519.2

MEDEDELINGEN VAN DE LANDBOUWHOGESCHOOL TE WAGENINGEN
NEDERLAND 61 (12) 1-19 (1961)

RANDOM VARIABLES RATHER THAN DISTRIBUTION FUNCTIONS, A DIFFERENT AND USEFUL APPROACH IN STATISTICS

by

NICOLAAS H. KUIPER

Department of Mathematics, Agricultural University, Wageningen,[2] *Netherlands*

(Received 27.7.'61)

The role of the notions random variable, random vector, random point etc. in statistical theory and applications, has been fairly modest so far, in papers, books, as well as in textbooks. In particular a tendency seems to exist, to replace these notions and words in every case as soon as possible by probability distributions.

Our main objective in this paper is to demonstrate the advantages obtained in case we consciously and on purpose not only mention random variables etc., but even work and do calculations with them.

With this in mind we introduce three innovations:

I. We use *a symbol* $\cong$ for the relation, called *isomory*, between two random variables (vectors), which relation consists in equality of their cumulative distribution (c.d.) functions.

II. We make frequent use of *functions of random variables:* If f is a real function, $\underline{x}$ a random variable, then a random variable $f(\underline{x})$ is defined.[1]

III. We also consider *empirical random variables* [8] and we even represent them by symbols. Hence we will for example talk about *the unknown* (empirical) *random variable* $\underline{x}$. The application of statistics often consists in the confrontation of an empirical random variable $\underline{x}$ and a theoretical random variable $\underline{y}$, by virtue of a value x which $\underline{x}$ has assumed in some experiment. For example it may happen that $\underline{y} = \underline{\chi}_f^2$ and that the null-hypothesis $\underline{x} \cong \underline{y} = \underline{\chi}_f^2$ must be tested.

In *chapter 1* we illustrate our point with *definitions and examples.* See also [8].

In *chapter 2 a standard representation for any random vector* is applied to normal random vectors, and we indicate our approach to the *analysis of variance theory* (FISHER). We give in particular the fundamental lemma for the analysis of variance. See also [2,8].

In *chapter 3*, we study correlation coefficients with our methods. For the

[1] Random variables, random vectors, random points, etc. will be recognised by a bar under.
[2] The author is now (8.10.'62) connected with the University of Amsterdam.

classical approach compare ANDERSON [1] chapter 4. A survey of the results in chapter 3 is given in section 3.1.

My thanks are due to M. KEULS for help with the approximations in sections 3.5. and 3.8., and to M. KEULS and L. C. A. CORSTEN for valuable critical remarks.

Remark I. W. KRUSKAL drew my attention to a paper by G. ELFVING, A simple method of deducing certain distributions connected with multivariate sampling, Skandinavisk Aktuarie tidskrift XXX (1947), p. 56–74, in which formula A ch. 3 of this paper is given.

Remark II. R. A. WIJSMAN drew my attention to his paper, Applications of a certain representation of the Wishart matrix, Annals of Math. Stat. 30 (1959), p. 597–601, in which a formula equivalent to formula B ch. 3 is given.

1. DEFINITIONS AND EXAMPLES

1.1. *The symbol $\cong$ for isomorous*

$\mathbf{R}^n$ is the n-dimensional euclidean vectorspace of sequences $a = (a_1, \ldots a_n)$, $b = (b_1, \ldots b_n)$ etc. of n real numbers, with inner product $ab = \Sigma_{i=1}^{n} a_i b_i$. The length of a is $|a| = \sqrt{aa}$. A probability distribution over $\mathbf{R}^n$ determines a theoretical random vector. For $n = 1$ it is a random variable. Following D. VAN DANTZIG two random vectors $\underline{x}$ and $\underline{y}$ are called *isomorous* if they have the same cumulative distribution (c.d.) functions. We use the symbol $\cong$ for this equivalence relation:

$$\underline{x} \cong \underline{y}$$

1.2. *Functions of $\underline{x}$*

If f is a measurable function from $\mathbf{R}^n$ to $\mathbf{R}^p$, and $\underline{x}$ is a random vector in $\mathbf{R}^n$, then $\underline{y} = f(\underline{x})$ *is defined to be* the random vector in $\mathbf{R}^p$ with probability for any Borel set $I \in \mathbf{R}^p$:

$$P(\underline{y} \in I) = P[f(\underline{x}) \in I] = P[\underline{x} \in f^{-1}(I)]$$

It is clear that in this case

$$\underline{x} \cong \underline{y} \text{ implies } f(\underline{x}) \cong f(\underline{y})$$

If f and g are two measurable functions, then the random vectors $\underline{y} = f(\underline{x})$ and $\underline{z} = g(\underline{x})$ may be stochastically dependent (for example if $f = g$). If in the sequel we consider two random vectors, $\underline{y}$ and $\underline{z}$, then stochastic dependence by virtue of some 'background' $\underline{x}$ will always be allowed, if it is not specifically excluded. In particular we occasionally meet the relation *equality*, say $\underline{y} = \underline{z}$. This means, also for empirical random variables, that $P(\underline{y} = \underline{z}) = 1$.

1.3. If c is a constant vector, then *the random vector c* is by definition the random vector $\underline{x}$ with $P(\underline{x} = c) = 1$.

1.4. *Limits**

A *sequence of random vectors* $\underline{v}_{(1)}, \underline{v}_{(2)}, \ldots$ has the random vector $\underline{v}$ as a limit*, if the sequence of c.d. functions of $\underline{v}_{(1)}, \underline{v}_{(2)}, \ldots$ converges, for almost every value in the domain, to the c.d. function of $\underline{v}$. Notation:

$$\lim{}^*_{k=\infty} \underline{v}_{(k)} \cong \underline{v}$$

1.5. *Direct sum* $\underline{u} \oplus \underline{v}$. If $\underline{u} = (\underline{u}_1, \ldots \underline{u}_p)$ and $\underline{v} = (\underline{v}_1, \ldots \underline{v}_q)$ are random vectors then $(\underline{u}, \underline{v})$ is the random vector $(\underline{u}_1, \ldots \underline{u}_p, \underline{v}_1, \ldots \underline{v}_q)$ and the direct sum $\underline{u} \oplus \underline{v}$ is the random vector $\underline{w} = (\underline{w}_1, \ldots \underline{w}_{p+q})$ with c.d. function

$$P(\underline{w}_1 \leqslant w_1, ..\underline{w}_{p+q} \leqslant w_{p+q}) =$$
$$= P(\underline{u}_1 \leqslant w_1, ..\underline{u}_p \leqslant w_p) \cdot P(\underline{v}_1 \leqslant w_{p+1}, ..\underline{v}_q \leqslant w_{p+q})$$

$\underline{u}$ and $\underline{v}$ are *stochastically independent* if and only if

$$(\underline{u}, \underline{v}) \cong \underline{u} \oplus \underline{v}$$

1.6. The *standard normal random variable* with probability density $(2\pi)^{-\frac{1}{2}} \exp(-\frac{1}{2} t^2)$ will be denoted by $\underline{\chi}_1$, or short $\underline{\chi}$.

1.7. *Symmetry*

$$\underline{x} \cong -\underline{x} \quad \text{(random variable)}$$

means that the random variable $\underline{x}$, which may be empirical or theoretical, is symmetric with respect to 0. Consequently $\underline{\chi} \cong -\underline{\chi}$ is a well known and true statement concerning the standard normal random variable $\underline{\chi}$.

1.8. *General normal random variable*

$$\underline{x} \cong \mu + \sigma\underline{\chi}$$

says that some (empirical) random variable $\underline{x}$ has the normal distribution with expectation μ and variance σ^2.

1.9. *The central limit theorem for a random variable*

If $\underline{x}_1$, $\underline{x}_2$, etc. are stochastically totally independent and mutually isomorous random variables with finite expectation μ and finite variance σ^2, then (weak law of large numbers)

$$\lim{}^*_{n=\infty} \frac{\Sigma_{i=1}^{n} \underline{x}_i}{n} \cong \mu$$

and (central limit theorem)

$$\lim{}^*_{n=\infty} \frac{\Sigma_{i=1}^{n} (\underline{x}_i - \mu)}{\sqrt{n}} \cong \sigma\,\underline{\chi}$$

1.10. *The standard normal random vector* $\underline{\chi}_n$ is defined by induction and starting from $\underline{\chi}_1 = \underline{\chi}$ by:

$$\underline{\chi}_{k+1} \cong \underline{\chi}_k \oplus \underline{\chi}_1 \qquad k \geqslant 1$$

The inner product of $\underline{\chi}_f$ with itself is $\underline{\chi}_f^2$, the '*Chi-square*' *random variable with dimension f*.

