THE JOINT DISTRIBUTION OF SERIAL CORRELATION COEFFICIENTS

BY M. H. QUENOUILLE

Rothamsted Experimental Station

1. Summary. An expression for the joint distribution of serial correlation coefficients, circularly defined, has been derived. It has been shown that this distribution possesses properties similar to those already encountered in the distribution of a single serial correlation coefficient, i.e. it is definedby different function forms for various subregions. The distribution thus found is of little use for computational purposes. Consequently, approximate forms have been investigated and the suitability of the ordinary partial correlation coefficient for large-sample testing has been inferred.

2. Introduction. Anderson [1] has derived the distribution of the serial correlation coefficient

$$r_l = \frac{\sum_{i=1}^{n} \epsilon_i \epsilon_{i+l} - \left(\sum_{i=1}^{n} \epsilon_i\right)^2 \Big/ n}{\sum_{i=1}^{n} \epsilon_i^2 - \left(\sum_{i=1}^{n} \epsilon_i\right)^2 \Big/ n},$$

where the ϵ_i are normally and independently distributed with mean μ and variance σ^2 and where a circular definition is employed, so that ϵ_{n+i} is defined to be equal to ϵ_i. However, in making a test of any series, we shall usually be faced with a set of serial correlation coefficients, so that we shall require a joint distribution function of $r_1, r_2, \cdots, r_m$ say. This distribution function is derived below by an extension of the method used by Koopmans [2].

It should be noted that Bartlett [3] has shown that for large samples the variances and covariances of the r_l are independent of the distribution of ϵ_i under fairly wide conditions. This means that the joint distribution function obtained for normal ϵ_i will often give a good approximation for non-normal ϵ_i and can be used as the basis for any test of the correlogram.

3. Conditions on the r_l. It is easily seen that the r_l cannot take all values from $+1$ to -1 independently. For example, r_2 cannot take a value near -1 if r_l takes a value near $+1$. As a result, there will be certain necessary conditions that the r_l will have to fulfil. It is not difficult to find these conditions, since, if $y_i (i = 1, 2, \cdots, n)$ are any set of variables, then

$$(1) \qquad \sum_{j=1}^{n} (\epsilon_{i+j} y_i)^2 = \left(\sum_{i=1}^{n} \epsilon_i^2\right) r_j y_l y_{l+j},$$

where ϵ_i may or may not be corrected for the mean and the double-suffix summation convention is employed.

Thus, provided $0 < m < n/2$, we will have

$$(2) \qquad R_m = \begin{vmatrix} 1 & r_1 & r_2 & \cdots & r_m \\ r_1 & 1 & r_1 & \cdots & r_{m-1} \\ r_2 & r_1 & 1 & \cdots & r_{m-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_m & r_{m-1} & r_{m-2} & \cdots & 1 \end{vmatrix} \geq 0$$

