DISTRIBUTION OF A SERIAL CORRELATION COEFFICIENT NEAR THE ENDS OF THE RANGE[1]

By M. M. Siddiqui

University of North Carolina

1. Introduction and summary. If $y_1, \cdots, y_N$ are observations on a stationary time series at equal intervals of time and it is known that $Ey_t = 0$ for $t = 1, \cdots, N$, the most natural definition of a serial correlation coefficient with lag unity would be

$$r^* = \left(\sum_{i=1}^{N-1} y_i y_{i+1}\right)\left[\left(\sum_{i=1}^{N-1} y_i^2\right)\left(\sum_{i=1}^{N-1} y_{i+1}^2\right)\right]^{-1/2}$$

if the denominator $\neq 0$. This is the ordinary correlation coefficient between $(y_1, \cdots, y_{N-1})$ and $(y_2, \cdots, y_N)$, except that instead of taking deviations from the sample mean, we have taken deviations from the population means. Due to the seemingly unsurmountable mathematical difficulties involved in obtaining the distribution of r^* even on the hypothesis of independence and normality of the observations, several alternative definitions have been proposed as approximations to r^*. However, it is desirable to consider some relevant properties of the distribution of r^*.

In this paper the distribution of r^* near the extremities of its range will be considered. The observations will be assumed to be distributed as independent $N(0, 1)$ variates. There is no loss of generality in assuming the variance to be unity as r^* is independent of the scale parameter. A geometrical approach suggested by Hotelling seemed to be particularly suitable in obtaining the order of contact of the distribution curve at $r^* = \pm 1$. Hotelling [1] shows how to determine the order of contact of frequency curves of some statistics with the variate axis at the ends of the range even though the actual distributions are unknown. It will be shown here that if for a number r_0 in [0, 1] and close to 1, $P(r^* \geqq r_0)$ is expanded in a series of powers of $(1 - r_0)$, the first non-zero coefficient is that of the power $(N - 2)/2$. Upper and lower bounds for the coefficient of this power will be calculated. The lower bound is positive and the upper bound gives an approximation for an upper bound on $P(r^* \geqq r_0)$.

2. Geometrical representation. Let $X_1, \cdots, X_N$ be N independent $N(0, 1)$ variates. Define

$$r^* = \left(\sum X_i X_{i+1}\right)\left[\left(\sum X_i^2\right)\left(\sum X_{i+1}^2\right)\right]^{-1/2} \tag{2.1}$$

where all the summations are from 1 to $N - 1$ and the denominator $\neq 0$, then r^* is a variate with range [0, 1].

Received August 20, 1957; revised February 7, 1958.

[1] Sponsored by the Office of Naval Research under the contract for research in probability and statistics at Chapel Hill. Reproduction in whole or in part is permitted for any purpose of the United States government.

For every set of observations $y_1, \cdots, y_N$ on these variates we take a point S with coordinates $(y_1, \cdots, y_N)$ in an N-dimensional Euclidean space, which may be regarded as a representation of the sample space. Denoting the origin by O, we see that the points S are distributed with spherical symmetry about O. Furthermore, a unique value of r^* corresponds to all the points on a straight line $0S$, excepting the origin. Let the straight line OS meet the $N - 1$-dimensional unit sphere in Q and Q', where Q is on the same side of the origin with S. Denoting by $(x_1, \cdots, x_N)$ the coordinates of Q, we have

$$\sum_{i=1}^{N} x_i^2 = 1, \tag{2.2}$$

which may also be taken as the equation of the unit sphere. The points Q and Q' may be considered to determine a unique value of r^*. Considering only the point Q, it is easily seen that the distribution of Q is uniform over the unit sphere; that is, denoting the total $(N - 1)$-dimensional surface area of (2.2) by S_{N-1}, the probability of Q falling in an area A on the sphere is

$$A/S_{N-1}.$$

For a given r_0 in $[-1, 1]$ there exists a set of points on the unit sphere such that for each point in this set the corresponding value of r^* lies in the interval $[r_0, 1]$, and for no other point. If this set of points covers an area A on the surface of the sphere (2.2), it follows that