The *noncentral Chi-square random variable* with dimension f and excentricity $|\gamma|$ is

$$(\underline{\chi}_f^{\gamma})^2 \cong [\underline{\chi}_{f-1} \oplus (\underline{\chi} + \gamma)]^2$$

STUDENT's and FISHER's noncentral random variables are:

$$\underline{t}_f^{\gamma} \cong \frac{\underline{\chi} + \gamma}{\sqrt{\underline{\chi}_f^2/f}}$$

and
$$\underline{F}_f^{p,\gamma} \cong \frac{(\underline{\chi}_p{}^{\gamma})^2/p}{(\underline{\chi}_f{}^2)/f}$$

In these formulas $\underline{\chi}$, $\underline{\chi}_f$ and $\underline{\chi}_p{}^{\gamma}$ are assumed stochastically independent. (Compare 3.8. for some other definitions).

1.11. *Some approximations*

KEULS [7] gave the following formula for an *approximation* by PATNAIK:

$$\underline{F}_q^{p,\gamma} \cong \lambda\, \underline{F}_q^{s} \text{ with } \lambda = \frac{p+\gamma^2}{p},\quad s = \frac{(p+\gamma^2)^2}{p+2\gamma^2}.$$

Well known is:

$$\sqrt{2\,\underline{\chi}_n{}^2} \cong \sqrt{2n\text{–}1} + \underline{\chi} \qquad \text{asymptotically}$$

or more precise:

$$\lim{}^*_{n=\infty} \left(\sqrt{2\,\underline{\chi}_n{}^2} - \sqrt{2n\text{–}1}\right) \cong \underline{\chi}$$

2. RANDOM VECTORS AND ANALYSIS OF VARIANCE

2.1. *The standard representation of a random vector* $\underline{y} \in \mathbf{R}^n$

We consider random vectors for which the first and second moments of each component are finite. If $\underline{y}$ has components $y_1, \ldots y_n$ then $\mu = E(\underline{y})$ is the vector with components $E(\underline{y}_1), \ldots E(\underline{y}_n)$. From now on it will be taken for granted that a vector like $\underline{y}$ is *a column of random variables.* We use matrix theory and designate the transposed of a matrix m by adding a prime: m'. The *covariance-matrix* V of $\underline{y}$ is

$$V = E(\underline{y} - \mu)(\underline{y} - \mu)'$$

We assume that V has rank n. This is equivalent to the condition that for no vector $a \in \mathbf{R}^n$ the inner product $a\,\underline{y}$, in matrix notation $a'\,\underline{y}$, is isomorous to a constant. In particular:

$$\sigma_i = \sqrt{E(\underline{y}_i - E\underline{y}_i)^2} > 0.$$

Let N be the $n \times n$-matrix with the numbers σ_i in the main diagonal and zero's elsewhere. Then the *correlation random vector of* $\underline{y}$ is

$$\underline{z} = N^{-1}(\underline{y} - \mu) \qquad \text{(matrix product)}$$

Observe that $E(\underline{z}) = 0$ and the diagonal terms of the *correlation matrix of* $\underline{y}$, which is the covariance matrix of $\underline{z}$,

$$R = E(\underline{z}\,\underline{z}'),$$

are all equal to one. The other elements, $\rho_{ij} = E(\underline{z}_i\,\underline{z}_j)$ for $i \neq j$, of R are the correlation coefficients. Clearly:

$$V = E(N\underline{z})(N\underline{z})' = NE(\underline{z}\,\underline{z}')N' = NRN$$

As V has rank n, so has the symmetric matrix R. Then there exists a unique triangular matrix T (zero's above the main diagonal) with positive elements in the main diagonal, hence with elements $t_{ii} > 0$, $t_{ij} = 0$ for $i < j$, for which: $R = TT'$.

For $n = 2$ one obtains

$$R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},\ T = \begin{pmatrix} 1 & 0 \\ \cos\eta & \sin\eta \end{pmatrix} \qquad \begin{matrix} \rho = \cos\eta, \\ \sin\eta > 0 \end{matrix}$$

The covariance matrix of $\underline{x} = T^{-1} \underline{z}$ is the unit matrix:

$$E\,(T^{-1}\underline{z})\,(T^{-1}\underline{z})' = T^{-1}\,E\,(\underline{z}\,\underline{z}')\,(T^{-1})' = T^{-1}T\,T'\,(T')^{-1} = \mathbf{1}.$$

We can now conclude to

Theorem: [2.1]: *Standard representation of a random vector $\underline{y}$.* If $\underline{y}$ is a random vector in $\mathbf{R}^n$ with correlation matrix R of rank n,[1]) then there exist: (a) a unique random vector $\underline{x}$ with $E\,(\underline{x}) = 0$, $E\,(\underline{x}\,\underline{x}') = 1$; (b) a unique *diagonal matrix* N and a unique *triangular matrix* T, both with positive elements in the diagonal, such that the following standard representation of $\underline{y}$ holds

$$\boxed{\underline{y} = E\,(\underline{y}) + S\,\underline{x} = \mu + NT\,\underline{x}} \qquad (E\,(\underline{y}) = \mu,\ NT = S)$$

$T\underline{x}$ is the *correlation vector* of $\underline{y}$; $TT' = R$ is the *correlation matrix*; $NRN = V$ is the *covariance matrix.*

If ρ is the correlation coefficient in case $n = 2$ and $\eta = \text{arc cos } \rho$, then the standard representation is for $n = 2$:

$$\underline{y} = \begin{pmatrix} \underline{y}_1 \\ \underline{y}_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \cos\eta & \sin\eta \end{pmatrix} \begin{pmatrix} \underline{x}_1 \\ \underline{x}_2 \end{pmatrix}$$

2.2. *The central limit theorem for a random vector*

Let $\underline{y}_{(i)}\ i = 1, 2, \ldots$ be an infinite sequence of totally independent random vectors each isomorous with $\underline{y} = \mu + NT\,\underline{x}$ as in theorem [2.1]. Then (weak law of large numbers)

$$\lim{}^*_{n=\infty} \frac{1}{n} \Sigma_{i=1}^{n}\ \underline{y}_{(i)} \cong \mu$$

This means that the mean is a consistent estimator for μ. Furthermore a form of the central limit theorem is expressed by

$$\lim{}^*_{n=\infty} \frac{1}{\sqrt{n}} \Sigma_{i=1}^{n} (\underline{y}_{(i)} - \mu) \cong N\,T\,\underline{\chi}_n$$

2.3. *The standard normal random vector and rotations*

If $\underline{x}$ is a random vector in $\mathbf{R}^n$ with $E(\underline{x}) = 0$ and $E(\underline{x}\,\underline{x}') = \mathbf{1}$ and if H is a linear transformation (matrix) for which $H\,\underline{x} \cong \underline{x}$ then

$$\mathbf{1} = E(\underline{x}\,\underline{x}') = E[(H\,\underline{x})\,(H\,\underline{x})'] = H(E\,\underline{x}\,\underline{x}')H' = H\,\mathbf{1}\,H' = HH'$$

and H is an orthogonal matrix.

Theorem: The standard normal random vector $\underline{\chi}_n$ has the property: If and only if the matrix H is orthogonal ($HH' = \mathbf{1}$) then

$$H\,\underline{\chi}_n \cong \underline{\chi}_n$$

Proof: The necessity of the orthogonality follows from $E\,(\underline{\chi}_n) = 0$ and $E\,(\underline{\chi}_n\,\underline{\chi}_n') = \mathbf{1}$ in view of the above remark.

