, $i=k$, $p=q$, or $p=r$ are zero. The number of times each type of term occurs in (3.18) will be $n(n-1)\cdots(n-s+1)$, where s is the number of distinct subscripts.

The evaluation of $E(S^2)$ involves taking the expectations term by term of the sextuple sum and collecting terms. A list of the different terms involved and the numbers of each can be compiled using David and Kendall's symmetric function tables (1949), however, for the sum (3.18), David and Mallows (1961) have given these results as Appendix 1 of their paper. Now each term is the expectation of a product of four H's and we require the probability that each of these products be positive. Hence,

$$(3.19) \quad E(S^2) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{p=1}^{n} \sum_{q=1}^{n} \sum_{r=1}^{n} Pr(x_i - x_j > 0, y_i - y_k > 0, x_p - x_q > 0, y_p - y_r > 0).$$

David and Mallows (1961) have used this technique to obtain an expression for the variance of Spearman's r_S, which is exact for n and to the order of ρ^{12}, for the non-truncated bivariate normal distribution. It is evident from the complications produced by truncation that this procedure cannot be applied to the truncated distribution.

Hence, the variance of Spearman's r_S is analytically intractable with the exception of the cases treated in the following section.

3.3.1 Special cases of the variance of Spearman's r_S

Case 1: $\rho=0$

From the previous section we have, when $\rho=0$,

$$(3.20)\quad E(S^2) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{p=1}^{n}\sum_{q=1}^{n}\sum_{r=1}^{n} Pr(x_i-x_j>0, x_p-x_q>0)$$

$$Pr(y_i-y_k>0, y_p-y_r>0).$$

The joint distributions of the independent x and the independent y values are dependent on the number of distinctly subscripted variates and given by

$$f(x_1,\ldots,x_{n_1}) = \prod_{i=1}^{n_1} \frac{Z(x_i)}{Q(a)}, \qquad a\leq x_i<\infty,$$

and

$$f(y_1,\ldots,y_{n_2}) = \prod_{i=1}^{n_2} Z(y_i), \qquad -\infty<y_i<\infty,$$

where n_1, n_2 are the number of distinct subscripts on the x and y values, respectively.

Appendix 1 of David and Mallows (1961) gives a complete list of the 58 different types of terms which arise in (3.20) and the number of times each occurs. Hence, we need only to evaluate the 58 products of probabilities and sum them with the appropriate number of times each occurs. This task is readily accomplished since

$$\Pr(x_i - x_j > 0,\ x_p - x_q > 0) \quad \text{or} \quad \Pr(y_i - y_k > 0,\ y_p - y_r > 0)$$

may take on only four values, depending upon the subscripts. In fact, writing u for either x or y with the subscripts i, j, p and q being distinct, we have

$$\Pr(u_i - u_j > 0, u_p - u_q > 0) = \Pr(u_i - u_j > 0)\Pr(u_p - u_q > 0) = \frac{1}{4}$$

$$\Pr(u_i - u_j > 0, u_j - u_k > 0) = \Pr(u_i - u_j > 0, u_k - u_i > 0) = \frac{1}{6}$$

$$\Pr(u_i - u_j > 0, u_i - u_k > 0) = \Pr(u_i - u_j > 0, u_k - u_j > 0) = \frac{1}{3}$$

$$\Pr(u_i - u_j > 0, u_i - u_j > 0) = \Pr(u_i - u_j > 0) = \frac{1}{2}.$$

These probabilities are most easily evaluated using nonparametric arguments, however, they may be evaluated directly using the joint distributions given above.

Since the above probabilities are independent of truncation, $E(S^2 \mid \rho=0)$ must also be independent of the truncation parameter, a. This result, as seen from earlier work in this chapter, also holds for $E(S \mid \rho=0)$, hence the variance of r_S is independent of a when $\rho=0$. Thus, as for the non-truncated distribution,

(3.21) $$\operatorname{var}(r_S \mid \rho=0) = \frac{1}{n-1} .$$

Case 2: $\rho = \pm 1$

By the very nature of these extreme correlations, y_i must be a linear function of x_i. Hence, the ranking of the y's would be identical to that of the x's (or a complete reversal). This fact would lead to a constant value of r_S and, thus,

(3.22) $$\operatorname{var}(r_S \mid \rho = \pm 1) = 0.$$

3.4 The expectation of Kendall's r_K

As in Section (3.2), let $(x_1, y_1), \ldots, (x_n, y_n)$ be a sample of n pairs of values from a truncated bivariate normal population with probability density

$$f(xy) = \frac{1}{2\pi Q(a)\sqrt{1-\rho^2}} \exp\{\frac{-1}{2(1-\rho^2)}[x^2 - 2\rho xy + y^2]\}$$

where

$$a \le x < \infty, \qquad -\infty < y < \infty.$$

For each pair of values (i, j) Kendall (1962) defines a term

$$t_{ij} = \mathrm{sgn}(x_i - x_j)\,\mathrm{sgn}(y_i - y_j).$$

Then the sample value of r_K is

$$r_K = \sum_{i<j}\sum t_{ij} / \binom{n}{2} .$$

This definition of r_K is equivalent to the computational form given in Chapter I. Hence to find $E(r_K)$ it is sufficient to find $E(t_{ij})$ since

$$E(r_K) = E\{\sum_{i<j}\sum t_{ij}\} / \binom{n}{2}$$

$$= E(t_{ij}). \tag{3.23}$$

Now the joint distribution of two independent pairs of values (x_1, y_1), (x_2, y_2) is

$$f(x_1, x_2, y_1, y_2) = \frac{1}{4\pi^2 Q^2(a)(1-\rho^2)} \exp\{\frac{-1}{2(1-\rho^2)}$$

$$[x_1^2 + x_2^2 - 2\rho(x_1 y_1 + x_2 y_2) + y_1^2 + y_2^2]\}.$$

To find $E(t_{ij})$ we could proceed as does Kendall by using the characteristic function and the fact that

$$\operatorname{sgn} u = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{e^{itu}}{it}\, dt$$

$$= +1 \quad \text{for } u > 0$$

$$= 0 \quad \text{for } u = 0$$

$$= -1 \quad \text{for } u < 0\,.$$

However, a direct evaluation is much simpler, as we can use the previous results for $E(r_S)$. We have

$$E(r_K) = E(t_{ij}) = \Pr(x_i - x_j > 0,\ y_i - y_j > 0) + \Pr(x_i - x_j < 0,\ y_i - y_j < 0)$$

$$- \Pr(x_i - x_j < 0,\ y_i - y_j > 0) - \Pr(x_i - x_j > 0,\ y_i - y_j < 0).$$

Clearly the first two terms of this expression are equal, as are the last two, since these probabilities do not depend on the particular i and j. In fact, the first term of this expression has already been evaluated as (3. 4),

$$\Pr(x_i - x_j > 0,\ y_i - y_j > 0) = R(a, \rho)$$

and evidently

$$Pr(x_i - x_j < 0,\ y_i - y_j < 0) = R(a, \rho)$$

also. For the last two terms we have similarly

$$Pr(x_i - x_j < 0,\ y_i - y_j > 0) = Pr(x_i - x_j > 0,\ y_i - y_j < 0) = \frac{1}{2}[1 - 2R(a, \rho)],$$

since the four events are mutually exclusive and exhaustive. Hence

$$E(r_K) = 2R(a, \rho) - \{1 - 2R(a, \rho)\}$$

(3.24) $$= 4R(a, \rho) - 1.$$

Thus $E(r_K)$ can be found from $R(a, \rho)$ which has been previously evaluated as (3.4). Writing $E(r_K | a, \rho)$ for $E(r_K)$, we note the following special cases already considered in Section (3.2.1).

$$E(r_K | a, -\rho) = 1 - 4R(a, \rho)$$

(3.25) $$= -E(r_K | a, \rho).$$

$$E(r_K | -\infty, \rho) = 1 - \frac{2}{\pi} \cos^{-1} \rho$$

(3.26) $$= \frac{2}{\pi} \sin^{-1} \rho$$

as in Kendall (1962) for the non-truncated case.

(3.27) $$E(r_K | a, 0) = 0 .$$

(3.28) $$E(r_K|0,\rho) = \frac{4}{\pi}\{\sin^{-1}\rho - \sin^{-1}\frac{\rho}{\sqrt{2}}\}.$$

A tabulation of $E(r_K)$ appears in Appendix A for

ρ: 0.05 (0.05) 0.95

a: $-\infty$, - 2.0 (0.2) 3.0 .

The details of the numerical evaluation may be found in Appendix B.

3.5 The variance of Kendall's r_K

To find the variance of Kendall's r_K in all possible samples, we require

$$E(r_K^2) = \frac{1}{\binom{n}{2}^2} E(\sum_{i<j}\sum t_{ij})^2$$

$$= \frac{1}{\binom{n}{2}^2} E(\sum_{i}<\sum_{j} \sum_{k<}\sum_{l} t_{ij}t_{kl})$$

$$= \frac{1}{\binom{n}{2}^2} \sum_{i} < \sum_{j} \sum_{k} < \sum_{l} E(t_{ij}t_{kl})$$

and there are three cases to consider.

i) If i=k, j=l each term is +1 and the expectation of each term is +1. There are $\binom{n}{2}$ such cases.

ii) If $i \neq k$, $j \neq l$, the expectation of the term $t_{ij}t_{kl}$ is

$$E(t_{ij}t_{kl}) = E(t_{ij})E(t_{kl}) = E^2(r_K).$$

There are $\binom{n}{2}\binom{n-2}{2}$ such cases.

iii) If $i=k$ or $j=l$ but not both, we have a term which may be evaluated by considering

$$(3.29) \quad E(t_{12}t_{13}) = E\{sgn(x_1-x_2)sgn(y_1-y_2)sgn(x_1-x_3)sgn(y_1-y_3)\}\ .$$

There are $6\binom{n}{3}$ such cases.

The characteristic function approach as in Kendall (1962) was found to be unsatisfactory for the evaluation of (3.29), hence, we proceed with the direct approach. $E(t_{12}t_{13})$ is the sum of eight terms, of which a typical one is

$$T = Pr(x_1-x_2 > 0,\ y_1-y_2>0,\ x_1-x_3 > 0,\ y_1-y_3 > 0).$$

The joint distribution of the three pairs of values (x_1, y_1), (x_2, y_2) and (x_3, y_3) is

$$f(x_1, x_2, x_3, y_1, y_2, y_3) = \frac{1}{(2\pi)^3 Q^3(a)(1-\rho^2)^{3/2}} \exp\{\frac{-1}{2(1-\rho^2)}$$

$$[x_1^2 + x_2^2 + x_3^2 - 2\rho(x_1y_1 + x_2y_2 + x_3y_3) + y_1^2 + y_2^2 + y_3^2]\}$$

where

$$a \leq x_i < \infty, \quad -\infty < y_i < \infty \text{ for } i = 1, 2, 3.$$

Let

$$u_1 = \frac{x_1 - x_2}{\sqrt{2}} \qquad v_1 = \frac{y_1 - y_2}{\sqrt{2}}$$

$$u_2 = \frac{x_1 - x_3}{\sqrt{2}} \qquad v_2 = \frac{y_1 - y_3}{\sqrt{2}}$$

$$u_3 = \frac{x_1 + x_2 + x_3}{\sqrt{3}} \qquad v_3 = \frac{y_1 + y_2 + y_3}{\sqrt{3}}$$

where

$$a\sqrt{3} \leq u_3 < \infty, \ \left(\frac{3a - \sqrt{3}u_3}{\sqrt{2}}\right) \leq u_1, u_2 \leq \left(\frac{\sqrt{3}u_3 - 3a}{\sqrt{2}}\right), \ -\infty < v_i < \infty$$

for $i = 1, 2, 3$.

Then the joint distribution of the values $(u_1, u_2, u_3, v_1, v_2, v_3)$ becomes

$$f(u_1, u_2, u_3, v_1, v_2, v_3) = \frac{4}{3(2\pi)^3 Q^3(a)(1-\rho^2)^{3/2}} \exp\{\frac{-1}{18(1-\rho^2)}$$

$$[12u_1^2 + 12u_2^2 + 9u_3^2 - 12u_1u_2$$

$$- 2\rho(12u_1v_1 - 6u_1v_2 - 6u_2v_1 + 12u_2v_2 + 9u_3v_3)$$

$$+ 12v_1^2 + 12v_2^2 + 9v_3^2 - 12v_1v_2]\} .$$

Hence

$$T = \int_{a\sqrt{3}}^{\infty} du_3 \int_0^{\frac{\sqrt{3}u_3 - 3a}{\sqrt{2}}} \int du_2 du_1 \int_0^{\infty} \int dv_1 dv_2 \int_{-\infty}^{\infty} f(u_1, u_2, u_3, v_1, v_2, v_3) dv_3 .$$

The v_3 - integral, including the u_3-exponent, is

$$I_1 = \int_{-\infty}^{\infty} \exp\{\frac{-1}{18(1-\rho^2)}(9v_3^2 - 18\rho u_3 v_3 + 9u_3^2) dv_3$$

$$= 2\pi\sqrt{1-\rho^2}\ Z(u_3) .$$

The v_1 - integral is

$$I_2 = \int_0^{\infty} \exp\{\frac{-1}{18(1-\rho^2)}[12v_1^2 - 12v_1v_2 - 2\rho(12u_1v_1 - 6u_2v_1)]\} dv_1$$

$$= \frac{\sqrt{2\pi}\sqrt{3(1-\rho^2)}}{2} P\{\frac{v_2 - \rho(u_2 - 2u_1)}{\sqrt{3(1-\rho^2)}}\} e^{\{\frac{[v_2 - \rho(u_2 - 2u_1)]^2}{6(1-\rho^2)}\}} .$$

Hence

$$T = \int_{a\sqrt{3}}^{\infty} du_3 \int_0^{\frac{\sqrt{3}u_3 - 3a}{\sqrt{2}}} du_2 du_1 \int_0^{\infty} f(u_1, u_2, u_3 v_2 | v_1 > 0) dv_2$$

where

$$f(u_1, u_2, u_3 v_2 | v_1 > 0) = \frac{2}{\sqrt{3}(2\pi)^{3/2} Q^3(a)\sqrt{1-\rho^2}} Z(u_3)$$

$$P\{\frac{v_2 - \rho(u_2 - 2u_1)}{\sqrt{3(1-\rho^2)}}\} \exp\{\frac{-2}{3(1-\rho^2)}[u_1^2 - u_1 u_2 + u_2^2 - \rho(2u_2 v_2 - u_1 v_2) + v_2^2]$$

$$- \frac{1}{4}[v_2 - \rho(u_2 - 2u_1)]^2\}$$

and this exponent reduces to

$$\frac{-2}{3(1-\rho^2)}[(1-\rho^2)(u_1^2 - u_1 u_2) + u_2^2(1 - \frac{\rho^2}{4}) - \frac{3}{2}\rho\, u_2 v_2 + \frac{3}{4} v_2^2].$$

The v_2-integral is

$$I_3 = \int_0^{\infty} P\{\frac{v_2 - \rho(u_2 - 2u_1)}{\sqrt{3(1-\rho^2)}}\} \exp\{\frac{-1}{2(1-\rho^2)}[v_2^2 - 2\rho u_2 v_2]\} dv_2.$$

$$= \exp(\frac{\rho^2 u_2^2}{2(1-\rho^2)}) \int_0^{\infty} P\{\frac{v_2 - \rho(u_2 - 2u_1)}{\sqrt{3(1-\rho^2)}}\} \exp\{-\frac{1}{2}[\frac{v_2 - \rho u_2}{\sqrt{1-\rho^2}}]^2\} dv_2$$

$$= \exp\left(\frac{\rho^2 u_2^2}{2(1-\rho^2)}\right)\sqrt{1-\rho^2}\int_{\frac{-\rho u_2}{\sqrt{1-\rho^2}}}^{\infty}\sqrt{2\pi}\,Z(w)P\left\{\frac{w}{\sqrt{3}}+\frac{2\rho u_1}{\sqrt{3(1-\rho^2)}}\right\}dw$$

where $w = \dfrac{v_2 - \rho u_2}{\sqrt{1-\rho^2}}$.