$$P(r^* \geqq r_0) = A/S_{N-1}.$$

We observe that $r^* = 1$ if and only if $x_i = \lambda x_{i-1}$, $i = 2, 3, \cdots, N$, $\lambda > 0$ and $x_1 \neq 0$, that is, $x_i = \lambda^{i-1} x_1$, $i = 2, 3, \cdots, N$, $\lambda > 0$ and $x_1 \neq 0$. Since the point $(x_1, \cdots, x_N)$ lies on (2.2), we obtain for the value of x_1, $x_1 = \pm c$ where

$$c = (1 - \lambda^2)^{1/2}(1 - \lambda^{2N})^{-1/2}. \tag{2.3}$$

Denote the variable point $(c, \lambda c, \cdots, \lambda^{N-1}c)$ by P and $(-c, -\lambda c, \cdots, -\lambda^{N-1}c)$ by P'. As λ varies from 0 to ∞, each of P and P' describes a curve for every point of which—excepting the two points of each curve obtained by $\lambda = 0$ and ∞—corresponds the value of $r^* = 1$.

Since both these curves are exactly alike, except for their position in space, we confine our attention to the curve

$$x_1 = c, \quad x_i = \lambda^{i-1} x_1, \qquad i = 2, \cdots, N, 0 < \lambda < \infty. \tag{2.4}$$

Further, from now on we reserve $(x_1, \cdots, x_N)$ to denote the point on curve (2.4) which corresponds to the parameter λ, and we use $(\epsilon_1, \cdots, \epsilon_N)$ to denote any other point on the unit sphere.

To find the probability of r^* exceeding a given value r_0 which is close to 1, we consider the points within a "tube" of geodesic radius θ on the surface of the sphere (2.2) with its axial curve (2.4).

Let the length of the curve (2.4) measured from $P_0(1, 0, \cdots, 0)$ to

$$P(x_1, \cdots, x_N)$$

be denoted by s, or more explicitly $s(\lambda)$, and an element of curve by ds. Denoting by primes the differential coefficient with respect to s, the direction cosines of the tangent to the curve at P are

$$x_1', x_2', \cdots, x_N',$$

where

$$x_i' = [(i-1)\lambda^{i-2}c + \lambda^{i-1}\, dc/d\lambda]\lambda', \qquad i = 1, 2, \cdots, N. \tag{2.5}$$

We note that

$$\sum_{i=1}^{N} x_i'^2 = 1, \tag{2.6}$$

and since

$$\sum_{i=1}^{N} x_i^2 = 1$$

we have

$$\sum_{i=1}^{N} x_i x_i' = 0. \tag{2.7}$$

Let the coordinate axes be rotated so that the new coordinates are denoted by the elements of a vector α. Let $\alpha = B\epsilon$ where

$$B = \begin{bmatrix} x_1' & x_2' & \cdots & x_N' \\ x_1 & x_2 & \cdots & x_N \\ b_{31} & b_{32} & \cdots & b_{3N} \\ \cdots & \cdots & \cdots & \cdots \\ b_{N1} & b_{N2} & \cdots & b_{NN} \end{bmatrix},$$

and

$$BB' = I. \tag{2.8}$$

Here I denotes the identity matrix, B' denotes the transpose of B, and ϵ and α denote the column vectors $(\epsilon_1, \cdots, \epsilon_N)$ and $(\alpha_1, \cdots, \alpha_N)$ respectively.

The α_1 axis is now parallel to the tangent of the curve at P and the α_2 axis coincides with the line OP.

The $(N-3)$-dimensional sphere given by the set of equations

$$\alpha_1 = 0, \qquad \alpha_2 = \cos\theta, \qquad \alpha_i = \beta_i \sin\theta, \qquad i = 3, 4, \cdots, N, \tag{2.9}$$

with

$$\sum_{i=3}^{N} \beta_i^2 = 1,$$

lies entirely on the $(N - 1)$-dimensional unit sphere

$$\sum_{i=1}^{N} \alpha_i^2 = 1 = \sum_{i=1}^{N} \epsilon_i^2. \tag{2.10}$$

The sphere (2.10) is the same as (2.2) with a change of notation. Each point on (2.9) is at a geodesic distance θ from P measured on the sphere (2.10). Further, since (2.9) lies in the plane $\alpha_1 = 0$, this hypersphere is perpendicular to the tangent of curve (2.4) at P.