The sufficiency can be proved *elementary* (without going into n dimensional integration theory for $n > 2$ as follows: For $n = 2$ sufficiency follows from

[1] If rank $R < n$, then the same representation exists but it is not unique, and some of the diagonal elements of N and/or T are necessarily zero.

the probability density function being invariant under rotations about 0. Then if $n \geqslant 2$ and if H_i is some elementary orthogonal transformation, which acts effectively as a rotation on 2 or 1 variables and is identity on the set of n–2 or n–1 remaining componentvariables, also $H_i \underline{\chi}_n \cong \underline{\chi}_n$. But any orthogonal transformation H can be written for some N in the form

$$H = H_N H_{N-1} \dots H_2 H_1 \text{ with } H_i \text{ elementary for } i = 1, \dots, N$$

Then

$$H \underline{\chi}_n = H_N \dots H_2 H_1 \underline{\chi}_n \cong H_N \dots \underline{\chi}_n \dots \cong H_N \underline{\chi}_n \cong \underline{\chi}_n$$

$$H \underline{\chi}_n \cong \underline{\chi}_n$$

2.4. *Analysis of variance* (FISHER)

In applications one often assumes that some empirical random vector $\underline{y}$, for example the set of random variables which appear in a Latin-square- or other design, has a general normal distribution, or as we want to express it, is isomorous to a general normal random vector. It then admits the standard representation (*Cf*, th. [2.1]):

$$\underline{y} \cong \mu + S \underline{\chi}_n = \mu + N T \underline{\chi}_n$$

The analysis of variance problem of FISHER *is in general: Given some assumptions on the pair* (μ, S) *and given a value* y *of* $\underline{y}$, *to test some hypothesis concerning* μ, S *or to estimate (by point or by confidence region)* (μ, S) *or the value of some function of* (μ, S).

The following case is very common and applies to regression on polynomials, latin squares and other designs: P, Q and F are linear subspaces of $\mathbf{R}^n$, of dimension p, q and f: $Q \subset P$; F is totally orthogonal to P;

It is assumed that

$$\underline{y} \cong \mu + \sigma \underline{\chi}_n, \ \sigma \text{ is an unknown scalar,}$$

and

$$\mu \in P$$

The problem is for example to test, in view of a value y which $\underline{y}$ has assumed in some experiment, the *nullhypothesis:* $\mu \in Q$.

We denote the orthogonal projection of a vector z into a linear subspace A by z_A. Then the nullhypothesis can be expressed by $\mu_P - \mu_Q = 0$ or, if D is the space of all vectors in P that are orthogonal to Q, by:

$$\text{nullhypothesis: } \mu_D = 0$$

The solutions of problems of this kind rests on

The main lemma of the analysis of variance [2.4] (Cf [2, 8]) *Let* $a \in \mathbf{R}^n$ *be a column vector,* D *and* F *totally orthogonal linear subspaces of the vector space* $\mathbf{R}^n$ *of dimensions* d *and* f; $a \in D$. *Then using inner products:*

$$a \underline{\chi}_n \cong |a| \cdot \underline{\chi} \,;$$

$$[(\underline{\chi}_n)_D]^2 \cong \underline{\chi}_d^2 \,;$$

$$[(a + \sigma \underline{\chi}_n)_D]^2 \cong \sigma^2 (\underline{\chi}_d^{\gamma})^2 \text{ with } \gamma = \frac{|a|}{\sigma}$$

and moreover: $(\underline{\chi}_n)_D$ and $(\underline{\chi}_n)_F$ are stochastically independent.

Immediate consequences are

$$\frac{a\,\underline{\chi}_n}{\sqrt{(\underline{\chi}_n)_F^2/f}} \cong |a| \cdot \underline{t}_f\,; \quad \frac{(\underline{\chi}_n)_D^2/d}{(\underline{\chi}_n)_F^2/f} \cong \underline{F}_f^d$$

and

$$\frac{[a + \sigma\,(\underline{\chi}_n)_D]^2/d}{\sigma^2\,(\underline{\chi}_n)_F^2/f} \cong \underline{F}_f^{d,\,\gamma}, \quad \sigma > 0, \quad \gamma = \frac{|a|}{\sigma}.$$

Applying the main lemma to the above problem we easily obtain:

$$\underline{y}_D^2 \cong [\mu_D + \sigma\,(\underline{\chi}_n)_D]^2 \cong \sigma^2\,(\underline{\chi}_d^\gamma)^2 \text{ with } \gamma = \frac{|\mu_D|}{\sigma}$$

and (in view of $\mu_F = 0$)

$$\underline{y}_F^2 = [\mu_F + \sigma\,(\underline{\chi}_n)_F]^2 \cong \sigma^2\,\underline{\chi}_f^2$$

These two are stochastically independent and therefore

$$\frac{\underline{y}_D^2/d}{\underline{y}_F^2/f} \cong \underline{F}_f^{d,\,\gamma} \qquad \text{with } \gamma = \frac{|\mu_D|}{\sigma}$$

The nullhypothesis $\mu_D = 0$ or equivalently $\gamma = 0$ is tested with the F-test by a right critical region of values greater then a constant c, such that the significance level has the preassigned value α:

$$P\,(\underline{F}_f^d > c) = \alpha$$

The power of the test is the function of γ:

$$P\,(\underline{F}_f^{d,\,\gamma} > c)$$

3. Correlation coefficients concerning normal random vectors

3.1. *Summary and results*

In 3.2. we define the random variable $\underline{r}_f^\rho$, *the random correlation coefficient with f degrees of freedom and population correlation coefficient* ρ, *and we prove the formula for* $|\rho| < 1$,

(*A*)

$$\frac{\underline{r}}{\sqrt{1-\underline{r}^2}} = \operatorname{cotg} \underline{h} \cong \frac{\underline{\chi} + \operatorname{cotg}\eta \cdot \sqrt{\underline{\chi}_{f+1}^2}}{\sqrt{\underline{\chi}_f^2}}$$

$$\underline{r} = \underline{r}_f^\rho, \qquad \eta = \arccos\rho$$

Of course: $\underline{r} \cong 1$ for $\rho = 1$; $\underline{r} \cong -1$ for $\rho = -1$

We get in particular:

(*A*°) $$\operatorname{cotg} \underline{h} \cong \frac{1}{\sqrt{f}}\,\underline{t}_f \qquad \text{for } \rho = 0$$

We also prove the theorem of Miss Harley [5]: $E\,(\underline{h}) = \eta$.

In 3.3 and 3.4 we prove that the usual random total or partial correlation coefficient is isomorous with $\underline{r}_f^\rho$.

In 3.5 we obtain and study an approximation, which was obtained in a different way by Miss Harley. The approximation is

(A_{app}) $$\operatorname{cotg} \underline{h} \cong \lambda \underline{t}_f^{\delta}$$
with
$$\lambda = \sqrt{\frac{2 + \operatorname{cotg}^2 \eta}{2f}} \text{ and } \delta = \operatorname{cotg} \eta \sqrt{\frac{2f+1}{2 + \operatorname{cotg}^2 \eta}}$$

In 3.6. we define the random variable $\underline{r}_f^{m,\rho}$, *the random m-tuple correlation coefficient with f degrees of freedom and population m-tuple correlation coefficient* ρ, and we prove the formula

(*B*)
$$\frac{\underline{r}^2}{1-\underline{r}^2} = \operatorname{cotg}^2 \underline{h} \cong \frac{(\underline{\chi} + \operatorname{cotg} \eta \cdot \sqrt{\underline{\chi}^2_{f+m}})^2 + \underline{\chi}^2_{m-1}}{\underline{\chi}_f^2}$$
$$\underline{r} = \underline{r}_f^{m,\rho}, \qquad \eta = \arccos \rho$$

In particular:

(B°) $$\operatorname{cotg}^2 \underline{h} \cong \frac{\underline{\chi}_m{}^2}{\underline{\chi}_f{}^2} \cong \frac{m}{f} \underline{F}_f^m \qquad \text{for} \qquad \rho = 0$$

and (see formula *A*):
$$(\underline{r}_f^{1,\rho})^2 = (\underline{r}^{\rho})^2 \qquad \text{for } m = 1$$

In 3.7. we indicate that the usual random total or *p*-partial *m*-tuple correlation coefficient is isomorous with $\underline{r}_f^{m,\rho}$.

In 3.8. we present two approximations of $\underline{r}_f^{m,\rho}$ which may be useful for practical purposes.