Then

$$I_3 = \sqrt{2\pi}\,\sqrt{1-\rho^2}\ \exp\left\{\frac{\rho^2 u_2^2}{2(1-\rho^2)}\right\}L\left(\frac{-\rho u_2}{\sqrt{1-\rho^2}}, \frac{-\rho u_1}{\sqrt{1-\rho^2}}; \frac{1}{2}\right).$$

Hence

$$T = \frac{2}{\sqrt{3}\,Q^3(a)2\pi}\int_{a/3}^{\infty} Z(u_3)du_3\int_0^{\frac{\sqrt{3}u_3-3a}{\sqrt{2}}}\int du_1du_2\exp\left\{\frac{-2}{3}(u_1^2 - u_1u_2+u_2^2)\right\}$$

$$L\left(\frac{-\rho u_2}{\sqrt{1-\rho^2}}, \frac{-\rho u_1}{\sqrt{1-\rho^2}}; \frac{1}{2}\right),$$

or finally,

$$T = \frac{1}{Q^3(a)}\int_{a/3}^{\infty} Z(u_3)du_3\int_0^{\frac{\sqrt{3}u_3-3a}{\sqrt{2}}}\int f(u_1, u_2; \tfrac{1}{2})du_1du_2\int_{\frac{-\rho u_1}{\sqrt{1-\rho^2}}}^{\infty}$$

$$\int_{\frac{-\rho u_2}{\sqrt{1-\rho^2}}}^{\infty} f(x_1, x_2; \tfrac{1}{2})dx_1dx_2$$

where

$$f(x, y;\rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\{\frac{-1}{2(1-\rho^2)}(x^2-2\rho xy+y^2)\} .$$

This cannot be simplified analytically. Thus the variance of Kendall's r_K also proves to be analytically intractable with the exception of the cases of the following section.

3.5.1 Special cases of the variance of Kendall's r_K

Case 1: $\rho=0$

It is evident from the previous section that the variance of Kendall's r_K depends on the evaluation of $E(t_{12}t_{13})$. When the x_i and y_i are independent, we have from (3.29)

$$E(t_{12}t_{13}) = E\{sgn(x_1-x_2)sgn(x_1-x_3)\} E\{sgn(y_1-y_2)sgn(y_1-y_3)\}$$

and the joint distributions of the x_i and y_i are

$$f(x_1, x_2, x_3) = \frac{Z(x_1)Z(x_2)Z(x_3)}{Q^3(a)}, \quad a\leq x_i<\infty,$$

and

$$f(y_1, y_2, y_3) = Z(y_1)Z(y_2)Z(y_3), \quad -\infty<y_i, <\infty ,$$

for i=1, 2, 3.

Now, writing u_i for either x_i or y_i, i=1, 2, 3, we have

$$E\{\mathrm{sgn}(u_1-u_2)\mathrm{sgn}(u_1-u_3)\} = Pr(u_1-u_2>0, u_1-u_3>0) + Pr(u_1-u_2<0, u_1-u_3<0)$$

$$- Pr(u_1-u_2>0,\ u_1-u_3<0) - Pr(u_1-u_2<0, u_1-u_3>0)$$

$$= 2\{Pr(u_1-u_2>0,\ u_1-u_3>0) - Pr(u_1-u_2>0, u_1-u_3<0)\}\ .$$

These two probability statements may be evaluated directly using the joint distributions of the x_i and y_i, i=1, 2, 3. However, since u_1, u_2 and u_3 are independent, identically distributed variates from a continuous distribution, we have immediately

$$Pr(u_1-u_2>0,\ u_1-u_3>0) = 1/3$$

and

$$Pr(u_1-u_2>0,\ u_1-u_3<0) = 1/6.$$

Hence, $E(t_{12}t_{13}) = 1/9$ and with the result (3.27), $E(r_K|\rho=0)=0$, we have

$$\mathrm{var}(r_K|\rho=0) = E(r_K^2|\rho=0)$$

$$= \frac{1}{\binom{n}{2}^2}\left[\binom{n}{2} + \frac{6\binom{n}{3}}{9}\ \right]$$

(3.30)
$$= \frac{1}{\binom{n}{2}}\left[1 + \frac{2(n-2)}{9}\right]$$

$$= \frac{2(2n+5)}{9n(n-1)}$$

as for the non-truncated distribution.

Case 2: $\rho = \pm 1$

As for r_S, we must have

$$\text{var}(r_K \mid \rho = \pm 1) = 0 .$$

IV. Empirical Moments of Rank Correlation Coefficients

4.1 Introduction

The work of the previous chapter demonstrated that the variances of the rank correlation coefficients r_S and r_K are analytically intractable when the parent distribution is the singly truncated bivariate normal. No attempt has been made to obtain analytical expressions for the higher moments of these rank correlation coefficients since this would be even a more formidable task than evaluating the variances. Also, no analytical results have yet been obtained for the higher moments of r_S and r_K in the simpler case when the parent distribution is the non-truncated bivariate normal. However, Sundrum (1953) has given approximations to the third and fourth moments of r_S for the non-truncated case when $\rho = {}^1/_2$ and David and Mallows (1961) point out that their approach is being used to evaluate the third moments of r_S and r_K.

Fieller, Hartley, and Pearson (1957), in an investigation of the normalizing property of R. A. Fisher's transformation

$$z = \tanh^{-1} r = \frac{1}{2} \log_e \frac{1+r}{1-r}$$

when applied to both r_S and r_K for the non-truncated bivariate normal parent distribution, found it necessary to determine empirically the variance of r_S. Their sampling estimates were found to be on the whole quite good when compared to the theoretical values obtained later by David and Mallows (1961). The success of their study encouraged the present empirical estimation of the variances of r_S and r_K for the singly truncated bivariate normal parent distribution.

4.2 Pilot Study

A pilot study was conducted to evaluate the accuracy of the empirical estimation technique and also to determine the computing time necessary for a comprehensive investigation. Values of r_S and r_K were calculated from the same samples of size 10 from non-truncated bivariate normal parent distributions generated on an electronic computer. Nine values of parent correlation were investigated, $\rho = 0.1(0.1)\,0.9$, and the first four crude moments of r_S and r_K obtained for 2500 samples generated for each value of ρ.

A comparison of the observed mean values of r_S and r_K with the known theoretical values proved to be entirely satisfactory. The observed mean values are not reproduced here as this comparison is useful only as a check on the representative character of the random samples.

The observed variances of r_S and r_K agreed with the known theoretical values to within an error of ± 0.004. As in the study conducted by Fieller, Hartley and Pearson (1957), a smoothing of the observed variances alleviated much of the extraneous sampling fluctuation reflected in the third decimal place. The present smoothing process consisted of fitting fourth degree orthogonal polynomials (see, for example, Goulden (1952)) to ten points, namely, the nine observed variances for parent correlation, $\rho = 0.1\,(0.1)\,0.9$, and the known value for the variance at $\rho = 0$. The inclusion of the variance for $\rho=0$ in the smoothing process was found to improve the variance estimates as might well be expected since the standard error of the estimates increases for small ρ. The known variance for $\rho = 1$ was omitted from the smoothing process since the inclusion of this point gave poorer estimates of the true values.

The results of the pilot study for the empirical estimation of the variances of r_S and r_K are given in Table 4.1. The observed and smoothed variance estimates are presented as are the correct theoretical values for comparison. It is noted that the smoothed variance estimates correspond very well to the true values, at least to three decimal places.

Table 4.1 Variance Estimation

	Variance r_S			Variance r_K		
ρ	observed	smoothed	theory	observed	smoothed	theory
0.1	.112	.110	.110	.062	.061	.061
0.2	.105	.106	.106	.061	.060	.060
0.3	.098	.099	.099	.057	.058	.058
0.4	.090	.090	.089	.054	.054	.054
0.5	.076	.078	.078	.049	.050	.050
0.6	.068	.064	.064	.048	.045	.045
0.7	.045	.048	.048	.036	.038	.038
0.8	.031	.031	.031	.030	.030	.030
0.9	.013	.013	.014	.019	.019	.019

A major drawback brought out by the pilot study was the lengthy computing time required to generate a sufficient number of sample values of r_S and r_K to give good estimates of their respective variances. The results of the pilot study as represented in Table 4.1 required about fifty minutes for generation on an IBM 7040 electronic computer. Due to the prohibitive time requirement for determining the effects of truncation on r_S and r_K, it was necessary to have the computing done on a CDC 1604 electronic computer at the Oak Ridge National Laboratory. Since the computing would be conducted by people unfamiliar with the problem and with only a limited amount of computing time available, it was deemed necessary to keep the study of the effects of truncation on r_S and r_K as simple as possible.

4.3 Methodology

A program was written for an electronic computer to empirically determine the first four crude moments of the distributions of r_S and r_K in singly truncated standardized bivariate normal distributions. These moments were obtained for four selected values of the point of truncation, namely $a = -2.0(1.0)\,1.0$. The sample sizes selected for investigation were n=5, 10 and 25, the number of samples taken being 5000 with n=5, 2500 with n=10 and 1000 with n=25. For each combination of a and n, nine different bivariate populations were considered, namely those with $\rho=0.1(0.1)\,0.9$. For each sample of size n, the sample values of r_S and r_K were calculated; hence sampling fluctuations in the moments of r_S will also be reflected in r_K.

Pairs of random deviates from a singly truncated bivariate normal population were generated internally on the electronic computer. Normal deviates were produced by using the transformation

$$v_1 = (-2 \ln u_1)^{\frac{1}{2}} \cos 2\pi u_2$$

$$v_2 = (-2 \ln u_1)^{\frac{1}{2}} \sin 2\pi u_2$$

where u_1 and u_2 are independent random variables from the uniform distribution from zero to one generated by the conventional multiplicative congruential method. Box and Muller (1958) have shown that v_1 and v_2 are independent and normally distributed with means equal to zero and variances equal to one. The v_1 and v_2 were then correlated by the transformation

$$x = v_1$$

$$y = \rho v_1 + (1-\rho^2)^{\frac{1}{2}} v_2 .$$

The resulting x and y have a bivariate normal distribution with correlation ρ, means zero and unit variances. Truncation was introduced by selecting only pairs of variates when $x = v_1 \geq a$ where a is the selected point of truncation.

Once a truncated sample of n pairs of variates $\{x_i, y_i\}$, $i=1, \ldots, n$, has been generated, the calculations of the sample values for Spearman's and Kendall's rank correlation coefficients, r_S and r_K, are readily accomplished by the methods of Section (1. 2).

The first four powers of the sample values of r_S and r_K were calculated. The computer then generated an independent paired sample and, following the same procedure, determined the powers of r_S and r_K for this new sample. The powers of r_K and r_S were aggregated and the process continued for the total number of samples N where N was determined by the sample size n. Hence, empirical estimates of the first four crude moments of r_S and r_K were obtained for the 108 combinations of the three parameters n, a and ρ.

4.4 Results of the empirical estimation

The first four crude moments of r_S and r_K were obtained empirically for singly truncated bivariate normal parent distributions.

The observed mean values of r_S and r_K are not reproduced here, however, comparisons with the theoretical values presented in Appendix A proved to be entirely satisfactory. Such comparisons serve only as a useful check on the representative character of the random samples.

The variances of r_S and r_K were calculated from the empirical crude moments for all the sets of parameters. As in the pilot study, for fixed values of a and n, fourth degree orthogonal polynomials were fitted to the nine observed variances for parent correlation $\rho = 0.1(0.1)0.9$ and the known value of the variance at $\rho = 0$. However, due to extreme sampling fluctuations, this smoothing process did not produce the excellent results found in the pilot study. As the use of orthogonal polynomials proved unsatisfactory, the technique adopted for smoothing the variances was similar to that employed in the study by Fieller, Hartley and Pearson (1957). For fixed a and n, the nine observed variances for parent correlation $\rho = 0.1(0.1)0.9$ were plotted along with the known variance for $\rho = 0$ and, by a rough graphical process, the smoothed variances were obtained. The results of the smoothing process together with the observed variance estimates are presented in Table 4.2. Figures 4.1 and 4.2 graphically represent the smoothed variance estimates of r_S and r_K respectively.