Changing back to the original coordinates we have $\epsilon = B'\alpha$ or

$$\epsilon_i = x_i'\alpha_1 + x_i\alpha_2 + b_{3i}\alpha_3 + \cdots + b_{Ni}\alpha_N, \qquad i = 1, \cdots, N.$$

Equations (2.9) become

$$\epsilon_j = x_j \cos\theta + \sin\theta \sum_{i=3}^{N} b_{ij}\beta_i, \qquad j = 1, 2, \cdots, N \tag{2.11}$$

with

$$\sum_{i=3}^{N} \beta_i^2 = 1.$$

3. The value of r^* near the curve. Let us calculate the value of r^* corresponding to a point $(\epsilon_1, \cdots, \epsilon_N)$ on the hypersphere (2.11). We have

$$r^* = \left(\sum_{j=1}^{N-1} \epsilon_j \epsilon_{j+1}\right) [(1 - \epsilon_1^2)(1 - \epsilon_N^2)]^{-1/2} \tag{3.1}$$

since

$$\sum_{j=1}^{N-1} \epsilon_j^2 = 1 - \epsilon_N^2 \quad \text{and} \quad \sum_{j=1}^{N-1} \epsilon_{j+1}^2 = 1 - \epsilon_1^2.$$

Inserting the values of ϵ's from (2.11) in terms of x's, using equations (2.4)–(2.8) and neglecting the terms of order $\sin^3\theta$, we have, after some algebraic simplification

$$\begin{aligned} \frac{1 - r^*}{\sin^2\theta} &= 1 + \frac{(1 - \lambda^2)(1 + \lambda^{2N})}{2\lambda^2(1 - \lambda^{2N-2})} + \frac{(1 - \lambda^2)(1 - \lambda^{2N})}{\lambda^2(1 - \lambda^{2N-2})^2} \\ &\cdot \left[\lambda^{N-1} \sum_{k=3}^{N}\sum_{i=3}^{N} b_{i1} b_{kN} \beta_i \beta_k - \frac{\lambda(1 - \lambda^{2N-2})}{1 - \lambda^2} \sum_{j=1}^{N-1}\sum_{k=3}^{N}\sum_{i=3}^{N} b_{ij} b_{k,j+1} \beta_i \beta_k \right. \\ &\left. - \frac{(1 - \lambda^{2N})}{2(1 - \lambda^2)} \frac{\left(\sum_{i=3}^{N} b_{i1}\beta_i\right)^2}{\lambda^2} + \lambda^2 \left(\sum_{i=3}^{N} b_{iN}\beta_i\right)^2\right] \end{aligned} \tag{3.2}$$

As an approximation replace the terms in the square bracket in (3.2) by their expectations. Since $\beta_3, \cdots, \beta_N$ are Cartesian coordinates of a point on the $(N-3)$-dimensional unit sphere

$$\sum_{i=3}^{N} \beta_i^2 = 1,$$

the distribution of $(\beta_3, \cdots, \beta_{N-1})$ is given by [see for example [3] p 385]

$$\frac{\Gamma((N-3)/2)}{\pi^{(N-3)/2}} \frac{d\beta_3 \cdots d\beta_{N-1}}{(1-\beta_3^2 - \cdots - \beta_{N-1}^2)^{1/2}}.$$

From this or from considerations of symmetry we have

$$E\beta_i = 0, \qquad E\beta_i^2 = 1/(N-2), \qquad i = 3, \cdots, N,$$
$$E\beta_i\beta_k = 0, \qquad i \neq k, \qquad i, k = 3, \cdots, N.$$

Rearranging the terms of (3.2) and using the orthogonal property of B, we have after simplification

$$\sin\theta \doteq (1-r^*)^{1/2}\left[1 + \frac{(1-\lambda^2)(1+\lambda^{2N})}{2\lambda^2(1-\lambda^{2N-2})} + \frac{(1-\lambda^{2N-4})(1-\lambda^{2N})}{(N-2)(1-\lambda^{2N-2})^2}\right]^{-1/2}. \tag{3.3}$$