Side remark: From formulas (*A*) and (*B*) one sees immediately (see section 1.3, 1.4 and 1.11)

$(A\infty)$ $$\underline{r}_{\infty}^{\rho} = \operatorname{lim}^*_{f=\infty} \underline{r}_f^{\rho} \cong \rho$$

$(B\infty)$ $$\underline{r}_{\infty}^{m,\rho} = \operatorname{lim}^*_{f=\infty} \underline{r}_f^{m,\rho} \cong \rho$$

These formulae express that $\underline{r}_f^{\rho}$ and $\underline{r}_f^{m,\rho}$ for $f = 1, 2, \ldots$ are *consistent sequences of estimators* of ρ.

3.2. *The random variable* $\underline{r}_f^{\rho}$.

We consider a random normal vector in $\mathbf{R}^2$ with the standard representation (See theorem [2.1]):

(3.2.1) $$\underline{z}_0 = \begin{pmatrix} \underline{x}_0 \\ \underline{y}_0 \end{pmatrix} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} + \begin{pmatrix} \sigma_x & 0 \\ 0 & \sigma_y \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \cos \eta & \sin \eta \end{pmatrix} \begin{pmatrix} \underline{u}_0 \\ \underline{v}_0 \end{pmatrix}$$
where $\begin{pmatrix} \underline{u}_0 \\ \underline{v}_0 \end{pmatrix} \cong \underline{\chi}_2, \eta = \arccos \rho,\ |\rho| < 1, 0 < \eta < \pi$

$\underline{x}_0$, $\underline{y}_0$, $\underline{u}_0$ and $\underline{v}_0$ are random variables and $\cos \eta = \rho$ is the (*population*) *correlation-coefficient* of $\underline{x}_0$ and $\underline{y}_0$. The standard normal random variables $\underline{u}_0$ and $\underline{v}_0$ are independent. Observe that we use the symbol $=$ and not $\cong$ in (3.2.1) (See section 1.2).

(3.2.1) can be written:

(3.2.2) $$\begin{cases} \underline{x}_0 = \mu_x + \sigma_x \underline{u}_0 \\ \underline{y}_0 = \mu_y + \sigma_y (\cos\eta \cdot \underline{u}_0 + \sin\eta \cdot \underline{v}_0) \end{cases}$$

In order to give our definition of the theoretical random variable $\underline{r}_f^\rho$, we restrict to the case $\mu_x = \mu_y = 0$:

(3.2.3) $$\begin{cases} \underline{x}_0 = \sigma_x \underline{u}_0 \\ \underline{y}_0 = \sigma_y (\cos\eta \cdot \underline{u}_0 + \sin\eta \cdot \underline{v}_0) \end{cases}$$

We consider a sample of size $f+1$ from the random vector $\underline{z}_0$. Equivalently we consider $f+1$ stochastically independent random vectors $\underline{z}_i$, $i = 1, \ldots f+1$, each isomorous with $\underline{z}_0$. We form the vectors $\underline{x}$, $\underline{y}$, $\underline{u}$ and $\underline{v} \in \mathbf{R}^{f+1}$ from the corresponding component random variables $\underline{x}_i$, $\underline{y}_i$, $\underline{u}_i$ and $\underline{v}_i$ for $i = 1, \ldots f+1$, and we obtain from (3.2.3):

(3.2.4) $$\begin{cases} \underline{x} = \sigma_x \underline{u} \\ \underline{y} = \sigma_y (\cos\eta \cdot \underline{u} + \sin\eta \cdot \underline{v}) \end{cases}$$

with $\underline{u} \cong \underline{v} \cong \underline{\chi}_{f+1}$, $\underline{u}$ and $\underline{v}$ stochastically independent.

Definition: The (theoretical) random correlation coefficient $\underline{r} = \underline{r}_f^\rho$ *with f degrees of freedom and population correlation coefficient* ρ *is:*

(3.2.5) $$\underline{r} = \cos\underline{h} = \frac{\underline{x}\,\underline{y}}{\sqrt{\underline{x}^2\,\underline{y}^2}}$$

where $\underline{x}\,\underline{y}$ and $\underline{x}^2 = \underline{x}\,\underline{x}$ are inner products and $\underline{h} = \arccos\underline{r}$.
Consequently:

(3.2.6) $$\frac{\underline{r}}{\sqrt{1-\underline{r}^2}} = \operatorname{cotg}\underline{h} = \frac{\underline{x}\underline{y}}{\sqrt{\underline{x}^2\,\underline{y}^2 - (\underline{x}\underline{y})^2}}$$

Here we substitute (3.2.4):

$$\operatorname{cotg}\underline{h} = \frac{\sigma_x \sigma_y \underline{u}(\cos\eta \cdot \underline{u} + \sin\eta \cdot \underline{v})}{\sigma_x \sigma_y \sqrt{\{\underline{u}^2 (\cos\eta \cdot \underline{u} + \sin\eta \cdot \underline{v})^2 - [\underline{u}(\cos\eta \cdot \underline{u} + \sin\eta \cdot \underline{v})]^2\}}}$$

As the term under the root sign equals

$$\underline{u}^2 [\cos^2\eta \cdot \underline{u}^2 + 2\sin\eta\cos\eta \cdot \underline{u}\underline{v} + \sin^2\eta \cdot \underline{v}^2] - (\cos\eta \cdot \underline{u}^2 + \sin\eta \cdot \underline{u}\underline{v})^2 = \\ = \sin^2\eta \cdot [\underline{u}^2\,\underline{v}^2 - (\underline{u}\underline{v})^2]$$

we find:

(3.2.7) $$\operatorname{cotg}\underline{h} = \frac{\dfrac{\underline{u}\,\underline{v}}{\sqrt{\underline{u}^2}} + \operatorname{cotg}\eta \cdot \sqrt{\underline{u}^2}}{\sqrt{\underline{v}^2 - \dfrac{(\underline{u}\,\underline{v})^2}{\underline{u}^2}}}$$

As $\underline{u}$ and $\underline{v}$ are independent, we have $\underline{v}|u \cong \underline{v}$ for any u; therefore for any value $u \neq 0$ of $\underline{u}$, $\dfrac{u\,\underline{v}}{\sqrt{u^2}}$ is the inner product of $\underline{v} \cong \underline{\chi}_{f+1}$ with a unit vector, and $\underline{v}^2 - \dfrac{(u\,\underline{v})^2}{u^2}$ is the square of the orthogonal projection of $\underline{v}$ on the space orthogonal to u, consequently in view of the main lemma [2.4]:

$$\frac{\underline{u}\,\underline{v}}{\sqrt{u^2}} \simeq \underline{\chi}, \qquad \underline{v}^2 - \frac{(\underline{u}\,\underline{v})^2}{u^2} \simeq \underline{\chi}_f^2$$

and they are stochastically independent.
Then also (unconditionally):

$$\frac{\underline{u}\,\underline{v}}{\sqrt{\underline{u}^2}} \simeq \underline{\chi}, \qquad \underline{v}^2 - \frac{(\underline{u}\,\underline{v})^2}{\underline{u}^2} \simeq \underline{\chi}_f^2$$

and they and $\underline{u}$ are stochastically totally independent. Furthermore $\underline{u}^2 \simeq \underline{\chi}_{f+1}^2$ and we have *the fundamental formula* (which we also could have used as definition)

$$(A) \qquad \frac{\underline{r}}{\sqrt{1-\underline{r}^2}} = \text{cotg}\, \underline{h} \simeq \frac{\underline{\chi} + \text{cotg}\, \eta \cdot \sqrt{\underline{\chi}_{f+1}^2}}{\sqrt{\underline{\chi}_f{}^2}}$$

where $\underline{r} = \underline{r}_f^{\rho}$; $\eta = \arccos \rho$; $\underline{\chi}$, $\underline{\chi}_f{}^2$ *and* $\underline{\chi}_{f+1}^2$ *are assumed stochastically totally independent.*

For $\rho = 0$ we obtain a known relation between $\underline{r}_f$ and $\underline{t}_f$:

$$(A^\circ) \qquad \frac{\underline{r}}{\sqrt{1-\underline{r}^2}} = \text{cotg}\, \underline{h} \simeq \frac{\underline{\chi}}{\sqrt{\underline{\ }_f{}^2}} = \frac{1}{\sqrt{f}}\, \underline{t}_f \quad \text{for } \rho = 0$$

Miss HARLEY [5] proved the following interesting theorem.

$$E(\underline{h}) = \eta$$

Observe that the left hand side is independent of the number of degrees of freedom f. DANIELS and KENDALL [3] gave a short proof of this result. We present a slightly different short proof.