The third and fourth central moments of r_S and r_K were, in general, not calculated. It was felt that the investigation of these higher moments would not add significantly to this study since the higher moments are of little use in describing the forms of the distributions, which have a "saw-tooth" shape for small n (see Kendall (1962)). Also, the sampling variation would not allow accurate estimation of these moments. For selected sets of

Table 4.2 Variances of r_K and r_S

r_K

a:	-2.0						-1.0					
n:	5		10		25		5		10		25	
	O[1]	S	O	S	O	S	O	S	O	S	O	S
ρ[2]:0.1	0.167	0.165	0.059	0.061	0.018	0.020	0.167	0.167	0.064	0.062	0.019	0.020
.2	.160	.162	.059	.061	.019	.019	.170	.165	.057	.061	.020	.020
.3	.152	.157	.061	.059	.019	.019	.161	.161	.061	.060	.017	.019
.4	.150	.151	.056	.056	.018	.018	.155	.156	.056	.057	.019	.019
.5	.144	.142	.052	.052	.016	.016	.149	.149	.054	.054	.018	.018
.6	.131	.130	.046	.046	.014	.014	.134	.139	.049	.049	.016	.016
.7	.111	.114	.037	.039	.012	.012	.123	.124	.042	.042	.013	.014
.8	.093	.093	.029	.031	.009	.009	.104	.103	.035	.034	.011	.011
.9	.063	.063	.021	.021	.005	.005	.074	.074	.024	.025	.007	.007

a:	0						1.0					
n:	5		10		25		5		10		25	
	O	S	O	S	O	S	O	S	O	S	O	S
ρ:0.1	0.165	0.167	0.060	0.062	0.020	0.020	0.173	0.167	0.063	0.062	0.020	0.020
.2	.168	.166	.061	.061	.020	.020	.166	.167	.063	.062	.022	.020
.3	.165	.165	.059	.060	.020	.020	.168	.166	.059	.061	.020	.020
.4	.162	.162	.058	.059	.019	.019	.165	.165	.060	.060	.019	.020
.5	.159	.157	.058	.058	.020	.019	.161	.162	.058	.059	.020	.020
.6	.148	.151	.055	.054	.018	.018	.155	.158	.059	.058	.020	.019
.7	.143	.142	.050	.050	.015	.016	.153	.152	.054	.055	.019	.019
.8	.129	.126	.045	.044	.015	.013	.145	.142	.051	.051	.018	.018
.9	.101	.102	.034	.036	.010	.010	.126	.126	.044	.045	.014	.014

[1] O for observed, S for smoothed

[2] $var(r_K | \rho=0)$ a all values

n	5	10	25
	.1667	.0617	.0204

Table 4.2 Variances of r_K and r_S (continued)

r_S

a:	-2.0						-1.0					
n:	5		10		25		5		10		25	
	O[1]	S	O	S	O	S	O	S	O	S	O	S
ρ[2]:0.1	.250	.247	.108	.110	.037	.041	.250	.249	.115	.110	.040	.041
.2	.239	.239	.105	.106	.038	.039	.252	.245	.103	.107	.041	.040
.3	.221	.227	.106	.100	.037	.037	.236	.236	.106	.102	.034	.038
.4	.214	.214	.093	.092	.033	.034	.222	.223	.094	.095	.035	.036
.5	.196	.195	.083	.081	.029	.029	.206	.206	.089	.086	.032	.033
.6	.169	.169	.067	.068	.022	.023	.180	.183	.076	.075	.028	.028
.7	.130	.136	.048	.053	.017	.017	.155	.155	.059	.060	.026	.022
.8	.098	.098	.030	.035	.010	.010	.120	.119	.043	.042	.014	.015
.9	.051	.050	.015	.013	.004	.003	.068	.069	.020	.020	.006	.006

a:	0						1.0					
n:	5		10		25		5		10		25	
	O	S	O	S	O	S	O	S	O	S	O	S
ρ:0.1	.246	.250	.111	.111	.040	.042	.257	.250	.113	.111	.041	.042
.2	.251	.248	.110	.109	.041	.041	.251	.250	.112	.110	.044	.041
.3	.245	.245	.104	.106	.042	.041	.251	.249	.105	.109	.041	.041
.4	.239	.238	.100	.102	.037	.040	.244	.248	.108	.108	.038	.040
.5	.229	.228	.098	.096	.039	.038	.236	.243	.102	.105	.039	.039
.6	.207	.212	.089	.088	.034	.034	.224	.233	.101	.100	.038	.038
.7	.191	.192	.078	.078	.025	.028	.216	.217	.090	.091	.035	.035
.8	.164	.162	.064	.062	.024	.020	.198	.196	.081	.078	.031	.030
.9	.108	.109	.039	.040	.012	.012	.155	.154	.060	.061	.020	.020

[1] O for observed, S for smoothed

[2] $var(r_S|\rho=0)$ a all values

n	5	10	25
	.2500	.1111	.0417

Figure 4.1 Variance of r_S

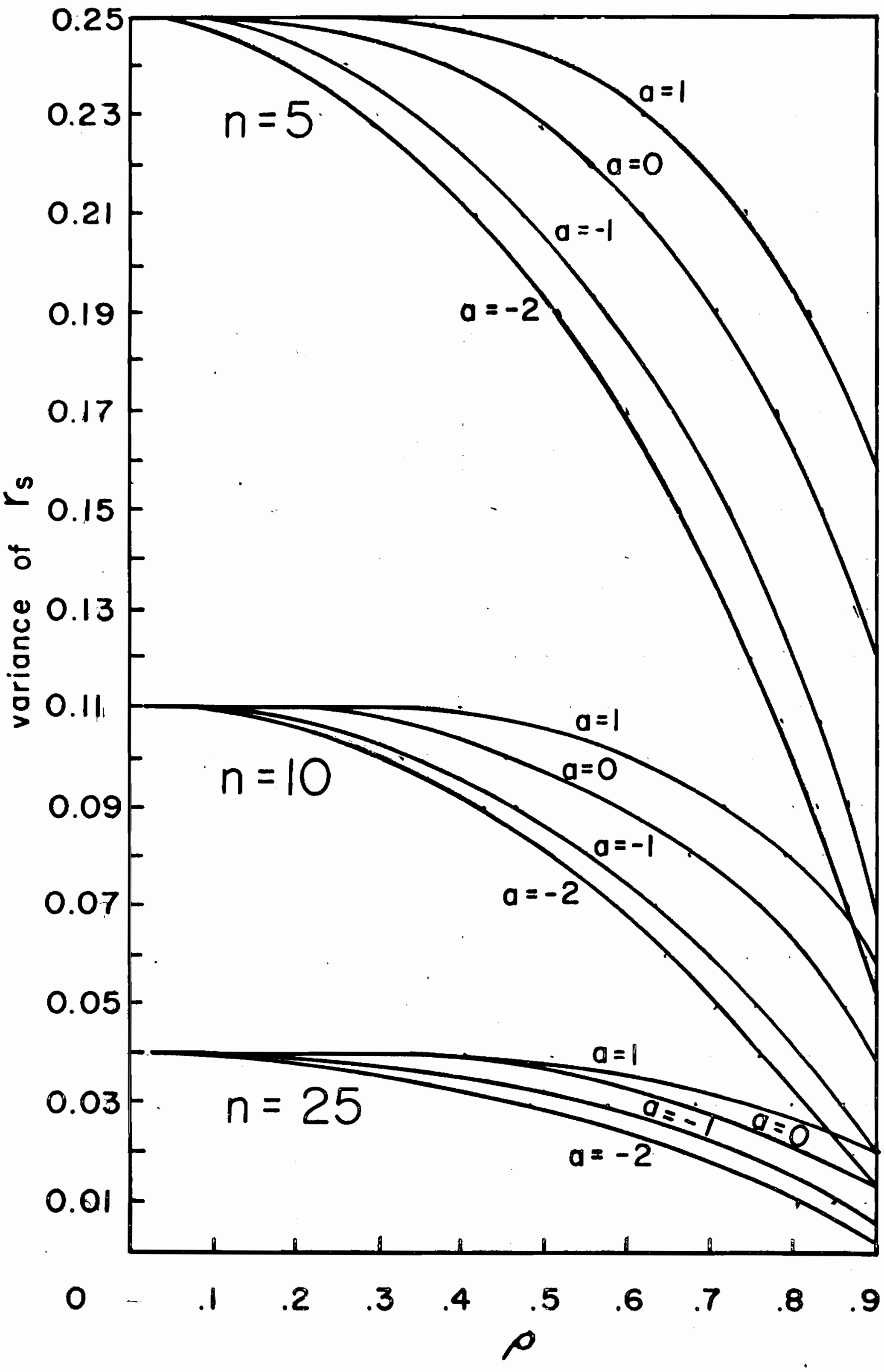
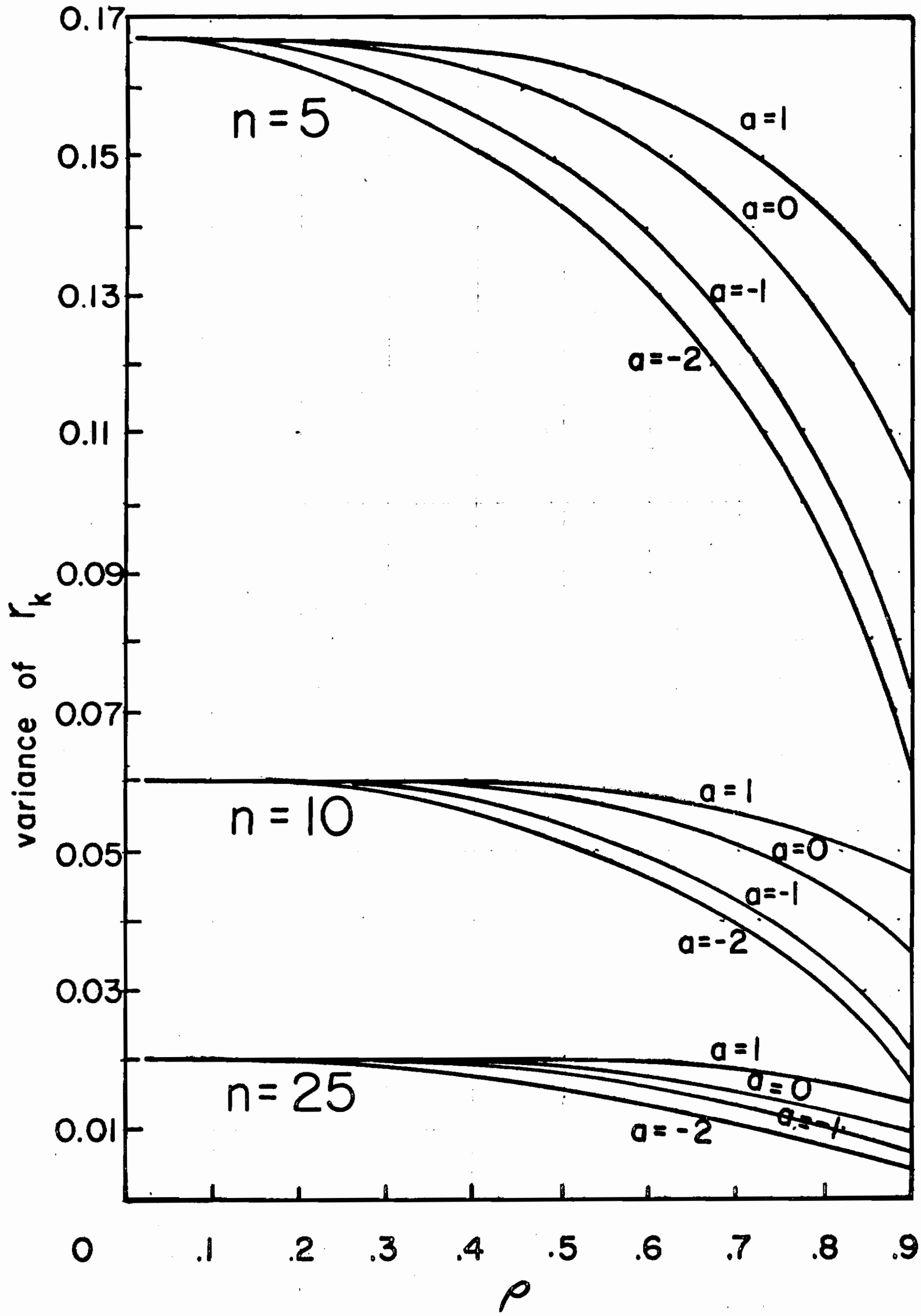


Figure 4.2 Variance of r_K



parameters, however, these higher moments were calculated to gain some insight into the effect of truncation on the distributions of r_S and r_K. The Pearson ratios, $\sqrt{b_1}$ and b_2, measurements of skewness and kurtosis, respectively, were determined from these moments. The Pearson ratios are given by

$$\sqrt{b_1} = m_3/m_2^{3/2}$$

$$b_2 = m_4/m_2^2$$

where

$$m_k = \frac{1}{N}\sum_{i=1}^{N}(r_i - \bar{r})^k, \quad k=1, \ldots, 4,$$ are the k^{th} central moments of r_S or r_K. These results are presented in Table 4.3 for r_K only. A limited investigation of these ratios for r_S showed that the effects of truncation as well as the magnitudes of the b_2 were similar for both coefficients; however, the skewness of r_S, measured by $\sqrt{b_1}$, was considerably larger in absolute value than that of r_K.

Table 4.3 Pearson Ratios for r_K

a	-2.0	-1.0	0.0	1.0
$\rho=0.2$				
	n=5	N=5000		
$\sqrt{b_1}$	-0.14	-0.14	-0.12	-0.12
b_2	2.6	2.5	2.6	2.5
$\rho=0.5$				
	n=5	N=5000		
$\sqrt{b_1}$	-0.46	-0.35	-0.34	-0.23
b_2	2.9	2.7	2.7	2.6
	n=10	N=2500		
$\sqrt{b_1}$	-0.36	-0.36	-0.22	-0.03
b_2	3.0	3.2	2.9	2.8
	n=25	N=1000		
$\sqrt{b_1}$	-0.37	-0.15	-0.19	-0.08
b_2	2.9	3.1	2.9	2.9
$\rho=0.8$				
	n=5	N=5000		
$\sqrt{b_1}$	-0.86	-0.76	-0.63	-0.45
b_2	4.0	3.6	3.2	2.9

4.5 Conclusions

The empirical investigation of the first four moments of r_S and r_K was conducted for two purposes. First, good estimates of the variances of r_S and r_K were obtained and the effects of truncation on these variances determined. The second purpose was to examine the higher moments to gain some insight into the effects of truncation on the distributions of r_S and r_K. Due to the inherent sampling errors of the higher moments, it is difficult to draw conclusions.

The graphical representations of the variances of r_S and r_K, Figures 4.1 and 4.2, respectively, may be extended symmetrically about $\rho=0$. These figures indicate that, as in the non-truncated case, the variances of r_S and r_K are continuous monotonic decreasing functions of ρ^2 for fixed a and n.

The effect of truncation on the variances of r_S and r_K is evident from Figures 4.1 and 4.2. For fixed ρ and n, the variance is a continuous bounded monotonic increasing function of a, the point of truncation. Although truncation increases the variance, it displays a stabilizing effect on the variances for changes in ρ. That is, the rate of change in variance, for increasing $|\rho|$, decreases for large truncation. Clearly, as the variance is a continuous monotonic decreasing function of ρ^2, an upper bound on the variance for increased truncation with fixed ρ and n is the variance for parent correlation of zero.

It was anticipated that, with estimates of the variances available, an investigation of the coefficients of variation of the rank correlation coefficients would indicate which of the two coefficients, r_S or r_K, would be the preferred estimator of parent correlation. However, this investigation proved fruitless, as the coefficients of variation were found to be the same, within sampling variation, for both r_S and r_K for any given point of truncation, parent correlation not close to unity and sample size. Similar results have been given by Kendall (1962) for the non-truncated distribution. A comparison of the coefficients of variation is given in Table 4.4.