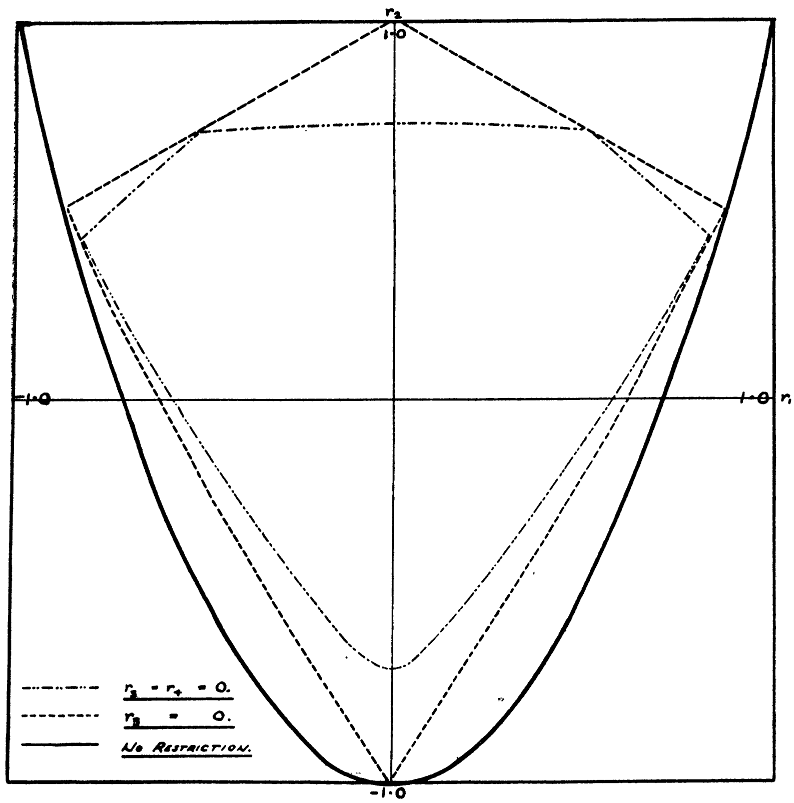

Fig. 1

as a necessary condition that the right-hand side of (1) be positive definite and this expression will impose necessary conditions upon the joint distribution of the r_l.

Fig. 1 gives the limits of possible values of r_1 and r_2 subject to (a) no restriction, (b) $r_3 = 0$, (c) $r_3 = r_4 = 0$.

4. Complex Integration in m Variables. Before finding the joint distribution function of the r_i some introductory remarks on complex integration involving m variables will be necessary.

We can evaluate an integral such as

$$\int_{-\infty}^{\infty} \cdots \int \frac{f(z_1, z_2 \cdots z_m)}{\prod_{j=1}^{m} (z_j - a_j)} \, dz_1 \cdots dz_m$$

where $\mathcal{I}(a_i) = 0$ and $f(z_1, z_2, \cdots, z_m)$ is regular in the region $\mathcal{I}(z_i) \geq 0$, by successive Cauchy integrations, so the integral has a value $(2\pi i)^m f(a_1, \cdots, a_m)$. In the same manner as for Cauchy integration, it will be possible to distort the contours over which we integrate so that we can evaluate

$$\int_S \cdots \int \frac{f(z_1 \cdots z_m)}{\prod_{j=1}^{m} (z_j - a_j)} \, dz_1 \cdots dz_m,$$

provided that $f(z_1, \cdots, z_m)$ is regular in the region defined by S, and $(a_1, \cdots, a_m)$ is enclosed in this region.

More generally, if we have an integral of the form

$$\int_S \cdots \int \frac{f(z_1 \cdots z_m)}{\prod_{j=1}^{m} (a_{ij} z_i - b_j)} \, dz_1 \cdots dz_m,$$

and we make the transformations $w_j = a_{ij} z_i$ and $b_j = a_{ij} c_i$, i.e. $\boldsymbol{W} = \boldsymbol{AZ}$, $\boldsymbol{C} = \boldsymbol{A}^{-1}\boldsymbol{B}$, it is possible, in the above manner, to evaluate the integral as

$$\pm \frac{(2\pi i)^m}{|\boldsymbol{A}|} f(c_1 \cdots c_m). \tag{3}$$

Suppose we now consider the integral

$$\int_S \cdots \int \frac{f(z_1 \cdots z_m)}{\prod_{j=1}^{n} (a_{ij} z_i - b_j)} \, dz_1 \cdots dz_m,$$

where $n \geq m$. We may select a set, g_k, of m equations $a_{ij} z_i = b_j$, and let $\boldsymbol{A}_k = [a_{ij}]$, $\boldsymbol{B}_k = [b_j]$, $\boldsymbol{C}_k = \boldsymbol{A}_k^{-1}\boldsymbol{B}_k = [c_{ik}]$. Then, we may carry out the integration as previously, in this case, summing a series of terms for various combinations of m equations out of the possible n. The value of the integral may then be written

$$(2\pi i)^m \sum_{g_k} \pm \frac{f(c_{1k}, \cdots c_{nk})}{|\boldsymbol{A}_k| \prod_{l \neq g_k} (a_{lj} c_{lk} - b_j)}, \tag{4}$$

where the summation occurs over the points $(c_{ik}, c_{2k}, \cdots, c_{nk})$ lying in the region defined by S, and the product term excludes the set of equations g_k. The ambiguity of sign in (3) and (4) arises from the Jacobian $|\boldsymbol{A}_k|^{-1}$, and the sign must be chosen which makes the transformation of $dz_1, \cdots, dz_m$ yield a positive element. It must be noted that it is possible to obtain several expansions of the form (4) according to the convention that is employed in defining "enclosure" for each of the variables.