Proof: First we observe that from elementary geometry it follows that the angle h between two vectors x and y, both different from zero, is equal to

$$h = \frac{\pi}{2}\, \text{mean}_{\,z \in S}\, [1 - \text{sign}\,(xz)\,(yz)]$$

where the function sign takes the values 1, 0, -1 in case $(xz)\,(yz)$ is > 0, $= 0$, < 0 respectively, where z runs over the sphere of unitvectors S, and the mean is taken with respect to the rotationinvariant continuous measure on S.

Next we compute, going back to the random variable $\underline{h}$:

$$E(\underline{h}) = E\, [\frac{\pi}{2}\, \text{mean}_{\,z \in S}\, [1 - \text{sign}\,(\underline{x}z)\,(\underline{y}z)]] = \text{mean}_{\,z \in S}\, E \frac{\pi}{2}\, [1 - \text{sign}\,(\underline{x}z)\,(\underline{y}z)]$$

But, as the pair $(\underline{x}, \underline{y})$ is invariant under orthogonal transformations, $(\underline{x}z)\,(\underline{y}z)$ is the same for any unitvector z. More precise, for any unitvectors z_1 and z_2 we have $(\underline{x}z_1)\,(\underline{y}z_1) \simeq (\underline{x}z_2)\,(\underline{y}z_2)$. In particular we may take for z the vector with first component 1 and other components zero. We then obtain

$$E(\underline{h}) = E \frac{\pi}{2}\, [1 - \text{sign}\, \underline{x}_0\, \underline{y}_0] = E \frac{\pi}{2}\, [1 - \text{sign}\, \underline{u}_0\, (\cos \eta \cdot \underline{u}_0 + \sin \eta \cdot \underline{v}_0)]$$

and there remains only to prove the theorem of Miss HARLEY for the case $f = 1$. As $(\underline{u}_0, \underline{v}_0) \simeq \underline{\chi}_2$ it is easily seen in the (u_0, v_0)-plane that

$$P\, [\underline{u}_0\, (\cos \eta \cdot \underline{u}_0 + \sin \eta \cdot \underline{v}_0) < 0] = \eta/\pi$$
$$P\, [u_0\, (\cos \eta \cdot \underline{u}_0 + \sin \eta \cdot \underline{v}_0) > 0] = (\pi - \eta)/\pi$$

Hence E sign $\underline{u}_0$ $(\cos\eta \cdot \underline{u}_0 + \sin\eta \cdot \underline{v}_0) = 1 \cdot (\pi - \eta)/\pi - 1 \cdot \eta/\pi = 1 - 2\eta/\pi$ and consequently $E(\underline{h}) = \eta$.

Remark: In case $\underline{x}$ and $\underline{y}$ are stochastically independent vectors then the formula $(A°)$ holds under more general assumptions. It is sufficient to assume that $\underline{y} \cong \sigma_y \underline{\chi}_{f+1}$ and $P(\underline{x} \neq 0) = 1$. This can be seen by following the above proof for the case $\rho = 0$ and writing $\underline{x}$ instead of $\underline{u}$. In particular $\underline{x}$ is allowed to be any constant vector $\neq 0$.

In view of the next sections we now consider the situation given in (3.2.4), but in $\mathbf{R}^n$ instead of $\mathbf{R}^{f+1}$, $n > f+1$. Hence $\underline{x}$, $\underline{y}$, $\underline{u}$ and $\underline{v}$ are random vectors in $\mathbf{R}^n$, $\underline{u} \cong \underline{v} \cong \underline{\chi}_n$, and $\underline{u}$ and $\underline{v}$ are stochastically independent. Let T be the $f+1$–dimensional linear subspace which consists of all vectors in $\mathbf{R}^n$ of which the last n–f–1 components are zero. The orthogonal projection of $\underline{x}$, etc. on T, denoted by $\underline{x}_T$ etc., is obtained from $\underline{x}$ by replacing the last n–f–1 by zero's. It is clear that by just omitting these last n–f–1 components we obtain the former case back and consequently

$$\text{(3.2.8)} \qquad \underline{r} = \frac{\underline{x}_T\, \underline{y}_T}{\sqrt{\underline{x}_T^2\, \underline{y}_T^2}} \cong \underline{r}_f^\rho$$

and also this $\underline{r}$ satisfies formula (A), in particular $(A°)$ in case $\rho = 0$.

In (3.2.4) the pair of independent standard normal random vectors $\underline{u}$ and $\underline{v}$ in $\mathbf{R}^n$ is invariant under orthogonal transformations. Hence if we first apply some orthogonal transformation H to $\underline{u}$, $\underline{v}$, $\underline{x}$ and $\underline{y}$, we obtain a pair $(H\underline{x}, H\underline{y})$ which is as a pair isomorous to $(\underline{x}, \underline{y})$. If we compute $\underline{r}$ from $(H\underline{x})_T$ and $(H\underline{y})_T$ we get the same results as from $\underline{x}_T$ and $\underline{y}_T$. It is equivalent to apply H^{-1} to the subspace T and to determine the correlation coefficient of $\underline{x}_{(H^{-1}T)}$ and $\underline{y}_{(H^{-1}T)}$. As H is an arbitrary orthogonal transformation this means that T can be replaced by any $f+1$-dimensional linear subspace of $\mathbf{R}^n$: (3.2.8) and (A) *holds for any* $f+1$*-dimensional linear subspace* T *of* $\mathbf{R}^n$.

3.3. *The sample correlation coefficient*

We now go back to the case (3.2 2) of the general normal vector in $\mathbf{R}^2$ and we consider a sample of size $n = f+2$. This yields random vectors

$$\text{(3.3.1)} \qquad \begin{cases} \underline{x} = \mu_x \cdot e + \sigma_x \underline{u} \\ \underline{y} = \mu_y \cdot e + \sigma_y (\cos\eta \cdot \underline{u} + \sin\eta \cdot \underline{v}) \end{cases}$$

where e is the vector with n components equal to 1. Let $T \subset \mathbf{R}^n$ be the $f+1$-dimensional subspace of all vectors orthogonal to e. As $e_T = 0$ it follows that $(\underline{x}_T, \underline{y}_T)$ is independent of (μ_x, μ_y) and yields the same if we replace (μ_x, μ_y) by $(0, 0)$. Hence we can apply (3.2.8).

$$\text{Let } \underline{\bar{x}} = \frac{1}{n} \Sigma_i \underline{x}_i, \quad \underline{\bar{y}} = \frac{1}{n} \Sigma_i \underline{y}_i$$

Then $\underline{x}_T$ and $\underline{y}_T$ are the vectors with components $\underline{x}_i - \underline{\bar{x}}$ for $i = 1, \ldots n$ and $\underline{y}_i - \underline{\bar{y}}$ for $i = 1, \ldots n$ respectively.

$$\underline{x}_T^2 = \Sigma_i (\underline{x}_i - \underline{\bar{x}})^2, \; \underline{y}_T^2 = \Sigma_i (\underline{y}_i - \underline{\bar{y}})^2,$$
$$\underline{x}_T\, \underline{y}_T = \Sigma_i (\underline{x}_i - \underline{\bar{x}})(\underline{y}_i - \underline{\bar{y}})$$

and $\underline{r}$ in (3.2.8) is seen to be the usual sample correlation coefficient.

The sample correlation coefficient is isomorous to $\underline{r}_f^\rho$. If the sample has size n then the number of degrees of freedom or dimension is $f = n-2$.

Remark. It follows from the remark in section 3.2 that if $\underline{x}$ and $\underline{y}$ are *stochastically independent* then the result holds even under the weaker assumptions $\underline{y} \cong \mu_y e + \sigma \underline{\chi}_n$, $\underline{x}$ an arbitrary (for example constant) random vector, for which $P(\underline{x}_T \neq 0) = 1$, that is for the present case: $P(\underline{x}_1 = \underline{x}_2 .. = \underline{x}_n) = 0$. For example $\underline{x}$ can be a vector which takes e.g. as only values the permutations of the sequence of the first n integers (SPEARMAN vector). Also in this case, and under the given conditions, $\underline{r}$ satisfies formula ($A°$). (Usual t-test).