The investigation of the higher moments was limited for reasons given in Section 4.4, and it is difficult to draw conclusions with much assurance. However, there is some evidence that, for increased truncation, the skewness (measured by $\sqrt{b_1}$) decreases when ρ is large. There is some indication, as one might well expect, that the shapes of the distributions of r_S and r_K roughly follow that of the product-moment correlation coefficient of the

Table 4.4 Coefficients of Variation

a	n	r_K ρ 0.2	r_K 0.5	r_K 0.8	r_S 0.2	r_S 0.5	r_S 0.8
	5	3.3	1.2	0.53	3.2	1.1	0.46
-2.0	10	2.0	0.71	0.31	2.0	0.67	0.26
	25	1.1	0.40	0.16	1.1	0.38	0.13
	5	5.4	1.9	0.86	5.3	1.9	0.80
0.0	10	3.3	1.2	0.51	3.2	1.1	0.46
	25	1.9	0.67	0.27	1.9	0.66	0.25
	5	7.4	2.7	1.2	7.3	2.6	1.1
1.0	10	4.6	1.6	0.71	4.5	1.6	0.66
	25	2.6	0.94	0.42	2.6	0.91	0.40

bivariate normal distribution (see, for example, Soper, et al. (1917)). However, due to the possible values of r_S or r_K for a given sample size, the distributions of r_S and r_K must be saw-toothed in shape for small n.

V. Applications and Conclusions

5.1 Introduction

The major areas of this research are:

1) correlation between ranks and variate values

2) expectations of r_S and r_K

3) variances of r_S and r_K

and

4) investigations of related properties.

In this chapter we shall unify these investigations with a brief discussion of the applications of and the drawbacks to the use of r_S and r_K.

5.2 Applications

The rank correlation coefficient r, where r may be either r_S or r_K, is of value in two applications for which the underlying parent distribution is the singly truncated bivariate normal. These problems, distinct in application, are related in theory.

One problem of interest is the association between the variables of the singly truncated bivariate normal distribution. In this problem one is interested in using the statistic r to make inferences about r_0, the expected value of the rank correlation coefficient of the truncated distribution.

The other problem of interest is to make, on the basis of truncated data, inferences concerning association in the underlying non-truncated bivariate normal population. This is essentially a problem of estimating the parent correlation ρ with the statistic r.

These two problems of inference are related. The statistic r not only provides an immediate estimate of r_0 but also may be used to estimate ρ. For known point of truncation (and sample size for $r=r_S$), the expected value of r was shown in Chapter III to be a function of ρ alone, say $E(r) = g(\rho)$. Hence, with the inverse transformation, $\rho=g^{-1}[E(r)]$. As the sample statistic r is an estimate of E(r), by assuming $r=E(r)$, $\hat{\rho}=g^{-1}(r)$. The tabulations of $E(r|a, n, \rho)$ given in Appendix A have direct application here. Kendall (1962) recommends that this use of r should be made only if $n \geq 30$. These uses of r will become clearer in the following section.

5.3 Tests of significance

The use of r for tests of significance is thoroughly discussed by Kendall (1962). This section contains a brief review of his methodology as it applies to the underlying singly truncated bivariate normal distribution.

Small sample permutation tests provide a nonparametric basis for tests of significance under the null hypothesis of no association. These tests are well known and tables are readily available (see, for example, Kendall (1962)). The results of our research are of primary value in large sample tests.

Kendall (1962) has shown that for any parent distribution the distribution of r tends to normality as n, the sample size, increase, provided r_0 is not too near unity. The distribution of r_K is known to approach normality more rapidly than that of r_S. For $n \geq 10$, the asymptotic normality of r_K may be used in tests of significance; r_S requires a slightly larger sample size. As Kendall (1962) has recommended these sample sizes for any parent distribution, the convergence to normality will be more rapid when the parent distribution is, in fact, the singly truncated bivariate normal. Indications of this are apparent in Table 4.3.

In the case of large samples, the test statistic for the null hypothesis of no association is

$$z = \frac{r}{\sqrt{\text{var}(r \mid \rho=0)}}$$

where z has the unit normal distribution under the null hypothesis. Expressions for $\text{var}(r \mid \rho=0)$ may be found in Chapter III, Section 3.31

for $r=r_S$ and Section 3.5.1 for $r=r_K$. The null hypothesis is rejected for large values of $|z|$.

If a value of r is found to be significant, we are then interested in setting confidence limits on the true value r_0. This may be done in the usual manner by treating

$$z = \frac{r - r_0}{\sqrt{\text{var}(r)}}$$

as a unit normal deviate. A problem now arises. Var(r) is a function of the unknown parameter ρ. In accordance with the usual practice in the theory of large samples, we may determine an estimate of ρ, $\hat{\rho}$, using the tables of Appendix A. Turning then to Chapter IV and using either Table 4.2 or Figures 4.1 or 4.2, from this value of $\hat{\rho}$ determine an estimate of var(r). This problem of determining var(r) is also encountered when using the general results of Kendall (1962). He gives an upper bound to var(r) for any underlying distribution. It is here that our variance estimates in Chapter IV play an important role. By assuming an underlying singly truncated bivariate normal distribution, we have smaller estimates of variance and, hence, shorter confidence intervals. This may best be illustrated with an example.

<u>Example:</u> In a bivariate sample of $n=25$ the value $r_K=0.340$ is found to be significantly different from zero by the asymptotic normal test. Assuming this sample is randomly drawn from a singly truncated bivariate normal distribution with the point of truncation $a=-1.0$, what can be said about r_0 and ρ ?

For the bivariate sample of $n=25$, the distribution of r_K may be taken to be normal with its mean estimated as 0.340. From Appendix A, we find $\hat{\rho}=0.60$ and Table 4.2 gives the corresponding value $\widehat{var(r)}=0.016$; i.e., a standard deviation of 0.127.

Now the probability of a deviation from the mean as large as 1.96 times the standard deviation (in absolute value) is approximately 0.05 under normal theory. We may therefore say that the approximate 95% confidence interval on r_0 is 0.340 $\pm$ (0.127)(1.96), i.e. $0.091 < r_0 < 0.589$. From Appendix A, the corresponding approximate 95% confidence interval on the parent correlation is $0.19 < \rho < 0.86$.

The above confidence intervals required the assumption of a singly truncated bivariate normal distribution with known point of truncation. If the point of truncation is not known, conservative confidence intervals may be determined using $var(r|\rho=0)$, a function of n alone, for var(r). As Kendall (1962) points out, the use of his upper bound to var(r), valid for any underlying distribution, generally gives very conservative confidence intervals and hence should not be used when the underlying distribution is known.

5.4 Problem areas

The analytic approach of Stuart (1954), as used in Chapter II, appears to give strong justification for replacing variate values by ranks when C_n is large. However, the interpretation of C_n is not entirely clear as it is a product-moment correlation coefficient between discrete and continuous variables. Olkin and Tate (1961) have investigated such product-moment correlation coefficients for various cases; however, their results are not applicable to this situation. Further research is needed in this area.

A basic assumption of this research has been that the standardized point of truncation is known. This will not be the case in applications where good estimates of the mean and variance of the truncated variable are not obtainable or are not known *a priori.*

The problem of tied ranks has not been considered in this research. Due to the underlying continuous distribution of the variates, the anlytical consideration of ties is not of primary importance.

However, in practice ties frequently occur. Kendall (1962) considers tied ranks in some detail.

The choice of using r_S or r_K for practical applications has not been decided. The distribution of r_K tends to normality more rapidly than that of r_S but r_S is the easier to calculate.

Appendix A - Tables of $E(r_K)$ and $E(r_S)$

Tabulations of the expectations of the rank correlation coefficients of Spearman and Kendall, r_S and r_K respectively, are given in this section. These expectations are tabulated for:

parent correlation $\rho = 0.05\ (0.05)\ 0.95$

Point of truncation $a = -\infty,\ -2.0\ (0.2)\ 3.0$

and for Spearman's coefficient,

sample size $n = 5,\ 10,\ 15,\ 20,\ 25,\ 50,\ 100,\ \infty$

as the expectation of Kendall's coefficient is independent of sample size. It is to be remembered that, for both r_S and r_K,

$$E(r \mid a, -\rho) = -E(r \mid a, \rho)\ .$$

E(r_K)

a \ ρ	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
-∞	.0318	.0638	.0959	.1282	.1609	.1940	.2276	.2620	.2972	.3333	.3707	.4097	.4505	.4936	.5399	.5903	.6468	.7129	.7978
-2.0	.0301	.0604	.0908	.1216	.1527	.1844	.2167	.2498	.2839	.3191	.3558	.3942	.4347	.4779	.5245	.5757	.6334	.7014	.7893
-1.8	.0295	.0590	.0888	.1188	.1493	.1804	.2121	.2446	.2781	.3129	.3491	.3872	.4275	.4705	.5171	.5685	.6266	.6954	.7847
-1.6	.0286	.0573	.0862	.1155	.1451	.1754	.2063	.2381	.2710	.3051	.3408	.3783	.4182	.4609	.5075	.5590	.6176	.6873	.7784
-1.4	.0276	.0553	.0832	.1115	.1402	.1694	.1994	.2304	.2624	.2957	.3306	.3675	.4068	.4492	.4955	.5471	.6061	.6768	.7702
-1.2	.0265	.0530	.0798	.1069	.1345	.1627	.1916	.2215	.2525	.2849	.3189	.3549	.3936	.4354	.4813	.5328	.5922	.6640	.7598
-1.0	.0252	.0506	.0761	.1020	.1284	.1553	.1831	.2118	.2416	.2728	.3058	.3409	.3786	.4197	.4651	.5163	.5759	.6488	.7472
-0.8	.0239	.0479	.0722	.0968	.1218	.1475	.1740	.2014	.2300	.2600	.2918	.3257	.3624	.4025	.4472	.4980	.5576	.6314	.7326
-0.6	.0226	.0453	.0682	.0914	.1151	.1395	.1646	.1907	.2179	.2466	.2771	.3098	.3453	.3843	.4281	.4782	.5377	.6121	.7160
-0.4	.0212	.0426	.0641	.0861	.1084	.1314	.1551	.1799	.2057	.2331	.2622	.2935	.3277	.3656	.4082	.4575	.5165	.5914	.6978
-0.2	.0199	.0400	.0602	.0808	.1018	.1235	.1459	.1692	.1937	.2196	.2474	.2773	.3102	.3467	.3881	.4363	.4947	.5697	.6783
0.0	.0187	.0374	.0564	.0757	.0954	.1158	.1368	.1588	.1820	.2066	.2329	.2615	.2929	.3280	.3681	.4151	.4726	.5474	.6578
0.2	.0175	.0350	.0528	.0709	.0894	.1085	.1283	.1490	.1708	.1940	.2190	.2461	.2761	.3098	.3485	.3942	.4506	.5250	.6368
0.4	.0163	.0328	.0494	.0663	.0837	.1016	.1202	.1397	.1602	.1822	.2058	.2316	.2601	.2924	.3296	.3739	.4291	.5027	.6155
0.6	.0153	.0307	.0462	.0621	.0784	.0952	.1126	.1309	.1503	.1710	.1934	.2178	.2450	.2758	.3115	.3544	.4082	.4808	.5942
0.8	.0143	.0287	.0433	.0582	.0734	.0892	.1056	.1228	.1411	.1606	.1817	.2049	.2308	.2602	.2945	.3358	.3881	.4596	.5731
1.0	.0134	.0269	.0406	.0546	.0689	.0837	.0991	.1153	.1325	.1510	.1710	.1929	.2175	.2456	.2784	.3183	.3690	.4392	.5524
1.2	.0126	.0253	.0381	.0512	.0647	.0786	.0932	.1084	.1247	.1421	.1610	.1818	.2052	.2320	.2634	.3018	.3510	.4196	.5322
1.4	.0118	.0238	.0359	.0482	.0609	.0740	.0877	.1021	.1174	.1339	.1518	.1716	.1938	.2193	.2494	.2863	.3340	.4010	.5127
1.6	.0112	.0224	.0338	.0454	.0574	.0698	.0827	.0963	.1108	.1264	.1433	.1621	.1833	.2076	.2364	.2719	.3180	.3834	.4940
1.8	.0106	.0212	.0319	.0429	.0542	.0659	.0781	.0910	.1047	.1195	.1356	.1534	.1736	.1968	.2244	.2585	.3030	.3668	.4760
2.0	.0100	.0201	.0302	.0406	.0513	.0624	.0739	.0861	.0992	.1132	.1285	.1455	.1647	.1869	.2133	.2460	.2891	.3511	.4588
2.2	.0096	.0191	.0287	.0386	.0487	.0592	.0701	.0817	.0941	.1074	.1220	.1381	.1565	.1777	.2030	.2345	.2761	.3365	.4425
2.4	.0092	.0182	.0274	.0367	.0463	.0563	.0667	.0777	.0895	.1021	.1160	.1315	.1489	.1692	.1935	.2238	.2640	.3227	.4270
2.6	.0089	.0175	.0262	.0350	.0442	.0537	.0636	.0741	.0852	.0973	.1106	.1253	.1420	.1615	.1848	.2140	.2528	.3098	.4122
2.8	.0087	.0168	.0251	.0335	.0423	.0513	.0608	.0707	.0814	.0930	.1056	.1197	.1357	.1543	.1767	.2048	.2423	.2978	.3983
3.0	.0085	.0163	.0242	.0322	.0405	.0492	.0582	.0677	.0779	.0890	.1011	.1146	.1299	.1478	.1693	.1964	.2326	.2865	.3851