5. Integral form for the joint distribution function. We can, without loss of generality, assume $\sigma^2 = 1$. Suppose that

$$p = \sum_{i=1}^{n} \epsilon_i^2 - \left(\sum_{i=1}^{n} \epsilon_i\right)^2 \Big/ n, \qquad q_l = \sum_{i=1}^{n} \epsilon_i \epsilon_{i+l} - \left(\sum_{i=1}^{n} \epsilon_i\right)^2 \Big/ n,$$

where $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ are independent, so that $r_l = q_l/p$. Then by a consideration of n dimensional space, we can see that p is distributed independently of $r_1, \cdots r_m$ so that their joint distribution can be written $g(p)h(r_1, \cdots, r_m)dp\, dr_1, \cdots, dr_m$. The joint distribution of p and $q_1, \cdots, q_m$ can thus be written

$$f(pq_1 \cdots q_m)\, dp\, dq_1 \cdots dq_m = \frac{g(p)}{p^m}\, h\left(\frac{q_1}{p}, \cdots \frac{q_m}{p}\right) dp\, dq_1 \cdots dq_m, \tag{5}$$

where it is not difficult to see that

$$g(p) = \frac{p^{\frac{1}{2}(n-3)} e^{-\frac{1}{2}p}}{2^{\frac{1}{2}(n-1)}\Gamma\left(\frac{n-1}{2}\right)}. \tag{6}$$

We can now find the joint distribution of p and $q_1, \cdots, q_m$ by inverting the characteristic function of these variables. This is given by

$$\frac{1}{(2\pi)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int \exp\left[-\frac{\Sigma\epsilon_j^2}{2} + i(\eta p + \theta_j q_j)\right] d\epsilon_1 \cdots d\epsilon_n,$$

$$= \frac{1}{(2\pi)^{\frac{1}{2}n}} \int_{-\infty}^{\infty} \cdots \int \exp\left[-\frac{\epsilon'\Delta\epsilon}{2}\right] d\epsilon_1 \cdots d\epsilon_n,$$

$$= 1/|\Delta|^{\frac{1}{2}},$$

where

$$\epsilon' = [\epsilon_1, \epsilon_2, \cdots, \epsilon_n]$$

and

$$|\Delta| = \prod_{l=1}^{n-1} (1 - 2i\eta - 2i\theta_j \kappa_{jl}), \qquad \kappa_{jl} = \cos\frac{2\pi_{j\,l}}{n},$$

$$= (1 - 2i\eta)^{n-1} \prod_{l=1}^{n-1} (1 - \kappa_j \kappa_{jl}), \qquad \kappa_j = \frac{2i\theta_j}{1 - 2i\eta},$$

so that the joint distribution of p and $q_1, \cdots, q_m$ is

$$f(p, q_1 \cdots q_m) = \frac{1}{(2\pi)^{m+1}} \int_{-\infty}^{\infty} \cdots \int \frac{1}{|\Delta|^{\frac{1}{2}}} \exp\{-i(\eta p + \theta_j q_j)\}\, d\eta\, d\theta_1 \cdots d\theta_m$$

$$= \frac{1}{(2\pi)^{m+1}} \int_{-\infty}^{\infty} e^{-i\eta p}(1 - 2i\eta)^{-\frac{1}{2}(n-2m-1)} \int_{S_\eta} \frac{1}{|\Delta|^{\frac{1}{2}}} \exp\left\{-\frac{(1 - 2i\eta)\kappa_j q_j}{2}\right\} \frac{d\kappa_1 \cdots d\kappa_m}{(2i)^m}\, d\eta, \tag{7}$$

where S_η is the region bounded by $\kappa_j = \pm \dfrac{2i\infty}{1 - 2i\eta}$. Now S_η can be replaced by region S enclosing the same set of singularities on the real hyperplane, and S can be chosen independent of η. Thus it will be possible to reverse the order of integration in (7) provided that $\int_{-\infty}^{\infty} | 1 - 2i\eta |^{-\frac{1}{2}(n-2m-1)} d\eta$ converges, i.e. provided $n > 2m + 3$. Then since