3.4. *The sample p-partial correlation coefficient*

We consider a set of $p + 2$ random variables which together form a general normal random vector *with expectation zero*. We assume that $\underline{x}_0$ as well as $\underline{y}_0$ together with the p other random variables form a random vector with correlation matrix of rank $p + 1$. This means that no linear function of $\underline{x}_0$ (or $\underline{y}_0$) and the other p random variables is a constant random variable. We are not so much interested now in the p other random variables individually as well as in the p dimensional euclidean vector space of their linear combinations (linear homogeneous functions).

There exists a suitable set of orthonormal combinations to be called $\underline{z}_{(1)0}, \dots \underline{z}_{(p)0}$, such that

(3.4.1)
$$\begin{cases} \underline{x}_0 = \alpha \cdot \underline{z}_{(1)0} + \beta \, \underline{u}_0 \\ \underline{y}_0 = \gamma \, \underline{z}_{(1)0} + \delta \, \underline{z}_{(2)0} + \varepsilon \, (\cos \eta \cdot \underline{u}_0 + \sin \eta \cdot \underline{v}_0) \end{cases}$$

where:
$$\begin{pmatrix} \underline{z}_{(1)0} \\ \vdots \\ \underline{z}_{(p)0} \\ \underline{u}_0 \\ \underline{v}_0 \end{pmatrix} \cong \underline{\chi}_{p+2}, \text{ and } 0 \leqslant \eta \leqslant \pi, \ \beta > 0, \ \varepsilon > 0.$$

(3.4.1) is obtained by a orthonormalisation process from the p originally given random variables. The number $\rho = \cos \eta$ is unique and it is called *the population p-partial correlation coefficient of* $\underline{x}_0$ *and* $\underline{y}_0$ *with respect to the p given random variables.* η is "the angle between the components of $\underline{x}_0$ and $\underline{y}_0$, orthogonal to $\underline{z}_{(1)0}, \dots \underline{z}_{(p)0}$ in the linear space of all random variables".

We take a sample of $f+p+1$ independent copies of $\underline{x}_0$, $\underline{y}_0$ etc. and combine them in vectors $\underline{x}$, $\underline{y}$ etc. Then (3.4.1) yields

(3.4.2)
$$\begin{cases} \underline{x} = \alpha \, \underline{z}_{(1)} + \beta \, \underline{u} \\ \underline{y} = \gamma \, \underline{z}_{(1)} + \delta \, \underline{z}_{(2)} + \varepsilon \, (\cos \eta \cdot \underline{u} + \sin \eta \cdot \underline{v}) \end{cases}$$

(3.4.3)
$$\underline{z}_{(1)} \cong \underline{z}_{(2)} \dots \cong \underline{z}_{(p)} \cong \underline{u} \cong \underline{v} \cong \underline{\chi}_{f+p+1}$$

and all random vectors in (3.4.3) are stochastically independent. The sample partial correlation coefficient is defined to be

(3.4.4)
$$\underline{r} = \frac{\underline{x}_{\underline{T}} \, \underline{y}_{\underline{T}}}{\sqrt{\underline{x}_{\underline{T}}^2 \, \underline{y}_{\underline{T}}^2}}$$

where $\underline{T}$ is the random linear subspace of dimension $f+1$, orthogonal to the vectors $\underline{z}_{(1)}, \underline{z}_{(2)} \ldots \underline{z}_{(p)}$, which are linearly independent with probability one, as their correlation matrix has rank p.

Under the condition that $\underline{z}_{(1)}, \ldots \underline{z}_{(p)}$ have the p linearly independent values $z_{(1)}, \ldots z_{(p)}$ we have for the conditional random vectors:

$$\underline{x} = \alpha\, z_{(1)} + \beta\, \underline{u}$$
$$\underline{y} = \gamma\, z_{(1)} + \delta\, z_{(2)} + \varepsilon\,(\cos\eta \cdot \underline{u} + \sin\eta \cdot \underline{v})$$

where $\underline{u}, \underline{v}, z_{(1)}, \ldots z_{(p)}$ are mutually independent; hence $\underline{u} \cong \underline{v} \cong \underline{\chi}_{f+p+1}$. T is the space orthogonal to $z_{(1)}, \ldots z_{(p)}$. Therefore $\underline{x}_{\mathrm{T}}$ and $\underline{y}_{\mathrm{T}}$ do not change if we omit $\alpha\, z_{(1)}$ from $\underline{x}$, and $\gamma\, z_{(1)} + \delta\, z_{(2)}$ from $\underline{y}$. But if then we compute $\underline{r}$ from (3.4.4) under the condition $\underline{\mathrm{T}} = \mathrm{T}$ we obtain a random variable which is isomorous to $\underline{r} = \underline{r}_f^{\rho}$ in formula (A). Then (3.4.4) is unconditionally isomorous to $\underline{r}_f^{\rho}$.

In case the expectation vector of the originally given normal random vector $\in \mathbf{R}^{p+2}$ *is not known to be zero,* we apply a construction as in section (3.3) projecting all vectors in $\mathbf{R}^n$ in the space orthogonal to the vector e, to obtain *the usual sample p-partial correlation coefficient* $\underline{r}$ which therefore *is also isomorous to* $\underline{r}_f^{\rho}$ *in formula A*. The dimension f is obtained from $n = f + p + 2$ in this case.

3.5. *Approximation of (A)*

If we substitute the approximation of section 1.11

$$\sqrt{\underline{\chi}^2_{f+1}} \cong \frac{1}{\sqrt{2}}\left(\sqrt{(2f+1)} + \underline{\chi}\right)$$

in (A) we get

$$\operatorname{cotg} \underline{h} \cong \frac{\operatorname{cotg}\eta \cdot \sqrt{(f+\frac{1}{2})} + \sqrt{(1^2 + \frac{1}{2}\operatorname{cotg}^2\eta)} \cdot \underline{\chi}}{\sqrt{\underline{\chi}_f^{\,2}}}$$

That is

(A_{app})

$$\frac{\underline{r}}{\sqrt{1-\underline{r}^2}} = \operatorname{cotg} \underline{h} \cong \lambda\, \underline{t}_f^{\delta} \text{ approximately}$$

with

$$\lambda = \sqrt{\frac{2 + \operatorname{cotg}^2\eta}{2f}}, \quad \delta = \operatorname{cotg}\eta \sqrt{\frac{2f+1}{2 + \operatorname{cotg}^2\eta}}$$

In TABLE 1 a comparison is made between cotg $\underline{h}$ and its approximation $\lambda \underline{t}_f^{\delta}$. Tables for these two random variables by F. N. DAVID [4] and N. L. JOHNSON and B. L. WELCH [6] were used for this purpose.

It should be noted that exactly the same approximation has been obtained by Miss. B. I. HARVEY [5] by trial and guess. Her method is to equate the *second* (non central) moments as well as the quotients of third to first moment of the following random variables:

$$g(\rho) \cdot \frac{\underline{r}\sqrt{f}}{\sqrt{1-\underline{r}^2}}, \quad \underline{t}_f^{\delta}$$

She then finds

$$\delta = \frac{\sqrt{(2f+1)}\,\rho}{\sqrt{2-\rho^2}}, \quad g(\rho) = \sqrt{2}\,\sqrt{\frac{1-\rho^2}{2-\rho^2}}.$$

This is easily seen to be equivalent to our above equations for λ and δ.

In accordance with our TABLE 1, her TABLE 2 suggests a very high agreement between the two random variables: however she restricts her considerations to upper α-levels. $\alpha \leqslant 0{,}40$. It has been pointed out by Miss HARLEY that the above approximation can also be used in the inverse way, to give upper α–points for the random variables $\underline{t}_f^\delta$ with the help of tables for $\underline{r}_f^\rho$.

From Table 1 we conclude that the given approximation is very satisfactory except for the case of lower-α-points with small f and large ρ.

TABLE 1. Comparison between exact and approximate upper and lower α-points of $\underline{r} = \underline{r}_f^\rho$.