a \ ρ	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
-∞	.0398	.0796	.1196	.1597	.2001	.2408	.2818	.3233	.3653	.4080	.4514	.4958	.5413	.5883	.6370	.6881	.7426	.8022	.8716
-2.0	.0377	.0755	.1134	.1516	.1901	.2291	.2686	.3087	.3497	.3915	.4345	.4787	.5244	.5719	.6216	.6742	.7306	.7927	.8654
-1.8	.0368	.0737	.1108	.1482	.1859	.2241	.2629	.3025	.3428	.3842	.4268	.4707	.5164	.5640	.6140	.6671	.7243	.7876	.8620
-1.6	.0357	.0716	.1076	.1440	.1807	.2180	.2559	.2946	.3342	.3750	.4170	.4606	.5060	.5536	.6039	.6576	.7158	.7806	.8572
-1.4	.0345	.0691	.1039	.1390	.1746	.2107	.2476	.2853	.3239	.3638	.4052	.4482	.4932	.5407	.5912	.6455	.7048	.7714	.8508
-1.2	.0331	.0663	.0997	.1334	.1676	.2025	.2380	.2745	.3120	.3509	.3913	.4336	.4781	.5253	.5759	.6307	.6912	.7598	.8425
-1.0	.0315	.0632	.0950	.1273	.1600	.1934	.2275	.2626	.2989	.3365	.3759	.4172	.4610	.5077	.5582	.6134	.6749	.7456	.8323
-0.8	.0299	.0599	.0902	.1208	.1519	.1837	.2163	.2499	.2847	.3211	.3591	.3993	.4422	.4883	.5384	.5938	.6562	.7291	.8200
-0.6	.0282	.0566	.0851	.1141	.1436	.1738	.2047	.2368	.2701	.3049	.3416	.3805	.4222	.4674	.5170	.5723	.6355	.7104	.8057
-0.4	.0265	.0532	.0801	.1074	.1353	.1638	.1931	.2235	.2552	.2885	.3236	.3612	.4016	.4457	.4945	.5495	.6132	.6899	.7896
-0.2	.0249	.0499	.0752	.1009	.1270	.1539	.1816	.2104	.2405	.2721	.3057	.3418	.3808	.4237	.4715	.5259	.5898	.6679	.7720
0.0	.0233	.0468	.0705	.0945	.1191	.1444	.1705	.1976	.2261	.2562	.2882	.3227	.3602	.4017	.4484	.5020	.5657	.6449	.7530
0.2	.0218	.0438	.0660	.0885	.1116	.1353	.1598	.1855	.2123	.2408	.2713	.3042	.3402	.3802	.4256	.4782	.5414	.6214	.7330
0.4	.0204	.0410	.0617	.0828	.1045	.1267	.1498	.1739	.1993	.2263	.2552	.2865	.3210	.3595	.4034	.4549	.5173	.5976	.7123
0.6	.0191	.0383	.0578	.0776	.0978	.1188	.1404	.1631	.1871	.2125	.2399	.2698	.3027	.3397	.3821	.4322	.4937	.5740	.6912
0.8	.0179	.0359	.0541	.0727	.0917	.1113	.1317	.1531	.1757	.1998	.2257	.2541	.2855	.3210	.3619	.4105	.4709	.5507	.6700
1.0	.0168	.0337	.0508	.0682	.0860	.1045	.1237	.1438	.1651	.1879	.2125	.2394	.2694	.3033	.3427	.3899	.4489	.5281	.6488
1.2	.0157	.0316	.0477	.0640	.0808	.0982	.1162	.1352	.1553	.1769	.2002	.2258	.2543	.2868	.3247	.3704	.4280	.5062	.6278
1.4	.0148	.0297	.0448	.0602	.0760	.0924	.1094	.1274	.1464	.1668	.1889	.2132	.2404	.2715	.3079	.3520	.4081	.4851	.6073
1.6	.0140	.0280	.0422	.0568	.0717	.0871	.1032	.1201	.1381	.1574	.1784	.2016	.2275	.2572	.2922	.3348	.3893	.4651	.5872
1.8	.0132	.0264	.0399	.0536	.0677	.0823	.0975	.1135	.1306	.1489	.1688	.1908	.2156	.2440	.2776	.3187	.3717	.4459	.5678
2.0	.0125	.0250	.0377	.0507	.0640	.0778	.0922	.1074	.1236	.1410	.1600	.1810	.2046	.2318	.2640	.3036	.3551	.4278	.5490
2.2	.0119	.0238	.0358	.0481	.0607	.0738	.0875	.1019	.1173	.1338	.1519	.1719	.1945	.2205	.2514	.2897	.3395	.4107	.5309
2.4	.0113	.0226	.0340	.0457	.0577	.0701	.0831	.0968	.1114	.1272	.1444	.1635	.1851	.2101	.2398	.2766	.3250	.3945	.5135
2.6	.0108	.0215	.0324	.0435	.0549	.0667	.0791	.0921	.1061	.1211	.1375	.1558	.1765	.2004	.2290	.2645	.3113	.3792	.4969
2.8	.0104	.0206	.0309	.0415	.0523	.0636	.0754	.0878	.1011	.1155	.1312	.1487	.1685	.1914	.2189	.2532	.2986	.3647	.4810
3.0	.0099	.0196	.0295	.0396	.0499	.0607	.0720	.0839	.0966	.1103	.1253	.1420	.1610	.1831	.2096	.2427	.2866	.3511	.4657

$E(r_S)$ n = 10

a \ ρ	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
-∞	.0434	.0869	.1304	.1741	.2179	.2620	.3064	.3511	.3963	.4419	.4881	.5349	.5826	.6313	.6812	.7326	.7861	.8427	.9051
-2.0	.0411	.0823	.1236	.1652	.2071	.2494	.2922	.3355	.3796	.4244	.4702	.5170	.5651	.6146	.6658	.7190	.7748	.8343	.9000
-1.8	.0402	.0804	.1208	.1615	.2025	.2440	.2861	.3288	.3722	.4166	.4621	.5087	.5568	.6064	.6580	.7119	.7688	.8296	.8971
-1.6	.0390	.0781	.1174	.1570	.1969	.2374	.2785	.3203	.3630	.4068	.4517	.4980	.5459	.5957	.6477	.7024	.7605	.8231	.8930
-1.4	.0376	.0754	.1133	.1516	.1903	.2295	.2694	.3102	.3519	.3948	.4390	.4849	.5325	.5823	.6347	.6903	.7497	.8144	.8874
-1.2	.0361	.0723	.1087	.1454	.1827	.2205	.2591	.2986	.3391	.3810	.4243	.4694	.5166	.5662	.6189	.6752	.7361	.8033	.8801
-1.0	.0344	.0689	.1036	.1388	.1744	.2106	.2477	.2857	.3249	.3655	.4077	.4519	.4984	.5478	.6005	.6575	.7199	.7897	.8709
-0.8	.0326	.0653	.0983	.1317	.1656	.2001	.2355	.2720	.3097	.3488	.3898	.4328	.4785	.5272	.5799	.6373	.7011	.7735	.8597
-0.6	.0308	.0617	.0929	.1244	.1565	.1893	.2230	.2577	.2938	.3314	.3709	.4127	.4572	.5052	.5574	.6151	.6800	.7551	.8465
-0.4	.0290	.0580	.0874	.1171	.1474	.1784	.2103	.2433	.2777	.3136	.3516	.3919	.4352	.4822	.5337	.5914	.6572	.7346	.8314
-0.2	.0272	.0545	.0820	.1100	.1385	.1677	.1979	.2291	.2617	.2960	.3323	.3711	.4129	.4587	.5094	.5667	.6330	.7125	.8145
0.0	.0255	.0510	.0769	.1031	.1299	.1574	.1858	.2153	.2461	.2787	.3133	.3505	.3909	.4353	.4849	.5415	.6080	.6893	.7963
0.2	.0238	.0477	.0719	.0965	.1217	.1475	.1742	.2020	.2312	.2621	.2950	.3306	.3693	.4123	.4606	.5164	.5827	.6652	.7768
0.4	.0223	.0447	.0673	.0904	.1139	.1382	.1633	.1895	.2171	.2463	.2776	.3115	.3486	.3900	.4370	.4917	.5575	.6408	.7564
0.6	.0208	.0418	.0630	.0846	.1067	.1295	.1531	.1778	.2038	.2314	.2611	.2934	.3289	.3687	.4142	.4676	.5326	.6163	.7354
0.8	.0195	.0392	.0590	.0793	.1000	.1214	.1436	.1669	.1914	.2176	.2457	.2764	.3104	.3486	.3925	.4445	.5085	.5922	.7141
1.0	.0183	.0367	.0554	.0744	.0938	.1139	.1348	.1567	.1799	.2046	.2313	.2605	.2929	.3296	.3719	.4224	.4852	.5685	.6926
1.2	.0172	.0345	.0520	.0698	.0881	.1071	.1267	.1474	.1693	.1927	.2180	.2458	.2767	.3118	.3526	.4015	.4630	.5455	.6713
1.4	.0162	.0324	.0489	.0657	.0829	.1008	.1193	.1388	.1595	.1817	.2057	.2321	.2616	.2952	.3344	.3818	.4418	.5234	.6503
1.6	.0152	.0305	.0461	.0619	.0782	.0950	.1125	.1310	.1505	.1716	.1944	.2195	.2476	.2798	.3175	.3633	.4218	.5022	.6296
1.8	.0144	.0288	.0435	.0584	.0738	.0897	.1063	.1238	.1423	.1622	.1839	.2078	.2347	.2655	.3018	.3460	.4029	.4819	.6095
2.0	.0136	.0273	.0411	.0553	.0698	.0849	.1006	.1171	.1347	.1537	.1743	.1971	.2228	.2522	.2871	.3298	.3851	.4627	.5900
2.2	.0129	.0259	.0390	.0524	.0662	.0805	.0954	.1111	.1278	.1458	.1655	.1872	.2117	.2400	.2735	.3147	.3684	.4444	.5711
2.4	.0123	.0246	.0370	.0498	.0628	.0764	.0906	.1055	.1214	.1386	.1573	.1781	.2015	.2286	.2608	.3006	.3527	.4271	.5529
2.6	.0117	.0234	.0352	.0473	.0597	.0727	.0861	.1004	.1155	.1319	.1498	.1696	.1921	.2181	.2491	.2875	.3380	.4107	.5354
2.8	.0111	.0223	.0335	.0450	.0569	.0692	.0820	.0956	.1101	.1257	.1428	.1618	.1833	.2083	.2381	.2752	.3242	.3952	.5185
3.0	.0106	.0212	.0319	.0429	.0542	.0660	.0782	.0912	.1050	.1200	.1363	.1545	.1752	.1991	.2278	.2637	.3112	.3805	.5023

$E(r_S)$ n = 15

a \ ρ	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
-∞	.0448	.0896	.1345	.1795	.2246	.2700	.3156	.3616	.4079	.4546	.5018	.5496	.5981	.6474	.6977	.7493	.8024	.8580	.9177
-2.0	.0424	.0849	.1275	.1703	.2135	.2570	.3010	.3456	.3908	.4368	.4836	.5315	.5804	.6306	.6823	.7358	.7914	.8498	.9130
-1.8	.0414	.0829	.1246	.1665	.2088	.2515	.2948	.3386	.3833	.4288	.4753	.5230	.5719	.6224	.6745	.7288	.7854	.8453	.9103
-1.6	.0402	.0805	.1210	.1618	.2030	.2447	.2869	.3300	.3738	.4187	.4647	.5120	.5609	.6115	.6642	.7193	.7772	.8390	.9064
-1.4	.0388	.0777	.1168	.1563	.1961	.2365	.2776	.3196	.3624	.4064	.4517	.4986	.5472	.5979	.6510	.7070	.7665	.8305	.9012
-1.2	.0372	.0745	.1121	.1500	.1883	.2273	.2670	.3076	.3493	.3922	.4366	.4828	.5310	.5816	.6350	.6919	.7530	.8196	.8942
-1.0	.0355	.0711	.1069	.1431	.1798	.2171	.2553	.2944	.3347	.3763	.4197	.4649	.5125	.5628	.6164	.6740	.7368	.8061	.8854
-0.8	.0336	.0674	.1014	.1358	.1707	.2063	.2428	.2803	.3190	.3592	.4012	.4454	.4921	.5419	.5954	.6536	.7179	.7902	.8746
-0.6	.0317	.0636	.0958	.1283	.1614	.1952	.2298	.2656	.3027	.3413	.3819	.4247	.4703	.5193	.5726	.6311	.6967	.7718	.8618
-0.4	.0299	.0599	.0901	.1208	.1520	.1840	.2168	.2508	.2861	.3231	.3621	.4034	.4478	.4958	.5485	.6071	.6736	.7514	.8470
-0.2	.0280	.0562	.0846	.1134	.1428	.1729	.2040	.2361	.2697	.3049	.3422	.3821	.4250	.4718	.5236	.5820	.6492	.7293	.8305
0.0	.0262	.0526	.0793	.1063	.1339	.1622	.1915	.2219	.2537	.2872	.3228	.3609	.4024	.4478	.4986	.5564	.6239	.7059	.8125
0.2	.0246	.0492	.0742	.0995	.1254	.1521	.1796	.2082	.2383	.2701	.3040	.3405	.3803	.4243	.4738	.5307	.5982	.6816	.7932
0.4	.0230	.0461	.0694	.0932	.1175	.1424	.1683	.1954	.2237	.2538	.2860	.3208	.3590	.4015	.4496	.5055	.5725	.6569	.7729
0.6	.0215	.0431	.0650	.0872	.1100	.1335	.1578	.1833	.2100	.2385	.2691	.3023	.3388	.3796	.4263	.4809	.5472	.6322	.7519
0.8	.0201	.0404	.0609	.0817	.1031	.1252	.1481	.1720	.1973	.2242	.2532	.2848	.3197	.3589	.4040	.4572	.5226	.6077	.7306
1.0	.0189	.0379	.0571	.0767	.0967	.1175	.1390	.1616	.1855	.2109	.2384	.2685	.3018	.3394	.3829	.4347	.4989	.5837	.7091
1.2	.0177	.0356	.0536	.0720	.0909	.1104	.1307	.1520	.1745	.1986	.2247	.2533	.2851	.3211	.3630	.4132	.4761	.5603	.6876
1.4	.0167	.0334	.0504	.0677	.0855	.1039	.1230	.1431	.1645	.1873	.2120	.2392	.2695	.3041	.3444	.3930	.4545	.5377	.6664
1.6	.0157	.0315	.0475	.0638	.0806	.0979	.1160	.1350	.1552	.1769	.2003	.2262	.2552	.2882	.3270	.3740	.4339	.5161	.6455
1.8	.0148	.0297	.0449	.0603	.0761	.0925	.1096	.1276	.1467	.1672	.1896	.2142	.2419	.2735	.3108	.3563	.4146	.4954	.6251
2.0	.0140	.0281	.0424	.0570	.0720	.0875	.1037	.1208	.1389	.1584	.1797	.2032	.2296	.2599	.2957	.3397	.3963	.4757	.6054
2.2	.0133	.0267	.0402	.0540	.0682	.0829	.0983	.1145	.1318	.1503	.1706	.1930	.2182	.2473	.2817	.3241	.3792	.4570	.5862
2.4	.0127	.0253	.0382	.0513	.0648	.0788	.0933	.1088	.1252	.1429	.1622	.1836	.2077	.2356	.2687	.3096	.3631	.4393	.5677
2.6	.0120	.0241	.0363	.0488	.0616	.0749	.0888	.1035	.1191	.1360	.1544	.1748	.1980	.2247	.2566	.2961	.3480	.4225	.5498
2.8	.0114	.0229	.0345	.0464	.0586	.0713	.0845	.0985	.1135	.1296	.1472	.1668	.1889	.2146	.2453	.2834	.3338	.4066	.5326
3.0	.0108	.0218	.0329	.0442	.0558	.0679	.0806	.0939	.1082	.1236	.1405	.1592	.1805	.2052	.2347	.2716	.3204	.3915	.5161