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} (1 - 2i\eta)^{-\frac{1}{2}(n-2m-1)} \exp\{-i\eta(p - \kappa_j q_j)\}\, d\eta$$

$$= \frac{(p - \kappa_j q_j)^{\frac{1}{2}(n-2m-3)}}{2^{\frac{1}{2}(n-2m-1)}\Gamma\left(\frac{n-2m-1}{2}\right)} \exp\{-\tfrac{1}{2}(p - \kappa_j q_j)\} \quad \text{for} \quad p \geq \kappa_j q_j,$$

$$= 0 \quad \text{for} \quad p \leq \kappa_j q_j,$$

we get

$$f(p, q_1 \cdots q_m) = \frac{e^{-\frac{1}{2}p}}{2^{\frac{1}{2}(n-1)}(2\pi i)^m \Gamma\left(\frac{n-2m-1}{2}\right)} \cdot \int_S \cdots \int \frac{(p - \kappa_j q_j)^{\frac{1}{2}(n-2m-3)}}{\left[\prod_{l=1}^{n-1}(1 - \kappa_j \kappa_{jl})\right]^{\frac{1}{2}}} d\kappa_1 \cdots d\kappa_m, \tag{8}$$

where S encloses the same singularities as S_η all of which lie in the region $p \geq \kappa_j q_j$. If we now use (5) and (6) we get

$$h(r_1 \cdots r_m) = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2m-1}{2}\right)} \cdot \frac{1}{(2\pi i)^m} \int_S \cdots \int \frac{(1 - \kappa_j r_j)^{\frac{1}{2}(n-2m-3)}}{\left[\prod_{l=1}^{n-1}(1 - \kappa_j \kappa_{jl})\right]^{\frac{1}{2}}} d\kappa_1 \cdots d\kappa_m. \tag{9}$$

In a similar manner, it is possible to derive for $n \geq 2m + 3$ the joint distribution of serial correlation coefficients, $\bar{r}_1, \cdots, \bar{r}_m$, uncorrected for the mean, in the form

$$h(\bar{r}_1 \cdots \bar{r}_m) = \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-2m}{2}\right)} \frac{1}{(2\pi i)^m} \int_S \cdots \int \frac{(1 - \kappa_j \bar{r}_j)^{\frac{1}{2}(n-2m-2)}}{\left[\prod_{l=1}^{n}(1 - \kappa_j \kappa_{jl})\right]^{\frac{1}{2}}} d\kappa_1 \cdots d\kappa_m. \tag{10}$$

6. Extension for variables in an autoregressive scheme. Madow [4] has shown how to extend the distribution of the serial correlation coefficient for uncorrelated variables to the case when the variables x_i are connected by a linear Markoff scheme, $x_i = \rho x_{i-1} + \epsilon_i$ with a normal distribution of the error ϵ_i. It is worth noting that the method used by Madow can be applied to derive the joint distribution of serial correlations of variables x_i, which are connected by a linear autoregressive scheme of order m, or less,

$$a_0 x_i + a_1 x_{i-1} + \cdots + a_m x_{i-m} = \epsilon_i,$$

where $\epsilon_1, \cdots, \epsilon_n$ are normally and independently distributed, and $\epsilon_{n+i} = \epsilon_i$.[1] Under these conditions, the expression (9) will be modified by a factor

$$(11) \qquad \left(\frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} \epsilon_i^2}\right)^{\frac{1}{2}(n-1)} = \frac{1}{(A + 2B_j r_j)^{\frac{1}{2}(n-1)}},$$

where

$$A = \sum_{k=0}^{m} a_k^2,$$

$$B_j = \sum_{k=0}^{m-j} a_k a_{k+j},$$

while (10) will be modified by a similar factor with n replacing $n - 1$.

7. Reduction of the distribution function integral. Using the method described in section 4, it is now possible to reduce the integral given in (9), if we observe that $\kappa_{jl} = \kappa_{jn-l}$ and assume n odd. We then have

$$h(r_1 \cdots r_m) = \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n-2m-1}{2}\right)} \frac{1}{(2\pi i)^m} \int_S \cdots \int \frac{(1 - \kappa_j r_j)^{\frac{1}{2}(n-2m-3)}}{\prod_{l=1}^{\frac{1}{2}(n-1)} (1 - \kappa_j \kappa_{jl})} d\kappa_1 \cdots d\kappa_m$$

$$(12) \qquad = \frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-2m-1}{2}\right)} \sum_{g_k} \frac{\begin{vmatrix} 1 & I \\ r & K_k \end{vmatrix}^{\frac{1}{2}(n-2m-3)}}{\prod_{l \neq g_k} \begin{vmatrix} 1 & I \\ \kappa_{jl} & K_k \end{vmatrix}},$$