			Lower α-point			Upper α-point	
α	f	ρ	exact c^l	approx. c^l_y	α'	exact c^u	approx. c^u_y
0,05	5	0	−.6694	−.6692	.0499	.6695	.6693
		0.5	−.1951	−.2015	.0487	.8863	.8862
		0.7	.1526	.1358	.0467		
		0.8	.3838	.3610	.0449		
		0.9	.6639	.6410	.0424	.9822	.9822
	10	0.8	.5402	.5353	.0480		
		0.9	.7559	.7496	.0457	.9678	.9677
	20	0.5	.1844	.1835	.0496	.7341	.7340
		0.9	.8086	.8066	.0475	.9535	.9534
	98	0.9	.8643	.8641	.0491	.9280	.9279
0,025	5	0.5	−.3534	−.3608	.0241	.9182	.9181
		0.8	+.2343	.1937	.0208		
		0.9	.5637	.5147	0.183	.98735	.98734
	10	0.9	.7061	.6936	.0223		
	20	0.5	.1120	.1106	.0246	.7665	.7663
		0.9	.7817	.7780	.0227	.9597	.9596
	98	0.5	.3389	.3387	.0249	.6357	.6356
		0.9	.8560	.8556	.0242	.9323	.9322

$$\underline{y} = \cos \operatorname{arccotg} \lambda\, \underline{t}_f^\delta \text{ or } \frac{\underline{y}}{\sqrt{1-\underline{y}^2}} = \lambda\, \underline{t}_f^\delta$$

The lower α-point of $\underline{r}$ is c^l, with $P(\underline{r} < c^l) = \alpha$

The lower α-point of $\underline{y}$ is c^l_y, with $P(\underline{y} < c^l_y) = \alpha$

$$\alpha' = P(\underline{r} < c^l_y)$$

The upper α-point of $\underline{r}$ is c^u, with $P(\underline{r} > c^u) = \alpha$

The upper α-point of $\underline{y}$ is c^u_y, with $P(\underline{y} > c^u_y) = \alpha$

$$\alpha'' = P(\underline{r} > c^l_y)$$

In the table α'' is not given, because it is almost equal to α in all cases presented in the table.

3.6. *The random variable* $\underline{r}_f^{m,\,\rho}$.

We consider a set of $m+1$ random variables which together form a general normal random vector in $\mathbf{R}^{m+1}$ *with expectation zero.* One of them is called $\underline{x}_0$. The other m are assumed to be linearly independent with probability one. As in section 3.4. we orthonormalise. We replace the m given random variables by m linear combinations to be called $\underline{z}_{(1)0}, \ldots \underline{z}_{(m)0}$, such that

(3.6.1) $$\underline{x}_0 = \alpha\,(\cos\eta\cdot\underline{z}_{(1)0} + \sin\eta\cdot\underline{u}_0)$$
and
(3.6.2) $$\begin{pmatrix} \underline{z}_{(1)0} \\ \vdots \\ \underline{z}_{(m)0} \\ \underline{u}_0 \end{pmatrix} \cong \underline{\chi}_{m+1}, \quad 0 \leqslant \eta \leqslant \frac{\pi}{2}, \alpha > 0.$$

Then $\rho = \cos\eta$, which is unique, is called the *population multiple correlation coefficient* of $\underline{x}_0$ with respect to the other random variables.

We take a sample of $f+m$ independent copies of $\underline{x}_0$ etc., and we combine these $\underline{x}_i$ $i=1, \ldots f+m$, etc. in random vectors in $\mathbf{R}^{f+m}$ which in view of (3.6.1) and (3.6.2) satisfy:

(3.6.3) $$\underline{x} = \alpha\,(\cos\eta\cdot\underline{z}_{(1)} + \sin\eta\cdot\underline{u}).$$
and
(3.6.4) $$\underline{z}_{(1)} \cong \ldots \cong \underline{z}_{(m)} \cong \underline{u} \cong \underline{\chi}_{f+m}$$

and the random vectors in (3.6.4) are stochastically independent.

Let $\underline{\mathrm{M}}$ be the m-dimensional subspace of $\mathbf{R}^{f+m}$ with basis $\underline{z}_{(1)}, \ldots \underline{z}_{(m)}$, who are linearly independent with probability one. Let $\underline{\mathrm{F}}$ be the f-dimensional space orthogonal to $\underline{\mathrm{M}}$.

The random multiple correlation coefficient $\underline{r}_f^{m,\,\rho} = \underline{r} = \cos\underline{h} \geqslant 0$ is defined by

(3.6.5) $$\frac{\underline{r}^2}{1-\underline{r}^2} = \mathrm{cotg}^2\,\underline{h} = \frac{(\underline{x}_{\underline{\mathrm{M}}})^2}{(\underline{x}_{\underline{\mathrm{F}}})^2} \qquad \text{for } \rho < 1$$
$$\text{and} \quad \underline{r} \cong 1 \qquad\qquad \text{for } \rho = 1$$

cotg $\underline{h}$ is the ratio between the lengths of the orthogonal projections $\underline{x}_{\underline{\mathrm{M}}}$ of $\underline{x}$ into $\underline{\mathrm{M}}$, and $\underline{x}_{\underline{\mathrm{F}}}$ of $\underline{x}$ into $\underline{\mathrm{F}}$. In the sequel we restrict to the case $\rho < 1$, hence $\sin\eta > 0$.

We now change temporarily to a *conditional* random situation. The condition we consider is that $\underline{z}_{(1)}, \ldots \underline{z}_{(m)}$ have the p linearly independent values $z_{(1)}, \ldots z_{(m)}$.

M (the value of $\underline{\mathrm{M}}$) is the space of dimension m spanned by $z_{(1)}, \ldots z_{(m)}$ · F is the space of dimension f orthogonal to $z_{(1)}, \ldots z_{(m)}$. Let B be the space of dimension 1 with basis the unitvector $b = (z_{(1)}^2)^{-\frac{1}{2}} \cdot z_{(1)}$. Let A be the space of dimension m–1 of all vectors in M that are orthogonal to b

Conditionally we find

$$\mathrm{cotg}^2\,\underline{h} = \frac{\underline{x}_{\mathrm{M}}^2}{\underline{x}_{\mathrm{F}}^2} = \frac{\alpha^2\,[\cos\eta\cdot z_{(1)} + \sin\eta\cdot(\underline{u}_{\mathrm{B}} + \underline{u}_{\mathrm{A}})]^2}{\alpha^2\,[0 + \sin\eta\cdot\underline{u}_{\mathrm{F}}]^2}$$

$$= \frac{[\operatorname{cotg}\eta \cdot z_{(1)} + (\underline{u}b)b + \underline{u}_A]^2}{\underline{u}_F{}^2}$$

$$= \frac{[\{\operatorname{cotg}\eta \cdot \sqrt{z^2{}_{(1)}} + (\underline{u}b)\}\, b]^2 + \underline{u}_A{}^2}{\underline{u}_F{}^2}$$

$$= \frac{[\operatorname{cotg}\eta \cdot \sqrt{z^2{}_{(1)}} + (\underline{u}b)]^2 + \underline{u}_A{}^2}{\underline{u}_F{}^2}$$

As $\underline{u}$ is stochastically independent of $\underline{z}_1, \ldots \underline{z}_p$, we have

$$\underline{u} \mid z_1, \ldots z_p \cong \underline{u} \cong \underline{\chi}_{f+m}$$

therefore

$$\underline{u}b \cong \underline{\chi},\ \underline{u}_A{}^2 \cong \underline{\chi}^2_{m-1},\ \underline{u}_F{}^2 \cong \underline{\chi}_f{}^2,$$

and as B, A and F are mutually orthogonal, these random variables are stochastically independent. We can write *conditionally:*

$$\operatorname{cotg}^2 \underline{h} \cong \frac{(\operatorname{cotg}\eta \cdot \sqrt{z^2{}_{(1)}} + \underline{\chi})^2 + \underline{\chi}^2_{m-1}}{\underline{\chi}_f{}^2}$$

where $\underline{\chi}$, $\underline{\chi}_{m-1}$ and $\underline{\chi}_f$ are also stochastically independent of $\underline{z}_{(1)}$.