$E(r_S)$ n = 20

a \ ρ	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
−∞	.0455	.0910	.1366	.1823	.2281	.2742	.3205	.3670	.4140	.4613	.5090	.5573	.6062	.6559	.7064	.7580	.8110	.8659	.9242
-2.0	.0431	.0862	.1295	.1730	.2168	.2610	.3056	.3508	.3967	.4432	.4906	.5390	.5884	.6390	.6910	.7446	.8001	.8580	.9198
-1.8	.0421	.0842	.1266	.1691	.2121	.2554	.2993	.3438	.3890	.4352	.4822	.5304	.5799	.6307	.6832	.7376	.7942	.8535	.9172
-1.6	.0408	.0818	.1229	.1644	.2062	.2485	.2914	.3350	.3795	.4249	.4715	.5194	.5687	.6198	.6728	.7281	.7860	.8473	.9135
-1.4	.0394	.0789	.1187	.1587	.1992	.2402	.2819	.3245	.3679	.4125	.4534	.5058	.5549	.6061	.6596	.7159	.7753	.8389	.9084
-1.2	.0378	.0757	.1138	.1523	.1913	.2308	.2711	.3123	.3546	.3981	.4431	.4898	.5385	.5896	.6435	.7007	.7619	.8281	.9016
-1.0	.0360	.0722	.1086	.1453	.1826	.2205	.2592	.2989	.3398	.3820	.4259	.4717	.5198	.5706	.6247	.6827	.7456	.8148	.8930
-0.8	.0342	.0685	.1030	.1379	.1734	.2095	.2465	.2846	.3239	.3647	.4072	.4519	.4992	.5495	.6035	.6622	.7267	.7989	.8824
-0.6	.0323	.0646	.0973	.1303	.1639	.1982	.2334	.2697	.3073	.3465	.3876	.4310	.4772	.5268	.5805	.6396	.7054	.7806	.8698
-0.4	.0303	.0608	.0915	.1227	.1544	.1869	.2202	.2547	.2905	.3280	.3675	.4095	.4544	.5030	.5562	.6153	.6823	.7602	.8552
-0.2	.0285	.0571	.0859	.1152	.1451	.1756	.2071	.2398	.2739	.3096	.3474	.3878	.4313	.4787	.5311	.5900	.6577	.7381	.8389
0.0	.0267	.0534	.0805	.1080	.1360	.1648	.1945	.2253	.2576	.2916	.3277	.3664	.4084	.4544	.5058	.5641	.6322	.7146	.8210
0.2	.0249	.0500	.0754	.1011	.1274	.1544	.1824	.2115	.2420	.2743	.3086	.3456	.3860	.4306	.4807	.5382	.6063	.6902	.8017
0.4	.0233	.0468	.0705	.0947	.1193	.1447	.1710	.1984	.2272	.2578	.2904	.3258	.3644	.4075	.4562	.5127	.5804	.6654	.7815
0.6	.0218	.0438	.0660	.0886	.1118	.1356	.1603	.1861	.2133	.2422	.2732	.3069	.3439	.3853	.4326	.4878	.5549	.6405	.7606
0.8	.0205	.0410	.0619	.0830	.1047	.1271	.1504	.1747	.2004	.2277	.2571	.2892	.3246	.3643	.4100	.4639	.5300	.6158	.7392
1.0	.0192	.0385	.0580	.0779	.0983	.1193	.1412	.1641	.1884	.2142	.2421	.2726	.3064	.3446	.3886	.4411	.5060	.5916	.7177
1.2	.0180	.0361	.0545	.0731	.0923	.1121	.1327	.1544	.1773	.2017	.2282	.2572	.2895	.3260	.3685	.4194	.4830	.5680	.6961
1.4	.0169	.0340	.0512	.0688	.0869	.1055	.1250	.1454	.1670	.1902	.2153	.2429	.2737	.3087	.3496	.3989	.4611	.5452	.6748
1.6	.0159	.0320	.0483	.0648	.0819	.0995	.1179	.1372	.1576	.1796	.2035	.2297	.2591	.2927	.3320	.3797	.4403	.5234	.6538
1.8	.0151	.0302	.0456	.0612	.0773	.0939	.1113	.1296	.1490	.1699	.1925	.2176	.2456	.2778	.3156	.3616	.4207	.5025	.6333
2.0	.0143	.0286	.0431	.0579	.0731	.0889	.1053	.1227	.1411	.1609	.1825	.2063	.2331	.2639	.3003	.3448	.4022	.4826	.6134
2.2	.0135	.0271	.0408	.0549	.0693	.0843	.0999	.1163	.1338	.1527	.1732	.1960	.2216	.2511	.2861	.3291	.3849	.4636	.5941
2.4	.0129	.0257	.0388	.0521	.0658	.0800	.0948	.1105	.1271	.1451	.1647	.1864	.2109	.2392	.2728	.3144	.3685	.4457	.5754
2.6	.0122	.0244	.0368	.0495	.0625	.0760	.0902	.1051	.1210	.1381	.1568	.1776	.2010	.2282	.2605	.3006	.3532	.4287	.5574
2.8	.0116	.0232	.0350	.0471	.0595	.0724	.0858	.1001	.1152	.1316	.1495	.1693	.1918	.2179	.2491	.2878	.3388	.4125	.5400
3.0	.0110	.0221	.0333	.0448	.0567	.0690	.0818	.0954	.1099	.1255	.1426	.1617	.1833	.2083	.2383	.2757	.3252	.3973	.5233

$E(r_S)$ n = 25

a \ ρ	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
∞	.0459	.0919	.1379	.1840	.2303	.2768	.3235	.3704	.4177	.4654	.5134	.5620	.6112	.6611	.7117	.7634	.[illegible]162	.8708	.9283
-2.0	.0435	.0870	.1307	.1746	.2188	.2634	.3085	.3541	.4003	.4472	.4950	.5436	.5933	.6442	.6964	.7500	.8054	.8630	.9239
-1.8	.0425	.0850	.1278	.1707	.2141	.2578	.3021	.3470	.3926	.4391	.4865	.5350	.5847	.6358	.6885	.7430	.7995	.8586	.9214
-1.6	.0412	.0826	.1241	.1659	.2081	.2508	.2941	.3381	.3829	.4288	.4757	.5239	.5736	.6249	.6781	.7335	.7914	.8524	.9178
-1.4	.0398	.0797	.1198	.1602	.2011	.2425	.2846	.3275	.3713	.4163	.4625	.5102	.5597	.6111	.6648	.7212	.7807	.8441	.9128
-1.2	.0382	.0764	.1149	.1538	.1931	.2330	.2737	.3152	.3579	.4017	.4471	.4941	.5432	.5946	.6487	.7061	.7673	.8334	.9062
-1.0	.0364	.0729	.1096	.1467	.1843	.2226	.2617	.3017	.3429	.3855	.4298	.4759	.5244	.5755	.6298	.6880	.7510	.8201	.8977
-0.8	.0345	.0691	.1040	.1392	.1750	.2115	.2489	.2873	.3269	.3680	.4110	.4560	.5036	.5542	.6086	.6675	.7321	.8043	.8872
-0.6	.0326	.0652	.0982	.1316	.1655	.2001	.2356	.2723	.3102	.3497	.3912	.4349	.4814	.5313	.5854	.6447	.7108	.7860	.8747
-0.4	.0306	.0614	.0924	.1239	.1559	.1886	.2223	.2571	.2932	.3311	.3709	.4132	.4585	.5074	.5609	.6203	.6876	.7656	.8603
-0.2	.0287	.0576	.0867	.1163	.1464	.1773	.2091	.2421	.2764	.3125	.3506	.3913	.4352	.4829	.5357	.5949	.6629	.7435	.8440
0.0	.0269	.0540	.0813	.1090	.1373	.1664	.1963	.2275	.2600	.2943	.3307	.3698	.4121	.4585	.5102	.5669	.6373	.7199	.8262
0.2	.0252	.0505	.0761	.1021	.1286	.1559	.1841	.2135	.2443	.2768	.3115	.3488	.3895	.4344	.4849	.5428	.6113	.6955	.8070
0.4	.0236	.0472	.0712	.0956	.1205	.1461	.1726	.2003	.2294	.2602	.2931	.3288	.3678	.4112	.4602	.5171	.5852	.6706	.7868
0.6	.0221	.0442	.0667	.0895	.1128	.1369	.1618	.1879	.2154	.2445	.2758	.3097	.3471	.3888	.4364	.4921	.5596	.6456	.7659
0.8	.0207	.0414	.0624	.0838	.1058	.1284	.1518	.1764	.2023	.2299	.2596	.2919	.3276	.3677	.4137	.4680	.5346	.6208	.7446
1.0	.0194	.0388	.0586	.0786	.0992	.1205	.1426	.1657	.1902	.2163	.2444	.2752	.3093	.3477	.3922	.4450	.5104	.5965	.7230
1.2	.0182	.0365	.0550	.0739	.0932	.1132	.1340	.1558	.1790	.2036	.2304	.2596	.2922	.3291	.3719	.4231	.4872	.5728	.7014
1.4	.0171	.0343	.0517	.0695	.0877	.1065	.1262	.1468	.1686	.1920	.2174	.2452	.2763	.3116	.3528	.4025	.4651	.5499	.6800
1.6	.0161	.0323	.0487	.0655	.0827	.1004	.1190	.1385	.1591	.1813	.2054	.2319	.2615	.2954	.3351	.3831	.4442	.5279	.6590
1.8	.0152	.0305	.0460	.0618	.0780	.0948	.1124	.1308	.1504	.1715	.1944	.2196	.2479	.2803	.3185	.3650	.4245	.5068	.6384
2.0	.0144	.0288	.0435	.0585	.0738	.0897	.1063	.1238	.1424	.1625	.1842	.2083	.2353	.2664	.3031	.3480	.4059	.4868	.6183
2.2	.0136	.0274	.0412	.0554	.0700	.0850	.1008	.1174	.1351	.1541	.1749	.1978	.2237	.2535	.2887	.3321	.3883	.4677	.5989
2.4	.0130	.0260	.0391	.0526	.0664	.0807	.0957	.1115	.1283	.1465	.1662	.1882	.2129	.2415	.2754	.3173	.3719	.4496	.5802
2.6	.0123	.0247	.0372	.0500	.0631	.0768	.0910	.1061	.1221	.1394	.1583	.1792	.2029	.2303	.2630	.3034	.3564	.4325	.5620
2.8	.0117	.0234	.0354	.0475	.0601	.0731	.0866	.1010	.1163	.1328	.1509	.1709	.1936	.2199	.2514	.2904	.3419	.4162	.5445
3.0	.0110	.0222	.0336	.0452	.0572	.0696	.0826	.0963	.1109	.1267	.1440	.1632	.1850	.2103	.2405	.2782	.3282	.4008	.5277

$E(r_S)$ n = 50

a \ ρ	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
−∞	.0468	.0937	.1406	.1876	.2347	.2821	.3296	.3773	.4254	.4738	.5226	.5718	.6215	.6718	.7227	.7744	.8271	.8609	.9366
-2.0	.0443	.0887	.1333	.1780	.2231	.2685	.3144	.3607	.4077	.4554	.5039	.5532	.6035	.6548	.7073	.7612	.8164	.8734	.9325
-1.8	.0433	.0867	.1303	.1741	.2182	.2628	.3078	.3535	.3999	.4471	.4953	.5444	.5948	.6464	.6995	.7542	.8106	.8691	.9302
-1.6	.0420	.0842	.1265	.1692	.2122	.2556	.2997	.3445	.3901	.4367	.4843	.5332	.5835	.6354	.6890	.7446	.8025	.8630	.9267
-1.4	.0406	.0813	.1222	.1634	.2050	.2472	.2900	.3337	.3783	.4240	.4709	.5194	.5695	.6215	.6757	.7324	.7919	.8548	.9219
-1.2	.0389	.0779	.1172	.1568	.1968	.2375	.2789	.3212	.3646	.4092	.4553	.5030	.5528	.6047	.6594	.7171	.7785	.8442	.9155
-1.0	.0371	.0743	.1118	.1496	.1879	.2269	.2667	.3075	.3494	.3927	.4377	.4845	.5337	.5854	.6403	.6990	.7622	.8311	.9073
-0.8	.0352	.0705	.1060	.1420	.1784	.2156	.2537	.2928	.3331	.3749	.4186	.4643	.5126	.5639	.6189	.6783	.7433	.8153	.8971
-0.6	.0332	.0665	.1001	.1341	.1687	.2040	.2402	.2775	.3161	.3563	.3985	.4429	.4901	.5407	.5955	.6554	.7219	.7971	.8849
-0.4	.0312	.0626	.0942	.1263	.1589	.1923	.2266	.2620	.2988	.3373	.3779	.4209	.4668	.5165	.5707	.6308	.6985	.7768	.8707
-0.2	.0293	.0587	.0884	.1186	.1493	.1807	.2132	.2467	.2817	.3184	.3573	.3986	.4432	.4916	.5451	.6050	.6736	.7546	.8546
0.0	.0274	.0550	.0829	.1112	.1400	.1696	.2001	.2319	.2650	.2999	.3370	.3767	.4197	.4668	.5193	.5787	.6478	.7310	.8370
0.2	.0257	.0515	.0776	.1041	.1311	.1590	.1877	.2176	.2490	.2821	.3174	.3554	.3968	.4424	.4936	.5523	.6215	.7064	.8179
0.4	.0240	.0482	.0726	.0974	.1228	.1489	.1760	.2042	.2338	.2652	.2987	.3350	.3747	.4187	.4686	.5263	.5952	.6814	.7978
0.6	.0225	.0451	.0680	.0912	.1150	.1396	.1650	.1915	.2195	.2492	.2811	.3156	.3536	.3961	.4444	.5009	.5693	.6562	.7769
0.8	.0211	.0422	.0637	.0855	.1078	.1309	.1548	.1798	.2062	.2343	.2645	.2974	.3338	.3746	.4213	.4765	.5439	.6311	.7555
1.0	.0197	.0396	.0597	.0802	.1012	.1228	.1453	.1689	.1938	.2204	.2491	.2804	.3151	.3543	.3995	.4531	.5194	.6065	.7339
1.2	.0185	.0372	.0561	.0753	.0950	.1154	.1366	.1589	.1824	.2076	.2348	.2646	.2977	.3353	.3788	.4309	.4959	.5826	.7122
1.4	.0174	.0350	.0527	.0708	.0894	.1086	.1286	.1496	.1719	.1957	.2216	.2499	.2815	.3175	.3595	.4099	.4735	.5594	.6907
1.6	.0164	.0329	.0497	.0668	.0843	.1024	.1213	.1412	.1622	.1848	.2094	.2364	.2666	.3010	.3414	.3902	.4523	.5371	.6695
1.8	.0155	.0311	.0469	.0630	.0796	.0967	.1146	.1334	.1534	.1748	.1981	.2238	.2527	.2857	.3245	.3718	.4322	.5158	.6488
2.0	.0147	.0294	.0444	.0596	.0753	.0915	.1084	.1263	.1452	.1656	.1878	.2123	.2398	.2715	.3088	.3545	.4133	.4955	.6286
2.2	.0139	.0279	.0420	.0565	.0713	.0867	.1028	.1197	.1377	.1571	.1783	.2016	.2280	.2583	.2942	.3383	.3955	.4761	.6089
2.4	.0132	.0265	.0399	.0536	.0677	.0823	.0976	.1137	.1308	.1493	.1695	.1918	.2170	.2461	.2806	.3232	.3788	.4578	.5900
2.6	.0125	.0251	.0379	.0509	.0643	.0782	.0928	.1081	.1244	.1421	.1613	.1827	.2068	.2347	.2679	.3091	.3630	.4403	.5716
2.8	.0119	.0239	.0360	.0484	.0612	.0745	.0883	.1029	.1185	.1354	.1538	.1742	.1973	.2241	.2561	.2959	.3482	.4238	.5539
3.0	.0112	.0226	.0342	.0461	.0582	.0709	.0841	.0981	.1130	.1291	.1467	.1663	.1885	.2143	.2451	.2835	.3343	.4081	.5368