4. Integral expression for $P(r^* \geqq r_0)$. To find the probability that r^* exceeds r_0 where $r_0 < 1$ and close to 1, we proceed in the following manner. For a given λ, r_0 determines a unique positive value of $\sin\theta$, hence a unique value of θ in the interval $[0, \pi/2]$, say $\theta_0(\lambda)$. From (3.3) it follows that for a given λ, $\theta \leqq \theta_0$ implies $r^* \geqq r_0$ and vice versa. By a theorem of Hotelling [2, p. 451] the $(N-1)$-dimensional "area" of a tube of length ds and geodesic radius θ on the surface of the $(N-1)$-dimensional sphere $\sum_{i=1}^{N} \epsilon_i^2 = 1$ is

$$\frac{\pi^{(N-2)/2}}{\Gamma(N/2)} \sin^{N-2}\theta \, ds.$$

The probability that a random point $(\epsilon_1, \cdots, \epsilon_N)$ falls in this elemental tube is the ratio of the $(N-1)$-dimensional area of the tube to the area of the unit sphere. This ratio equals

$$(2\pi)^{-1}\sin^{N-2}\theta \, ds.$$

Remembering that for $r^* = 1$ there correspond two curves on the unit sphere, one traced by P and the other by P' and noting that changing the signs of all ϵ's in (3.1) does not change the value of r^*, we obtain

$$P(r^* \geqq r_0) = \pi^{-1} \int_0^\infty \sin^{N-2}\theta_0 \, ds(\lambda), \tag{4.1}$$

where the variable of integration is λ. This can be written as

$$P(r^* \geqq r_0) \doteq \pi^{-1}(1 - r_0)^{(N-2)/2} \int_0^\infty [g(\lambda)]^{-(N-2)/2} h(\lambda)\, d\lambda, \tag{4.2}$$

where

$$g(\lambda) = 1 + \frac{(1 - \lambda^2)(1 + \lambda^{2N})}{2\lambda^2(1 - \lambda^{2N-2})} + \frac{(1 - \lambda^{2N-4})(1 - \lambda^{2N})}{(N - 2)(1 - \lambda^{2N-2})^2} \tag{4.3}$$

and

$$h(\lambda) = \frac{ds}{d\lambda} = \left[\sum_{i=1}^N \left(\frac{dx_i}{d\lambda}\right)^2\right]^{1/2} = \left[\frac{1}{(1 - \lambda^2)^2} - \frac{N^2\lambda^{2N-2}}{(1 - \lambda^2)^{2N}}\right]^{1/2}. \tag{4.4}$$

We note here that E. S. Keeping [4] has studied the integral of $h(\lambda)$ over the range $[0, \infty]$.

If we change the variable of integration from λ to $1/\lambda$ we observe that the integral in (4.2) remains unchanged, hence the integral from 0 to 1 is the same as from 1 to ∞. Writing J for the integral in (4.2), we have

$$J = 2 \int_0^1 [g(\lambda)]^{-(N-2)/2} h(\lambda)\, d\lambda. \tag{4.5}$$

By considering the sign of the differential coefficient of $g(\lambda)$ in the interval $[0, 1]$ we find that $g(\lambda)$ is a monotonically decreasing function of λ, and

$$g(0) = \infty, \qquad g(1) = [N/(N - 1)]^2.$$

Write

$$\chi(\lambda) = \frac{1}{g(\lambda)}; \tag{4.6}$$

then $\chi(\lambda)$ is a monotonically increasing function of λ in $[0, 1]$ with

$$\chi(0) = 0 \quad \text{and} \quad \chi(1) = (1 - 1/N)^2. \tag{4.7}$$

5. Bounds on the integral J. From (4.5) and (4.6) we have

$$J = 2 \int_0^1 [\chi(\lambda)]^{(N-2)/2} h(\lambda)\, d\lambda. \tag{5.1}$$

Now $\chi(\lambda)$ can be written as

$$\chi(\lambda) = 2\lambda^2 \left(\frac{1 - \lambda^{2N-2}}{1 - \lambda^{2N}}\right)^2 \left[1 + \frac{N\lambda^2(1 - \lambda^{2N-4})}{(N - 2)(1 - \lambda^{2N})}\right]^{-1}. \tag{5.2}$$

Make the transformation

$$\lambda = e^{-x/N} \tag{5.3}$$

in (5.1), then

$$J = \frac{2^{(N-2)/2}}{N}$$

(5.4)
$$\cdot \int_0^\infty \frac{\left(\cosh \frac{x}{N} - \coth x \sinh \frac{x}{N}\right)^{N-2} \left(\operatorname{cosech}^2 \frac{x}{N} - N^2 \operatorname{cosech}^2 x\right)^{1/2} dx}{\left[1 + \frac{N}{N-2}\left(\cosh \frac{2x}{N} - \coth x \sinh \frac{2x}{N}\right)\right]^{(N-2)/2}}.$$