where $I = (1, 1, \cdots, 1)$, $r' = (r_1, r_2, \cdots, r_m)$, $\underline{\kappa}_{jl} = (\kappa_{jl}, \cdots, \kappa_{ml})$ and K_k is the matrix formed from a set g_k of the m matrices $\underline{\kappa}_{jl}$ arranged in order. The factors in the summation can most easily be determined if we put $\begin{vmatrix} 1 & I \\ r & K_k \end{vmatrix} \propto A(r_1, \cdots, r_{m-1}) - r_m$ and sum over the region for which $r_m \leq A(r_1, \cdots, r_{m-1})$. To demonstrate the manner in which formula (12) works, we shall consider $m = 2$. From formula (2) we can see that a limit to the possible values that r_2 can take is given by $r_2 = 2r_1^2 - 1$ i.e. by the curve $(\cos\theta, \cos 2\theta)$

[1] This is a sufficient condition for $x_{n+i} = x_i$.

in the (r_1, r_2) plane. It is not difficult to see that there are ${}^nC_2$ possible terms in (12) and that each of these terms is proportional to the $\frac{1}{2}(n - 2m - 3)$th power of the distance from a line in the (r_1, r_2) planes. These lines are the joins of the points $(\cos 2\pi i/n, \cos 4\pi i/n)$, $i = 1, \cdots, \frac{1}{2}(n - 1)$ and the joins of such points on the curve $(\cos \theta, \cos 2\theta)$ give the outer limits of the possible values of r_1 and r_2. It can also be seen that these points correspond to the equations $\kappa_j \kappa_{jl} = 1$ (each of these equations determines a plane in 4-dimensional complex space), while the joins of these points correspond to the singularities defined by and terms arising from pairs of these equations. Furthermore, since the sum of residues in any plane is zero, the sum of contributions, taken with appropriate signs, arising from lines through any of these points is zero, i.e. the sum of all possible terms involving any particular κ_{jl} will disappear. This leads to several possible expansions for $h(r_1, \cdots, r_m)$.

If we consider the particular case $n = 9$, then each term in the expansion (12) is proportional to the distance from one of the lines joining $(\cos 2\pi i/9 \cos 4\pi i/9)$, $i = 1, 2, 3, 4$. These lines may be denoted by l_{ij}. Then the contribution from l_{ij} is given by

$$3 \frac{\kappa_{1i}\kappa_{1j} - (\kappa_{1i} + \kappa_{1j})r_1 + \frac{1}{2}(r_2 + 1)}{(\kappa_{1j} - \kappa_{1i})(\kappa_{1i} - \kappa_{1k})(\kappa_{1i} - \kappa_{1l})(\kappa_{1j} - \kappa_{1k})(\kappa_{1j} - \kappa_{1l})},$$

where $j > i$ and $\kappa_{1\alpha} = \cos \dfrac{2\pi\alpha}{q}$.

The values of this expression are:

$$\begin{aligned}
&l_{12}, && -1.979 + 2.938\, r_1 - 1.563\, r_2, \\
&l_{13}, && 0.926 - 2.106\, r_1 + 3.959\, r_2, \\
&l_{14}, && 1.053 - 0.832\, r_1 - 2.396\, r_2, \\
&l_{23}, && -5.012 - 3.959\, r_1 - 6.065\, r_2, \\
&l_{24}, && 3.033 + 6.897\, r_1 + 4.502\, r_2, \\
&l_{34}, && -4.086 - 6.065\, r_1 - 2.106\, r_2,
\end{aligned}$$

where, for example, the contribution from l_{12} acts in the region for which $1.563\, r_2 \leq -1.979 + 2.938\, r_1$. Fig. 2 demonstrates the configuration for this case. It is seen that the frequency surface is a tetrahedron. As particular examples of the identities mentioned above we have

$$\begin{aligned}
l_{12} + l_{13} + l_{14} &= 0, \\
-l_{12} + l_{23} + l_{24} &= 0, \\
-l_{13} - l_{23} + l_{34} &= 0.
\end{aligned}$$

For a general value of m, we shall find that the hyperplanes joining sets of m points $(\cos 2\pi i/n, \cos 4\pi i/n, \cdots, \cos 2\pi m i/n)$ will be singularities on the frequency hypersurface. The hyperplanes passing through sets of m successive points will give the limits of possible values of $r_1, \cdots, r_m$. Furthermore, the sum of contributions (with appropriate signs) to the frequency function from the set of $\frac{1}{2}(n - 2m + 1)$ hyperplanes passing through any point will be zero.