$E(r_S)$ n = 100

a \ ρ	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
−∞	.0473	.0946	.1420	.1894	.2370	.2848	.3327	.3809	.4294	.4781	.5273	.5768	.6268	.6773	.7284	.7801	.8327	.8862	.9409
-2.0	.0448	.0896	.1345	.1798	.2253	.2711	.3174	.3642	.4116	.4596	.5084	.5581	.6087	.6603	.7130	.7669	.8221	.8787	.9370
-1.8	.0437	.0876	.1315	.1758	.2203	.2653	.3108	.3569	.4037	.4513	.4998	.5493	.6000	.6519	.7051	.7599	.8163	.8744	.9346
-1.6	.0425	.0850	.1278	.1708	.2142	.2581	.3026	.3478	.3938	.4407	.4887	.5380	.5886	.6408	.6946	.7504	.8083	.8684	.9313
-1.4	.0410	.0821	.1233	.1650	.2070	.2496	.2928	.3369	.3819	.4279	.4753	.5241	.5745	.6268	.6812	.7381	.7977	.8603	.9266
-1.2	.0393	.0787	.1183	.1583	.1988	.2398	.2816	.3243	.3681	.4131	.4595	.5076	.5577	.6100	.6649	.7228	.7843	.8498	.9203
-1.0	.0375	.0750	.1129	.1510	.1898	.2291	.2693	.3104	.3527	.3964	.4413	.4890	.5385	.5906	.6458	.7047	.7680	.8367	.9122
-0.8	.0355	.0712	.1071	.1434	.1802	.2177	.2561	.2956	.3363	.3785	.4225	.4686	.5173	.5689	.6242	.6839	.7490	.8210	.9022
-0.6	.0335	.0672	.1011	.1354	.1704	.2060	.2425	.2802	.3191	.3597	.4022	.4470	.4946	.5456	.6006	.6609	.7276	.8029	.8901
-0.4	.0315	.0632	.0952	.1275	.1605	.1942	.2288	.2646	.3017	.3406	.3815	.4248	.4711	.5211	.5757	.6361	.7041	.7625	.8760
-0.2	.0296	.0593	.0893	.1197	.1508	.1825	.2152	.2491	.2844	.3215	.3606	.4024	.4473	.4961	.5500	.6102	.6792	.7603	.8601
0.0	.0277	.0556	.0837	.1122	.1414	.1713	.2021	.2341	.2676	.3028	.3402	.3803	.4236	.4711	.5239	.5838	.6532	.7367	.8425
0.2	.0259	.0520	.0783	.1051	.1324	.1605	.1895	.2198	.2514	.2848	.3205	.3588	.4005	.4465	.4981	.5572	.6268	.7121	.8235
0.4	.0243	.0486	.0733	.0984	.1240	.1504	.1777	.2062	.2361	.2677	.3016	.3382	.3782	.4226	.4729	.5310	.6004	.6869	.8035
0.6	.0227	.0455	.0686	.0921	.1162	.1409	.1666	.1934	.2217	.2516	.2838	.3187	.3570	.3998	.4485	.5055	.5742	.6616	.7826
0.8	.0213	.0427	.0643	.0863	.1089	.1321	.1563	.1816	.2082	.2366	.2671	.3003	.3370	.3781	.4253	.4808	.5487	.6364	.7612
1.0	.0199	.0400	.0603	.0810	.1022	.1240	.1468	.1706	.1957	.2226	.2515	.2831	.3181	.3576	.4032	.4573	.5241	.6117	.7395
1.2	.0187	.0375	.0566	.0760	.0960	.1166	.1380	.1604	.1842	.2096	.2371	.2671	.3006	.3385	.3824	.4349	.5004	.5876	.7178
1.4	.0176	.0353	.0532	.0715	.0903	.1097	.1299	.1511	.1736	.1977	.2237	.2523	.2843	.3205	.3628	.4138	.4779	.5643	.6962
1.6	.0166	.0333	.0502	.0674	.0851	.1034	.1225	.1426	.1638	.1867	.2114	.2386	.2691	.3039	.3446	.3939	.4565	.5419	.6749
1.8	.0156	.0314	.0474	.0636	.0804	.0976	.1157	.1347	.1549	.1765	.2001	.2260	.2551	.2884	.3276	.3753	.4362	.5204	.6541
2.0	.0148	.0297	.0448	.0602	.0760	.0924	.1095	.1275	.1466	.1672	.1896	.2144	.2422	.2741	.3118	.3578	.4172	.4999	.6338
2.2	.0140	.0281	.0424	.0570	.0720	.0876	.1038	.1209	.1391	.1587	.1800	.2036	.2302	.2608	.2970	.3415	.3992	.4804	.6141
2.4	.0133	.0267	.0403	.0541	.0684	.0831	.0985	.1148	.1321	.1508	.1711	.1937	.2191	.2485	.2833	.3263	.3823	.4619	.5950
2.6	.0127	.0254	.0383	.0514	.0650	.0790	.0937	.1092	.1257	.1435	.1629	.1845	.2088	.2370	.2705	.3121	.3665	.4444	.5765
2.8	.0120	.0241	.0364	.0489	.0618	.0752	.0892	.1039	.1197	.1367	.1553	.1759	.1992	.2263	.2586	.2987	.3515	.4277	.5587
3.0	.0113	.0228	.0345	.0465	.0588	.0716	.0849	.0990	.1141	.1303	.1481	.1679	.1903	.2163	.2474	.2862	.3374	.4119	.5415

$E(r_S)$ $n = \infty$

a \ ρ	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
−∞	.0477	.0955	.1434	.1913	.2394	.2876	.3359	.3846	.4334	.4826	.5321	.5819	.6322	.6829	.7341	.7859	.8383	.8915	.9453
−2.0	.0452	.0905	.1359	.1816	.2275	.2738	.3205	.3677	.4155	.4639	.5131	.5631	.6140	.6659	.7188	.7727	.8278	.8841	.9415
−1.8	.0442	.0884	.1329	.1775	.2225	.2679	.3138	.3603	.4075	.4555	.5044	.5543	.6053	.6574	.7109	.7657	.8221	.8799	.9392
−1.6	.0429	.0859	.1291	.1725	.2164	.2607	.3056	.3511	.3976	.4449	.4933	.5429	.5938	.6463	.7004	.7563	.8141	.8740	.9360
−1.4	.0414	.0829	.1246	.1666	.2090	.2520	.2957	.3401	.3855	.4320	.4797	.5288	.5796	.6322	.6869	.7439	.8035	.8659	.9314
−1.2	.0397	.0795	.1195	.1599	.2007	.2422	.2844	.3275	.3715	.4170	.4638	.5123	.5627	.6153	.6705	.7287	.7901	.8555	.9253
−1.0	.0378	.0758	.1140	.1526	.1916	.2314	.2719	.3135	.3562	.4002	.4459	.4935	.5434	.5958	.6513	.7104	.7739	.8425	.9173
−0.8	.0359	.0719	.1081	.1448	.1820	.2199	.2586	.2985	.3395	.3821	.4265	.4730	.5220	.5740	.6296	.6896	.7549	.8268	.9074
−0.6	.0339	.0679	.1021	.1368	.1720	.2080	.2449	.2829	.3222	.3632	.4061	.4512	.4992	.5505	.6059	.6664	.7334	.8087	.8954
−0.4	.0318	.0638	.0961	.1288	.1621	.1961	.2310	.2671	.3047	.3439	.3851	.4288	.4755	.5259	.5808	.6416	.7099	.7884	.8815
−0.2	.0299	.0599	.0902	.1209	.1523	.1843	.2174	.2516	.2872	.3246	.3641	.4062	.4515	.5007	.5549	.6156	.6848	.7661	.8656
0.0	.0280	.0561	.0845	.1134	.1428	.1730	.2041	.2364	.2702	.3058	.3435	.3839	.4276	.4755	.5287	.5889	.6588	.7424	.8481
0.2	.0262	.0525	.0791	.1062	.1337	.1621	.1914	.2219	.2539	.2876	.3236	.3622	.4043	.4507	.5027	.5622	.6322	.7178	.8292
0.4	.0245	.0491	.0740	.0994	.1253	.1519	.1795	.2082	.2384	.2704	.3045	.3414	.3818	.4266	.4773	.5358	.6056	.6926	.8092
0.6	.0229	.0460	.0693	.0930	.1173	.1423	.1683	.1953	.2238	.2541	.2866	.3218	.3604	.4036	.4527	.5101	.5793	.6671	.7883
0.8	.0215	.0431	.0649	.0872	.1100	.1335	.1578	.1834	.2103	.2389	.2697	.3032	.3402	.3817	.4293	.4853	.5536	.6419	.7669
1.0	.0201	.0404	.0609	.0818	.1032	.1253	.1482	.1723	.1977	.2248	.2540	.2859	.3212	.3611	.4070	.4615	.5288	.6170	.7452
1.2	.0189	.0379	.0572	.0768	.0969	.1177	.1393	.1620	.1860	.2117	.2394	.2698	.3035	.3417	.3860	.4390	.5050	.5927	.7234
1.4	.0178	.0357	.0538	.0723	.0912	.1108	.1312	.1526	.1753	.1996	.2259	.2548	.2870	.3236	.3663	.4177	.4823	.5693	.7018
1.6	.0167	.0336	.0507	.0681	.0860	.1044	.1237	.1440	.1654	.1885	.2135	.2410	.2717	.3068	.3479	.3976	.4607	.5467	.6805
1.8	.0158	.0317	.0478	.0643	.0811	.0986	.1168	.1360	.1564	.1783	.2020	.2282	.2576	.2912	.3307	.3788	.4403	.5251	.6596
2.0	.0150	.0300	.0452	.0608	.0768	.0933	.1106	.1288	.1481	.1689	.1915	.2165	.2445	.2768	.3148	.3612	.4211	.5045	.6392
2.2	.0142	.0284	.0429	.0576	.0727	.0884	.1048	.1221	.1405	.1602	.1818	.2056	.2324	.2633	.2999	.3448	.4030	.4849	.6193
2.4	.0135	.0270	.0407	.0546	.0690	.0839	.0995	.1159	.1334	.1522	.1728	.1956	.2212	.2509	.2861	.3294	.3860	.4662	.6001
2.6	.0128	.0256	.0386	.0519	.0656	.0798	.0946	.1102	.1269	.1449	.1645	.1863	.2109	.2393	.2732	.3151	.3699	.4485	.5816
2.8	.0121	.0243	.0367	.0494	.0624	.0759	.0900	.1049	.1208	.1380	.1568	.1776	.2012	.2285	.2611	.3016	.3549	.4317	.5636
3.0	.0114	.0230	.0349	.0469	.0594	.0723	.0857	.1000	.1152	.1316	.1496	.1695	.1922	.2184	.2498	.2889	.3406	.4157	.5463

Appendix B - Numerical Analysis

B.1 Expectations of r_S and r_K

In Chapter III it has been shown that, (3.16) and (3.24),

$$E(r_S) = \frac{12}{n+1}\{R(a,\rho) + (n-2)R^1(a,\rho) - \frac{n-1}{4}\}$$

and

$$E(r_K) = 4\,R(a,\rho) - 1$$

where $R(a,\rho)$ and $R^1(a,\rho)$, (3.4) and (3.12) respectively, have been given as

$$R(a,\rho) = \frac{1}{Q^2(a)}\int_0^\infty Z(u)\,P\left(\frac{\rho u}{\sqrt{1-\rho^2}}\right)Q(u + a\sqrt{2})\,du$$

and

$$R^1(a,\rho) = \frac{1}{Q^3(a)}\int_a^\infty Z(x)dx\int_a^\infty Z(y)Q(y)\,P\{\frac{\rho(x-y)}{\sqrt{2(1-\rho^2)}}\}dydx.$$

Let

$$I_1 = Q^2(a)R(a,\rho) = \int_0^\infty Z(u)\,Q\left(\frac{-\rho u}{\sqrt{1-\rho^2}}\right)Q(u + a\sqrt{2})\,du$$

and

$$I_2 = Q^3(a) R^1(a, \rho) = \int_0^\infty Z(x+a) \int_0^\infty Z(y+a) Q(y+a) Q\{\frac{\rho(y-x)}{\sqrt{2(1-\rho^2)}}\} dy dx,$$

then

$$E(r_K) = \frac{4}{Q^2(a)} I_1 - 1$$

and

$$E(r_S) = \frac{12}{n+1}\{\frac{I_1}{Q^2(a)} + \frac{n-2}{Q^3(a)} I_2 - \frac{n-1}{4}\}.$$

The values of I_1, I_2 and $Q(a)$ were evaluated numerically on the IBM 7040 electronic computer at Virginia Polytechnic Institute for

ρ: 0.05 (0.05) 0.95

a: $-\infty$*, -2.0 (0.2) 3.0

and aggregated with

n: 5, 10, 15, 20, 25, 50, 100, ∞

to form $E(r_S)$ and $E(r_K)$.

---

* a = -6 was actually used for this value.

The numerical integrations of I_1 and I_2 were accomplished using Gauss' arbitrary interval formulae, given as

$$I_1 = \int_0^L f(y)\,dy = \frac{L}{2}\sum_{i=1}^{m} w_i f(y_i)$$

and

$$I_2 = \int_0^L \int_0^L f(x, y)\,dx\,dy = \frac{L^2}{4}\sum_{i=1}^{m}\sum_{j=1}^{m} w_i w_j f(y_i, y_j)$$

where

$$y_i = \frac{L}{2} x_i + \frac{L}{2}$$

and the abscissas x_i and the weights w_i are tabulated by, for example, Abramowitz and Stegun (1964). The range of integration, L, and the number of points taken, m, were determined by the point of truncation, a. The values used were:

a	m	L
-6.0	96	12
-2.0-0.0	96	8
0.2-3.0	64	6.