Then *unconditionally:*

$$\operatorname{cotg}^2 \underline{h} \cong \frac{(\operatorname{cotg}\eta \cdot \sqrt{\underline{z}^2{}_{(1)}} + \underline{\chi})^2 + \underline{\chi}^2_{m-1}}{\underline{\chi}_f{}^2}$$

and so, as

$$\underline{z}_{(1)} \cong \underline{\chi}_{f+m}$$

we finally obtain the formula

(*B*)
$$\frac{\underline{r}}{1-\underline{r}^2} = \operatorname{cotg}^2 \underline{h} \cong \frac{(\operatorname{cotg}\eta \cdot \sqrt{\underline{\chi}^2_{f+m}} + \underline{\chi})^2 + \underline{\chi}^2_{m-1}}{\underline{\chi}_f{}^2}, \qquad \eta = \arccos \rho$$

where $\underline{\chi}$, $\underline{\chi}_{m-1}$, $\underline{\chi}_f$ and $\underline{\chi}_{f+m}$ are in this formula assumed to be stochastically totally independent.

In particular if the population multiple correlation coefficient $\rho = \cos\eta = 0$ we get

$$\operatorname{cotg}^2 \underline{h} \cong \frac{\underline{\chi}^2 + \underline{\chi}^2_{m-1}}{\underline{\chi}_f{}^2}$$

That is:

(*B°*)
$$\frac{\underline{r}^2}{1-\underline{r}^2} = \operatorname{cotg}^2 \underline{h} \cong \frac{\underline{\chi}^2_m}{\underline{\chi}_f{}^2} = \frac{m}{f}\, \underline{\mathrm{F}}^m_f \qquad \text{for } \rho = 0$$

3.7. *The sample m-tuple correlation coefficient*

The usual *sample m*-tuple correlation coefficient refers to the more general case where the expectation of $(\underline{x}_0, \underline{z}_{(1)0}, \ldots \underline{z}_{(m)0})$ is not known to be zero. We then apply a construction as in section 3.3. projecting all vectors in $\mathbf{R}^n$ ($n = f + m + 1$) in the space of all vectors orthogonal to the vector $e \in \mathbf{R}^n$. We conclude that *the usual random n-sample m-tuple correlation coefficient* $\underline{r}$ *of one random variable with respect to m other random variables with correlation matrix of rank m, and with population m-tuple correlation coefficient* ρ ($0 \leqslant \rho \leqslant 1$) *is isomorous to* $\underline{r}_f^{m,\rho}$ *in formula* (*B*). The dimension f is $f = n-m-1$ {*or* (B°) *in case* $\rho = 0$}. Also the *random n-sample p-partial m-tuple correlation coefficient concerning a sample of size* $n = f + m + p + 1$ of copies of a random normal vector in $\mathbf{R}^{m+p+1}$ with population p-partial m-tuple correlation coefficient ρ is isomorous to $\underline{r}_f^{m,\rho}$. The dimension f is $f = n-m-1-p$. We do not enter into the details of the proof of these last statements.

3.8. *Two approximations for* $\underline{r} = \underline{r}_f^{m,\rho}$.

Let $\underline{\gamma}_a$ be the *gamma random variable* with probability density function

$$\begin{cases} (\Gamma(a))^{-1} e^{-t}\, t^{a-1} & \text{for } t > 0 \\ 0 & \text{for } t \leqslant 0 \end{cases}$$

For $g = 2a$ a positive integer, one has the known relation

$$\underline{\chi}_{2a}^2 \cong 2\underline{\gamma}_a$$

which can be used to define the left hand side for any a > 0. Furthermore one defines for any $g > 1$, $\gamma \geqslant 0$, the non-central chi-square random variable by

$$(\underline{\chi}_g^\gamma)^2 \cong \underline{\chi}_{g-1}^2 + (\underline{\chi} + \gamma)^2$$

and the non-central F by

$$\underline{F}_f^{g,\gamma} \cong \frac{(\underline{\chi}_g^\gamma)^2/g}{\underline{\chi}_f^2/f}$$

The first approximation of formula *B* for

(3.8.1) $$\underline{r} = \underline{r}_f^{m,\rho}, \text{ or } \underline{h} = \arccos \underline{r}, \text{ or } \operatorname{cotg}^2 \underline{h} = \frac{\underline{r}^2}{1-\underline{r}^2},$$

is obtained by first finding real constants λ and $g > 0$ such that

(3.8.2) $$(\underline{\chi} + \operatorname{cotg} \eta \sqrt{\underline{\chi}_{f+m}^2})^2 + \underline{\chi}_{m-1}^2$$

and

(3.8.3) $$\lambda\, \underline{\chi}_g^2$$

have the same first and second moments. After straightforward computations one finds then on dividing by $\underline{\chi}_f^2$ the *approximation of formula* (*B*):

(3.8.4) $$\boxed{\operatorname{cotg}^2 \underline{h} \simeq \lambda \frac{\underline{\chi}_g^2}{\underline{\chi}_f^2} = \lambda \frac{g}{f} \underline{F}_f^g}$$

where: $$\lambda = 2 - \frac{m}{m+f\rho^2} + \frac{\rho^2}{1-\rho^2} = 2 - \frac{m}{m+f\rho^2} + \operatorname{cotg}^2 \eta$$

and $$g = \frac{m + f\,\rho^2}{\lambda\,(1 - \rho^2)} = \frac{m + f\cos^2\eta}{\lambda\sin^2\eta}$$

A very satisfactory and useful *second approximation*, which is better than the first, was obtained by KEULS [7] who replaced the numerator (3.8.2) of (B) by

(3.8.5) $$\lambda\,(\underline{\chi}_g^{\gamma})^2$$

where λ, γ and g are chosen such that the first three moments of (3.8.2) and (3.8.5) coincice. It turns out that then also the fourth moments are practically the same. On substitution one finds the resulting approximation:

(3.8.6) $$\underline{h} = \arccos \underline{r}_f^{m,\rho}$$

$$\boxed{\operatorname{cotg}^2 \underline{h} \cong \lambda \frac{g}{f}\ \underline{F}_f^{g,\gamma}}$$

with $$\lambda = 1 + \frac{b\operatorname{cotg}^2\eta}{b + \sin\eta},\ b = \sqrt{(1 + \frac{m}{f})}$$

$$\gamma^2 = \frac{bf\operatorname{cotg}^2\eta}{\lambda^2\sin\eta}$$

$$g = \frac{(m + f)\operatorname{cotg}^2\eta + m}{\lambda} - \gamma^2$$

October 1960, Landbouwhogeschool, Wageningen, The Netherlands.

REFERENCES

[1]. ANDERSON, T. W.: Introduction to multivariate analysis, Chapter 4, 1958.

[2]. CORSTEN, L. C. A.: Vectors, a tool in statistical regression theory. Diss. Landbouwhogeschool, Wageningen (1958) = Mededelingen van de Landbouwhogeschool te Wageningen **158** (1), 1958 = Mededelingen van het Inst. voor Rassenonderzoek van Landbouwgewassen te Wageningen, nr. 35.

[3]. DANIELS, H. E. – KENDALL, M. G.: Short proof of Miss HARLEY's theorem on the correlation coefficient. Biometrika **45**, 571–572 (1958).

[4]. DAVID, F. N.: Tables of the correlation coefficient, 1938.

[5]. HARLEY, B. I.: Some properties of an angular transformation for the correlation coefficient. Biometrika **43**, 219–223 (1956).

———: Further properties of an angular transformation of the correlation coefficient. Biometrika **43**, 273–275 (1957).

———: Relation between the distributions of non-central t and of a transformed correlation coefficient. Biometrika **44**, 219 (1957).

[6]. JOHNSON, N. L. and WELCH, B. L.: Applications of the non-central t-distribution. Biom. **31**, 362 (1939).

[7]. KEULS, M.: La puissance du critère F dans l'analyse de la variance de plans en blocs au hasard, nomogrammes pour le choix du nombre de répétitions. Biométrie-Praximetrie. Tome 1, pp. 65–81, 1960.

[8]. KUIPER, N. H.: Variantie-analyse. Statistica Neerlandica **6**, 149–194 (1952).

———: Een symbool voor isomoor. Statistica Neerlandica **13**, 295–317 (1959).

KUIPER, N. H.: Random variables and random vectors. Bulletin de l'Institut Agronomique et des stations de recherches. Gembloux, 1960. p. 344–355.
———: Tests concerning random points on a circle. Indag-Math. **22**, 38–47 (1960) = Proc. Amsterdam.
———: Linear algebra and geometry. Chapter 16: Applications in Statistics. North-Holland Publishing Company, Amsterdam, 1962. Dutch edition (1959).
———: On a metric in the space of random variables. Statistica Neerlandica 16, p. 231–235 (1962).