8. Integral approximation for the distribution function. The expression (12) is, of course, difficult to use in practice and we require an approximation similar to that of Koopmans. For this we make use of the integral expression (10)

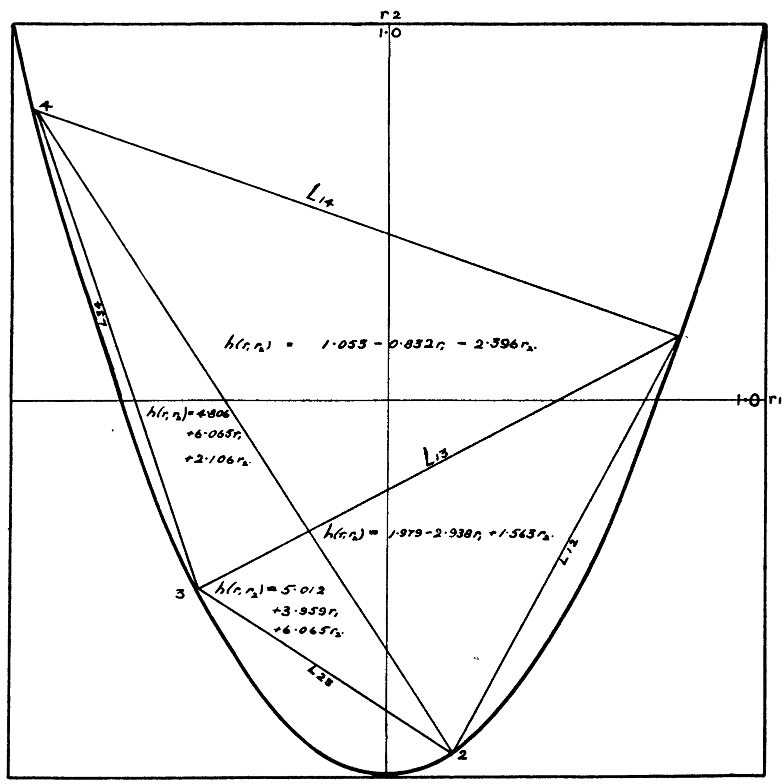

FIG. 2

for the joint distribution function of $\bar{r}_1, \cdots, \bar{r}_m$ and approximate to the factor $\left[\prod_{l=1}^{n} (1 - \kappa_j \kappa_{jl})\right]^{-\frac{1}{2}}$. This can be done without undue difficulty, but the resulting multiple integral does not appear to be capable of easy reduction. This is hardly surprising, since from the nature of the distribution of the r_i we should expect this approximation to involve R_m raised to a suitable power, and this conjecture is strengthened by the following considerations:

a) The distribution of $\bar{r}_1$ may be obtained by considering the two sets of observations $x_1, x_2, \cdots, x_{n-1}, x_n$ and $x_2, x_3, \cdots, x_n, x_1$ as unrelated, and using

the distribution of the ordinary correlation coefficient corresponding to $n + 3$ pairs of observations. (Dixon [6] Quenouille [7]). In the same manner, the m sets of observations $x_1, x_2, \cdots, x_{n-1}, x_n; x_2, x_3, \cdots, x_n; \cdots x_m, x_{m+1}, \cdots, x_{m-2}, x_{m-1}$, might be considered as unrelated and the joint distribution of their correlations, given by Garding (5), will involve R_m raised to a suitable power.
b) The outer limits for the joint distribution of $r_1, r_2, \cdots, r_m$ or $\bar{r}_1, \bar{r}_2, \cdots, \bar{r}_m$ for large n, will be provided by the equations $R_p = 0$, $(p = 1, \cdots, m)$. An investigation of the properties of the functions, $R_1, R_2, \cdots, R_m$ might therefore be expected to throw light upon the joint distribution of $r_1, r_2, \cdots, r_m$ or $\bar{r}_1, \bar{r}_2, \cdots, \bar{r}_m$.
c) R_p is a quadratic in r_p and may be put equal to $R_{p-2}(r'_p - r_p)(r_p - r''_p)$, where r'_p and r''_p are functions of $r_1, r_2, \cdots, r_{p-1}$ giving the limits of the values that r_p can take for any particular values of $r_1, \cdots, r_{p-1}$. Let $Q_p = R_p/R_{p-1}$, then Q_p is likewise a quadratic in r_p, taking all values between r'_p and r''_p and