The Q(x) were evaluated using Hastings' approximation (1955),

$$Q(x) = Z(x)\sum_{i=1}^{5} b_i t^i + \epsilon, \qquad 0 \leq x < \infty,$$

where

$t = (1 + \rho x)^{-1}$ $\rho = .2316419$

$b_1 = .319381530$ $b_4 = -1.821255978$

$b_2 = -.356563782$ $b_5 = 1.330274429$

$b_3 = 1.781477937$ $|\varepsilon| < 7.5 \times 10^{-8}$

and $Q(-x) = 1 - Q(x)$.

Checks on the numerical integration of I_1 and I_2 were made with the special results of $R(a, \rho)$ and $R^1(a, \rho)$ of Sections 3.21 and 3.22, respectively. Also, the non-truncated values of $E(r_K)$ and $E(r_S)$ are well known. Numerical integration for this case, $a = -\infty$, should yield the poorest results since the least number of points per unit interval were used. The accuracy of these values exceeded those reported in Appendix A. Hence, the values of $E(r_K)$ and $E(r_S)$, tabulated in Appendix A, are correct with, of course, the exception of rounding in the last decimal place.

B.2 Evaluation of C(a)

In Chapter II it has been shown, (2.11) and (2.12), that

$$\text{(B.1)} \qquad C(a) = \frac{\sqrt{3}\{\sqrt{2}\, R(a/2) - R(a)\}}{R(a)\sqrt{R^2(a) + a\,R(a) - 1}}$$

$$(B.2) \qquad = \frac{\sqrt{3}\,\{\frac{1}{\sqrt{\pi}}\,Q(a\sqrt{2}) - Z(a)Q(a)}{Q(a)\sqrt{Q^2(a) + aZ(a)Q(a) - Z^2(a)}} .$$

Values of C(a) were determined on the IBM 7040 computer at Virginia Polytechnic Institute for

a: -7.0 (1.0) 15.0, 25, 50 .

C(a) was evaluated using (B. 2) and Hastings' approximation (see previous section) when $a \leq 3$. For $a \geq 2$, C(a) was evaluated using (B. 1) where R(a) is given by the Laplace continued fraction expansion of Mills' ratio (see, for example, Sheppard (1939))

$$R(a) = \frac{1}{a+}\ \frac{1}{a+}\ \frac{2}{a+}\ \frac{3}{a+} \ldots,\quad a > 0.$$

Twenty-five terms of R(a) were used when $a \leq 7$, while ten terms were used in a double precision made when $a \geq 8$. Shenton (1954) gives a discussion of the number of terms of R(a) required to achieve a given accuracy as a function of a. A number of hand evaluations of C(a) were made to ensure the accuracy of the values reported in Section 2. 4.

B. 3 Evaluation of C(a, ρ)

In Chapter II we have shown with (2. 25) that

$$C(a, \rho) = \frac{\sqrt{3}\,\{\frac{1}{\sqrt{\pi}} f_1(a, \rho) - \rho Z(a) f_2(a, \rho)\}}{Q(a)\sqrt{Q^2(a) + \rho^2(aZ(a)Q(a) - Z^2(a))}}$$

where

$$f_1(a, \rho) = L(\frac{a\sqrt{2}}{\sqrt{2-\rho^2}}, \frac{a\sqrt{2}}{\sqrt{2-\rho^2}}; \frac{\rho^2}{2-\rho^2})$$

and

$$f_2(a, \rho) = 2L(a, \frac{a\rho}{\sqrt{2-\rho^2}}; \frac{\rho}{\sqrt{2-\rho^2}}) - Q(a) .$$

Values of $C(a, \rho)$ were determined on the IBM 7040 computer at Virginia Polytechnic Institute for

a: -5.0 (1.0) 6.0

and

ρ: 0.1(0.1) 0.8, 0.85, 0.90 (0.02) 0.98.

Owen (1956) has developed a method for evaluating the volume of a standardized bivariate normal distribution. Modifying his results, we have

$$L(h, k, \rho) = \frac{1}{2} Q(h) + \frac{1}{2} Q(k) - T(h, u_h) - T(k, u_k) - \{ {0 \atop \frac{1}{2}} ,$$

where the choice is 0 if $hk > 0$ or if $hk=0$ and $h+k \geq 0$, and

1/2 otherwise, where

$$T(h, u) = \frac{1}{2\pi} \int_0^u \frac{\exp\{-h^2(1+x^2)/2\}}{1+x^2} \, dx,$$

and

$$u_h = \frac{k}{h\sqrt{1-\rho^2}} - \frac{\rho}{\sqrt{1-\rho^2}} \; ; \; u_k = \frac{h}{k\sqrt{1-\rho^2}} - \frac{\rho}{\sqrt{1-\rho^2}} \; .$$

Special cases of $T(h, u)$ are:

$$T(h, \; u > 1) = \frac{1}{2}Q(h) + Q(hu)[\frac{1}{2} - Q(h)] - T(uh, \frac{1}{u})$$

$$T(0, u) = \frac{1}{2\pi} \arctan u$$

$$T(h, 1) = \frac{1}{2} Q(h)[1-Q(h)]$$

$$T(h, -u) = -T(h, u)$$

$$T(-h, u) = T(h, u)$$

$$T(h, 0) = 0$$

$$T(h, \infty) = \frac{1}{2} P(h) \text{ if } h \geq 0$$
$$= \frac{1}{2} Q(h) \text{ if } h < 0 \; .$$

This method was used to evaluate the L-functions of $f_1(a, \rho)$ and $f_2(a, \rho)$. The $T(h, u)$ integrals were evaluated using twenty points with Gauss' arbitrary interval formulae. It will be noted that the ranges on the integrals were always less than unity.

The $Q(t)$ were evaluated using

$$Q(t) = \frac{1}{2} - Z(t)\,\overline{R}(t) \quad \text{for } 0 \leq t \leq 3$$

and

$$Q(t) = Z(t)R(t) \quad \text{for } t > 3 \,;$$

also $Q(-t) = 1 - Q(t)$. $\overline{R}(t)$, the continued fraction of Shenton (1954) is given by

$$\overline{R}(t) = \frac{t}{1-}\,\frac{t^2}{3+}\,\frac{2t^2}{5-}\,\frac{3t^2}{7+}\cdots, \quad t \geq 0,$$

and converges rapidly for small t. Twenty-eight terms of $\overline{R}(t)$ were used while thirty terms of $R(t)$ were used in a double precision mode.

Checks on the numerical evaluation of $C(a, \rho)$ were made with the special results of Section 2.3 and also by tabulated values of the L-functions of $f_1(a, \rho)$ and $f_2(a, \rho)$. These checks all ensured the accuracy of the values reported in Section 2.4.

Bibliography

1. Abramowitz, M. and Stegun, I. A. Editors. (1964) Handbook of Mathematical Functions. National Bureau of Standards.

2. Aitkin, M. A. (1964) "Correlation in a singly truncated bivariate normal distribution". Psychometrika, 29, 263-270.

3. Aitkin, M. A. and Hume, M. W. (1965) "Correlation in a singly truncated bivariate normal distribution II. Rank Correlation". Biometrika, 52, 639-643.

4. Aitkin, M. A. and Hume, M. W. (1966) "Correlation in a singly truncated bivariate normal distribution III. Correlation between ranks and variate values". Biometrika, to be published.

5. Box, G. E. P. and Muller, M. E. (1958) "A note on the generation of random normal deviates". Ann. Math. Statist., 29, 610-611.

6. David, F. N. and Kendall, M. G. (1949) "Tables of symmetric functions, Part I." Biometrika, 36, 431-439.

7. David, F. N. and Mallows, C. L. (1961) "The variance of Spearman's rho in normal samples". Biometrika, 48, 19-28.

8. Esscher, F. (1924) "On a method of determining correlation from the ranks of variates". Skand. Akt., 7, 201.

9. Fieller, E. C., Hartley, H. O., Pearson, E. S. (1957) "Tests for rank correlation coefficients I". Biometrika, 44, 470-481.

10. Fieller, E. C., Pearson, E. S. (1961) "Tests for rank correlation coefficients II". Biometrika, 48, 29-40.

11. Goulden, C. H. (1952) Methods of Statistical Analysis. (second edition), Chapter 10, John Wiley and Sons, New York.

12. Greiner, R. (1909) "Uber das Fehlersystem der Kollektiomasslehre" Zeitschrift fur Mathematik und Physik, 57, 121, 225 and 337.

13. Hastings, Jr., C. (1955) Approximations for Digital Computers. Princeton University Press.

14. Kendall, M. G. (1938) " A new measurement of rank correlation" Biometrika, 30, 81-92.

15. Kendall, M. G. (1962) Rank Correlation Methods. (Third Edition). Charles Griffin, London.

16. Mills, J. P. (1926) "Table of ratio: area to bounding ordinate, for any portion of the normal curve". Biometrika, 18, 395-400.

17. Moran, P. A. P. (1948) "Rank correlation and product-moment correlation". Biometrika, 35, 203-206.

18. Olkin, I. and Tate, R. F. (1961) "Multivariate correlation models with mixed discrete and continuous variables". Ann. Math. Statist., 32(2), 448-465.

19. Owen, D. B. (1956) "Tables for computing bivariate normal probabilities". Ann. Math. Statist., 27, 1075-1090.

20. Pearson, K. (ed.)(1952) Tables for Statisticians and Biometricians II. Cambridge University Press.

21. Shenton, L. R. (1954) "Inequalities for the normal integral including a new continued fraction". Biometrika, 41, 177-189.

22. Sheppard, W. F. (1898) "On the application of the theory of error to cases of normal distributions and normal correlations" Phil. Trans., 192A, 101 et seq.

23. Sheppard, W. F. (1939). The Probability Integral. Cambridge University Press.

24. Soper, H. E., Young, A.W., Cave, B. M., Lee, A. and Pearson, K. (1917) "On the distribution of the correlation coefficient in small samples. Appendix II to the papers of "Student" and R. A. Fisher." Biometrika, 17, 328-413.

25. Spearman, C. (1904) "The proof and measurement of association between two things". Am. Jour. Psych. 15, 88.

26. Stuart, A. (1954) "The correlation between variate-values and ranks in samples from a continuous distribution". Brit. J. Statist. Psych. , 7(1), 37-44.

27. Stuart, A. (1955) "The correlation between variate-values and ranks in samples from distributions having no variance". Brit. J. Statist. Psych., 8, 25-27.

28. Sundrum, R. M. (1953) "Moments of the rank correlation coefficient τ in the general case". Biometrika, 40, 409-420.

Acknowledgements

The author wishes to express his sincere appreciation to Professor Murray A. Aitkin for suggesting this research and for his guidance during the course of its completion. Co-authored papers on Chapters II and III are to appear in Biometrika* and the work of Chapter IV is being prepared for publication.

Gratitude is extended to Dr. Donald R. Jensen, who kindly consented to read this dissertation in manuscript form. His general comments have added greatly to this dissertation.

Appreciation is expressed to Dr. Boyd Harshbarger for his general assistance and encouragement to the graduate student.

To Dottie Gladstone, the author is indebted for the preparation of the final typed copies.

The author wishes, also, to express his gratitude for the financial assistance from the National Institutes of Health.

*see, Aitkin and Hume in the bibliography

Vita

Merril Wayne Hume was born in San Francisco, California, on April 30, 1939. Being from a military family, his primary and secondary schooling was gained throughout the world. He entered the College of William and Mary in 1956, majored in Physics, and received a B. S. degree in 1960. Upon graduation he was employed as a research assistant for Technical Operations, Inc., Fort Monroe, Virginia.

The author entered the Department of Statistics at Virginia Polytechnic Institute in September 1961. In the fall of 1963, he married Miss Nancy Elizabeth Johnson, whom he met at college. In June of 1963, he was awarded an M. S. degree in Statistics. With the completion of his Ph. D., he plans to remain temporarily at Virginia Polytechnic Institute as an Instructor in the Department of Statistics.

Merril W. Hume

ABSTRACT

Considerable attention has been devoted to the rank correlation coefficients of Spearman and Kendall, denoted by r_S and r_K respectively. These coefficients were first proposed as measures of association between two groupings, requiring no assumptions on the parent distribution of the observations. Later work considered the distributions of r_S and r_K when the parent distribution is the bivariate normal. This study is an investigation of the moments and related properties of Spearman's r_S and Kendall's r_K when the underlying distribution is the singly truncated bivariate normal.

The singly truncated bivariate normal distribution arises in many situations, notably in acceptance sampling and examination selection. In both situations two evaluations (measurements or test scores), which are assumed to follow the bivariate normal distribution, are considered for each individual. The first test determines that an individual is accepted if its score exceeds a "cutting score", otherwise rejected. At some later time the second evaluation is made on accepted individuals. Hence, the pairs of observations determined from accepted individuals may be considered as arising from a singly truncated bivariate normal distribution.

It is often of interest to estimate the underlying parent correlation of the paired observations in the above situation. The rank correlation coefficients r_S and r_K serve this purpose with two strong justifications for their use. First, as ranks are substituted for variate values, they are computationally simple. Second, there are situations where it is impossible or not practical to obtain quantitative measurements on one or both variables but rankings can be made.

This study of the rank correlation coefficients of Spearman and Kendall closely parallels previous considerations of these coefficients arising from a non-truncated bivariate normal distribution. However, it is evident from the complications produced by truncation that analytical results are more difficult to obtain, and, as in the non-truncated case, few of the distributional problems of these coefficients can be fully investigated.

The first investigation of this study is the determination of the correlation between ranks and variate values for the two marginal distributions of the singly truncated bivariate normal distribution. It is generally the case that if this correlation is high, conclusions drawn from the ranks will be similar to those drawn from the variate values. The correlation between ranks and variate values, for all practical purposes, was found to be little affected by truncation for both marginal distributions.

The next investigation is the determination of the expectations of the rank correlation coefficients. These results are obtained in terms of integrals which have been evaluated numerically for selected values of parent correlation, point of truncation and sample size.

The variances of the rank correlation coefficients r_S and r_K proved to be analytically intractable. The first four moments of the two coefficients were determined empirically with the use of an electronic computer and Monte Carlo methods. This empirical study was conducted primarily to obtain accurate estimates of the variances, however, the higher moments indicate the effects of truncation on the distributional forms of r_S and r_K. With knowledge of the variances, a comparison of the coefficients of variation of r_S and r_K was made to indicate which of the rank correlation coefficients best serves as an estimator of parent correlation.