Using elementary expansions of hyperbolic functions in power series, for example,

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots$$

for every x and for $|x| < \pi$,

$$\coth x = \frac{1}{x} + \frac{x}{3} - \frac{x^3}{45} + \frac{2x^5}{945} + \cdots,$$

and after some binomial and exponential expansions, we finally obtain

$$[\chi(e^{-x/N})]^{(N-2)/2} = e^{-1}\left[1 - \frac{x^2}{6} + \frac{x^4}{40} + \cdots\right] + \frac{3e^{-1}}{2N}\left(1 + \frac{2x^2}{9} - \frac{x^4}{30} + \cdots\right) + O(N^{-2}),$$

where this expansion is valid for $|x| < \pi$.
Similarly

$$h(e^{-x/N})\, d(e^{-x/N}) = -\frac{1}{2\sqrt{3}}\left[1 - \frac{x^2}{10} + \frac{135}{12600}x^4 + \cdots + O(N^{-2})\right] dx.$$

We split the range of integration in (5.4) into the ranges [0, 1] and [1, ∞]. Denoting the integral from 0 to 1 by J_1, we have, omitting the terms $O(N^{-1})$,

(5.5)
$$J_1 = 3^{-1/2}e^{-1} \int_0^1 \left(1 - \frac{x^2}{6} + \frac{x^4}{40} + \cdots\right) \cdot \left(1 - \frac{x^2}{10} + \frac{137}{12600}x^4 + \cdots\right) dx = 3^{-1/2}e^{-1}(.9216) = 0.196.$$

Hence

(5.5)
$$J = 0.196 + 2\int_0^{e^{-1/N}} [\chi(\lambda)]^{(N-2)/2} h(\lambda)\, d\lambda.$$

Denote by J_2 the second term on the right hand side of (5.6) and substitute $\lambda = y^{1/2}$ so that

$$J_2 = 2^{(N-2)/2} \int_0^{e^{-2/N}} \frac{y^{(N-3)/2}}{(1-y)} \left(\frac{1-y^{N-1}}{1-y^N}\right)^{N-2} \tag{5.7}$$
$$\cdot \left[1 + \frac{Ny(1-y^{N-2})}{(N-2)(1-y^N)}\right]^{-(N-2)/2} \left[1 - \frac{N^2 y^{N-1}(1-y)^2}{(1-y^N)^2}\right]^{1/2} dy.$$

It can easily be shown that

$$\frac{e^{-1/2}}{(1+y)^{(N-2)/2}} < \left[1 + \frac{Ny(1-y^{N-2})}{(N-2)(1-y^N)}\right]^{-(N-2)/2} < \frac{1}{(1+y)^{(N-2)/2}}. \tag{5.8}$$

The other factors in the integrand can be expanded by the binomial theorem, e.g.

$$\left[1 - \frac{N^2 y^{N-1}(1-y)^2}{(1-y^N)^2}\right]^{1/2} = 1 - \frac{N^2 y^{N-1}(1-y)^2}{2(1-y^N)^2} - \frac{N^4 y^{2N-2}(1-y)^4}{8(1-y^N)^4} - \cdots.$$

We then have

$$e^{-1/2} Q < J_2 < Q, \tag{5.9}$$

where

$$Q = 2^{(N-2)/2} \int_0^{e^{-2/N}} \frac{y^{(N-3)/2}\, dy}{(1+y)^{(N-2)/2}} \tag{5.10}$$
$$\cdot \left[\frac{1}{1-y} - (N-2)y^{N-1} - \frac{N^2}{2} \frac{y^{N-1}(1-y)}{1-y^N} + \cdots\right].$$

We observe that we have to evaluate integrals of type

$$M(p, q, e^{-2/N}) = \int_0^{e^{-2/N}} y^p (1+y)^{-q}\, dy \tag{5.11}$$

and

$$L(p, q, e^{-2/N}) = \int_0^{e^{-2/N}} y^p (1+y)^{-q} (1-y)^{-1}\, dy, \tag{5.12}$$

where $q = (N-2)/2$ and $p = sq + b$, $s > 0$.