$$\int_{r''_p}^{r'_p} Q_p^s \, dr_p = \frac{R_{p-2}^s}{R_{p-1}^s} \int_{r''_p}^{r'_p} (r'_p - r_p)^s (r_p - r''_p)^s \, dr_p$$

$$= \frac{B(s + 1, \frac{1}{2})}{Q_{p-1}^s} \cdot \left(\frac{r'_p - r''_p}{2}\right)^{2s+1}.$$

But, by expanding R_p as a bordered determinant, it is not difficult to show that $r'_p - r''_p = 2Q_{p-1}$, so that

$$\int_{r''_p}^{r'_p} Q_p^s \, dr_p = \frac{\Gamma(s + 1)}{\Gamma(s + \frac{3}{2})} \cdot \pi^{\frac{1}{2}} \cdot Q_{p-1}^{s+1}.$$

In particular, if

$$f(r_1 \cdots r_m) = \frac{\Gamma(\frac{1}{2}n + 1) \cdots \Gamma(\frac{1}{2}n - m + 2)}{\Gamma(\frac{1}{2}n + \frac{1}{2}) \cdots \Gamma(\frac{1}{2}n - m + \frac{3}{2})} \cdot \frac{1}{\pi^{m/2}} Q_m^{\frac{1}{2}(n-2m+1)}, \tag{13}$$

and if we integrate with respect to $r_m, r_{m-1}, \cdots r_2$ in turn, we get

$$\int_{r''_2}^{r'} \cdots \int_{r''_m}^{r'_m} f(r_1 \cdots r_m) \, dr_m \cdots dr_2 = \frac{\Gamma(\frac{1}{2}n + 1)}{\Gamma(\frac{1}{2}n + \frac{1}{2})\pi^{\frac{1}{2}}} (1 - r^2)^{\frac{1}{2}(n-1)},$$

which is the approximate distribution of the first serial correlation coefficient, uncorrected for the mean, as given by Dixon [6].

The importance of this lies in the fact that the integral corresponding to that of Koopman's for the joint distribution is

$$\frac{\Gamma(\frac{1}{2}n)}{\Gamma(\frac{1}{2}n - m)} \cdot \frac{2^{\frac{1}{2}n}}{\Pi^m} \int_S \cdots \int \frac{\begin{vmatrix} 1 & \underline{I} \\ r & \mathbf{X} \end{vmatrix}^{\frac{1}{2}n-m-1}}{|\mathbf{Y}|^{\frac{1}{2}n}} \prod_{i=1}^{m} \left[\sin \tfrac{1}{2}nx_i \begin{vmatrix} & 0 & \\ \dfrac{d}{dx_i} & \kappa(x_i) & \begin{matrix} I \\ X \end{matrix} \end{vmatrix} \right] dx_1 \cdots dx_m$$

where $r' = [r_1, \cdots, r_m]$,

$$\mathbf{X} = \begin{bmatrix} \cos x_1 & \cos x_2 & \cdots & \cos x_m \\ \cos 2x_1 & \cos 2x_2 & \cdots & \cos 2x_m \\ \cdots & \cdots & \cdots & \cdots \\ \cos mx_1 & \cos mx_2 & \cdots & \cos mx_m \end{bmatrix},$$

$$\mathbf{I} = [1, 1, \cdots 1],$$

$$\mathbf{Y} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \cos x_1 & \cos x_2 & \cdots & \cos x_m \\ \cdots & \cdots & \cdots & \cdots \\ \cos (m-1)x_1 & \cos (m-1)x_2 & \cdots & \cos (m-1)x_m \end{bmatrix},$$

$$\kappa'(\theta) = [\cos \theta, \cos 2\theta, \cdots, \cos m\theta],$$

and S is the region given by $\begin{vmatrix} 1 & I \\ r & X \end{vmatrix} \geq 0$. This suggests, by analogy, that the joint distribution function is a polynomial in r_m of degree $2(\frac{1}{2}n - m - 1) + 3 = n - 2m + 1$ which vanishes only when $R_m = 0$. The equation satisfies these conditions, and in addition, it reduces to the known form when $m = 1$ and can be integrated to give this same form. Thus there is a strong suggestion that (13) gives an approximate distribution of $r_1, r_2, \cdots, r_m$, uncorrected for the mean.