Substituting $y = e^{-2/N} z$ and expanding $(1 + z e^{-2/N})^{-q}$ in powers of $(1-z)$ and integrating term by term, we obtain

$$M(p, q, e^{-2/N}) = \frac{e^{-2(p+1)/N}}{(p+1)(1+e^{-2/N})^q} F\left(1, q, p+2, \frac{1}{1+e^{2/N}}\right)$$

and

$$L(p, q, e^{-2/N}) = \frac{e^{-2(p+1)/N}}{(1+e^{-2/N})^q} \sum_{k=0}^{\infty} F\left(1, q, p+k+2, \frac{1}{1+e^{2/N}}\right).$$

If $s > 1$, $b > 0$ and $x > 0$

$$F(1, q, sq + b, x) = 1 + \frac{q}{sq + b} x + \frac{q(q + 1)}{(sq + b)(sq + b + 1)} x^2 + \cdots$$

$$> 1 + \frac{q}{sq + b} x + \frac{q^2}{(sq + b)^2} x^2 + \cdots$$

$$= \left(1 - \frac{qx}{sq + b}\right)^{-1}$$

and

$$F(1, q. sq + b, x) < 1 + \frac{q}{sq + b} x$$

$$+ \frac{q(q + 1)}{(sq + b)(sq + b + 1)} x^2[1 + x + x^2 + \cdots].$$

Since $q = O(N)$, omitting the terms of $O(N^{-1})$ we have

$$\frac{2s}{2s - 1} < F\left(1, q, sq + b, \frac{1}{1 + e^{2/N}}\right) < \frac{1 + s + 2s^2}{2s^2}.$$

A systematic calculation then shows that

$$\frac{.542}{2^{(N-2)/2}} [1 + O(N^{-1})] < L(p, q, e^{-2/N}) < \frac{.629}{2^{(N-2)/2}} [1 + O(N^{-1})].$$

Denoting the integrals of successive terms in (5.10) by Q_1, Q_2, etc., as they occur in order and neglecting the sign, we see that

$$Q_1 = 2^{(N-2)/2} L\left(\frac{N - 3}{2}, \frac{N - 2}{2}, e^{-2/N}\right).$$

Hence

$$0.542 < Q_1 < 0.629.$$

Similar calculations on the following terms show that

$$Q < .629 - .065 - .101 + .029 - .005 = .487$$

and

$$Q > .542 - .066 - .103 + .028 - .006 = .395.$$

The terms diminish very rapidly and the later terms do not affect the second decimal place. Thus from (5.9)

$$.239 < J_2 < .487,$$

and since

$$J = J_1 + J_2 = .196 + J_2$$

therefore

$$.435 < J < .683. \tag{5.13}$$

These calculations are valid to two decimal places and $O(N^{-1})$. Finally, the first term, P_0, in the expansion of $P(r^* \geqq r_0)$ in powers of $(1 - r_0)$ is

$$P_0 = J/\pi(1 - r_0)^{(N-2)/2}.$$

It is easy to see that the first term in the expansion of $P(r^* \leqq -r_0)$ is the same as P_0.

If the population mean is known to be zero, the frequency function of the ordinary correlation coefficient, r, for a sample of size N is given by

$$f(r) = \frac{\Gamma\left(\frac{N-1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{N-2}{2}\right)}(1 - r^2)^{(N-4)/2}.$$

Therefore the first term in the expansion of $P(r \geqq r_0)$ in powers of $(1 - r_0)$ is approximately

$$P \doteq 2^{(N-3)/2}\pi^{-\frac{1}{2}}(N-2)^{-\frac{1}{2}}(1 - r_0)^{(N-2)/2}.$$

Hence

$$P_0/P \doteq 2^{-(N-3)/2}(N-2)^{\frac{1}{2}}\pi^{-\frac{1}{2}}J,$$

which tends to zero as N tends to infinity.

6. Acknowledgement. The author wishes to express his indebtedness to Professor Harold Hotelling for suggesting this problem and for many helpful suggestions during the preparation of this paper.

REFERENCES

[1] Harold Hotelling, "An application of analysis situs to statistics," *Bull. Amer. Math. Soc.*, Vol. 33 (1927), pp. 467–476.

[2] Harold Hotelling, "Tubes and spheres in n-spaces and a class of statistical problems," *Amer. J. Math.*, Vol. 61 (1939), pp. 440–460.

[3] Harald Cramér, *Mathematical Methods of Statistics*, Princeton University Press, 1951.

[4] E. S. Keeping, "A significance test for exponential regression," *Ann. Math. Stat.*, Vol. 22 (1950), pp. 180–198.