An alternative form for the constant factor in (13) may be obtained if we note that

$$\frac{\Gamma(\frac{1}{2}n - m + 2)}{\Gamma(\frac{1}{2}n - m + \frac{3}{2})\pi^{\frac{1}{2}}} = \frac{1}{2^{n-2m+2}} \frac{\Gamma(n - 2m + 3)}{[\Gamma(\frac{1}{2}n - m + \frac{3}{2})]^2}.$$

d) Now r_p' and r_p'' can be written in the forms $(S_{p-1} + R_{p-1})/R_{p-2}$ and $(S_{p-1} - R_{p-1})/R_{p-2}$, where

$$S_{p-1} = (-1)^{p-1} \begin{vmatrix} r_1 & r_2 & r_3 & \cdots & 0 \\ 1 & r_1 & r_2 & \cdots & r_{p-1} \\ r_1 & 1 & r_1 & \cdots & r_{p-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{p-2} & r_{p-3} & r_{p-4} & \cdots & r_1 \end{vmatrix}.$$

Thus

$$R_p = R_{p-2}\left(\frac{S_{p-1} + R_{p-1}}{R_{p-2}} - r_p\right)\left(r_p - \frac{S_{p-1} - R_{p-1}}{R_{p-2}}\right)$$

$$= \frac{R_{p-1}^2}{R_{p-2}}\left[1 - \left(\frac{r_p R_{p-2} - S_{p-1}}{R_{p-1}}\right)^2\right]$$

$$Q_p = Q_{p-1}(1 - r_{1p+1.23\cdots}^2)$$

where $$r_{1,p+1.23\cdots} = T_{p-1}/R_{p-1},$$

and

$$T_{p-1} = \begin{vmatrix} r_p & r_1 & r_2 & \cdots & r_{p-1} \\ r_{p-1} & 1 & r_1 & \cdots & r_{p-2} \\ r_{p-2} & r_1 & 1 & \cdots & r_{p-3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_1 & r_{p-2} & r_{p-3} & \cdots & 1 \end{vmatrix}.$$

Therefore, if we make a change of variable to $r_{1,p-1.23\cdots}$, $r_{1,p.23\cdots}$, $\cdots$ $r_{13.2}$, r_1, we find that the new variables which correspond exactly to partial correlation coefficients are, in fact, independently distributed as such, with 3 degrees of freedom more than in the case where the sets of variables are distinct observations.

While the above properties do not prove that the r_i or $\bar{r}_i$ may be tested using partial or multiple correlation coefficients, this conjecture has been verified elsewhere and it has been shown [8] that, with certain adjustments, a test can be derived which is applicable to fairly short series.

REFERENCES

[1] R. L. Anderson, "Distribution of the serial correlation coefficient," *Annals of Math. Stat.*, Vol. 13 (1942), pp. 1–13.

[2] T. Koopmans, "Serial correlation and quadratic forms in normal variables," *Annals of Math. Stat.*, Vol. 13 (1943), pp. 14–33.

[3] M. S. Bartlett, "On the theoretical specification of sampling properties of autocorrelated time series," *Roy. Stat. Soc. Suppl.*, Vol. 8 (1946), pp. 27–41.

[4] W. G. Madow, "Note on the distribution of the serial correlation coefficient," *Annals of Math. Stat.*, Vol. 16 (1945), pp. 308–310.

[5] L. Garding, *Proceedings of Lund University Mathematical Seminars*, Vol. 5, pp. 185–202.

[6] W. J. Dixon, "Further contributions to the problem of serial correlation," *Annals of Math. Stat.*, Vol. 15 (1944), pp. 119–144.

[7] M. H. Quenouille, "Some results in the testing of the serial correlation coefficient," *Biometrika*, Vol. 35 (1948), pp. 261–7.

[8] M. H. Quenouille, "Approximate tests of correlation in time series 1," *Roy. Stat. Soc. Suppl.*, Vol. 11 (1949).