The Annals of Mathematical Statistics
1969, Vol. 40, No. 1, 188–196

ON SERIAL CORRELATION

BY S. JAMES PRESS

University of Chicago

1. Introduction and summary. Define

$$P = U/V \equiv \sum_{i=1}^{m} \lambda_i x_i^2 / \sum_{i=1}^{r} x_i^2, \qquad 1 \leqq m \leqq r, \tag{1.1}$$

where the x_i's are independent and identically distributed as $N(0, 1)$, and $0 < \lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_m$. We will give various representations for the distribution of P, and show how a special case of this distribution is useful for testing for correlation in a time series. We will also consider independence of the errors in the normal regression problem.

Let n be the number of observations, and k the number of coefficient parameters in a univariate linear regression model. Take $r = n - k$ and $m = r - 1$. Suppose the x_i's are the linear functions of the observations of the dependent variable obtained by a "Theil transformation" (see Theil [15]). It will be shown in Section 2 that under these conditions, P is an appropriate test statistic for independence.

In Section 3, two different characteristic function representations are developed for P; one involves an infinite series of complex valued gamma functions, while the other involves a doubly infinite series of real valued gamma functions.

Section 4 discusses the fact that for $r = m + 1$, P is distributed as a linear combination of correlated beta variates. An appropriate beta distribution approximation is given for comparison.

A numerical tabulation of the distribution of P for the case of $r = m + 1$ is in preparation. The cdf. of P is found by numerically inverting the characteristic function of a related linear combination of central chi-square variates.

2. Testing for serial correlation.

2.1 *History of the problem.* Analysis of the problem associated with serial correlation has had a long history. For this reason it is appropriate at this point to review briefly some of the efforts which have preceded this one.

The difficulties (such as inefficient estimators) which can arise as a result of the presence of serial correlation in a problem were considered at least as early as 1921 by Yule [21], and more recently (1949) by Cochrane and Orcutt [5]. Proposals for handling the testing problem really began with a series of papers involving John von Neumann. The series began with a preliminary study involving the need for the work, and examination of moments of an appropriate test statistic by von Neumann, Kent, Bellinson, and Hart [16] in 1941. That work was immediately followed by development of appropriate distribution

Received 6 February 1968.

theory in von Neumann [17], 1941; and von Neumann [18], 1942. Finally, in von Neumann and Hart [19], 1942, the pertinent distribution (discussed below) was tabulated.

Let $w = (w_1, \cdots, w_n)$ be a vector of n observations from some normal process. Let

$$\bar{w} = n^{-1} \sum_1^n w_i, \qquad s^2 = n^{-1} \sum_1^n (w_i - \bar{w})^2$$

denote the sample mean and variance, respectively. Consider the problem of testing the hypothesis:

H: all w_k's are mutually independent and homoscedastic, vs. the alternative, A: the covariance matrix of w is arbitrary (nonscalar).

These are the only hypotheses considered throughout the present paper. To test H vs. A, the work of von Neumann suggests use of the statistic

$$Q_{n-1} = \sum_1^{n-1} (w_{i+1} - w_i)^2 / \sum_1^n (w_i - \bar{w})^2,$$

where the subscript of Q denotes the number of degrees of freedom of the denominator quadratic form. The rationale is that $E(Q_{n-1}) = 2(1 - n^{-1})$ under H (or about 2 in large samples) and $E(Q_{n-1})$ can vary between about 0 and 2, or between 2 and 4, under A, depending upon the actual relationship among the w_k's. Therefore, an observed value of Q_{n-1} close to 2 militates in favor of H. Von Neumann derived an expression for the $(\frac{1}{2}(n - 1) - 1)$ derivative of the density of Q_{n-1}, under H, in the case of n odd, and developed integral relationships for expressing the $(\frac{1}{2}(n - 1) - 1)$ derivative, for n even, in terms of the results for n odd.

Koopmans [9], 1942, considered a related problem and found the density of Q_{n-1} in explicit form in terms of an integral.

R. L. Anderson [1], 1942, in his thesis, treated the serial correlation problem for the special case in which the covariance matrix is circular, and R. L. Anderson and T. W. Anderson, in 1950 [2], showed how the distribution of the sample circular serial correlation coefficient can be used in certain types of regression problems.

Dixon [6] took up R. L. Anderson's problem in 1944 and developed the moments of the circular serial correlation coefficient (in addition to moments of related serial correlation statistics).

T. W. Anderson [3], 1948, examined the power of tests based on the ratio of quadratic forms in normal variates against certain alternatives. He showed, for example, that no test is uniformly most powerful against the alternative that the errors follow a simple first order autoregressive scheme.

In 1940, Mauchly [10] considered the same problem when repeated observations are available (which is not usually the case in time series analysis).

In 1967, Bloch and Watson [4] studied the posterior distribution of the cell probabilities in a multinomial distribution. It will be seen in Section 4 that the distribution of a linear combination of the cell probabilities is related to that of P, defined in (1.1).

Moran [11], 1950, considered the special case of serial correlation in a regression on a single independent variable.

Durbin and Watson in 1950 and 1951 [7], [8], reexamined the serial correlation problem specifically for the case of multiple regression. They treated the statistic (1950), p. 424.

$$Q_n = \sum_1^{n-1} (w_{i+1} - w_i)^2 / \sum_1^n w_i^2, \tag{2.1}$$

where now the w_i's are the residuals in a regression (Q_n is reduced to our canonical form in Section 2.3). Although they were unable to find exact significance levels (the principal problem being that the usual residuals are correlated regardless of whether or not the disturbances in the regression are correlated), they did find and tabulate bounds for the significance levels which depend upon the sample size and the number of independent variables in the regression. These results do not yield positive results in all cases; i.e., sometimes the user of the test is left in doubt as to whether or not to accept H.

In 1961, Theil and Nagar [14] showed that under certain conditions the Durbin-Watson procedure could be simplified. However, the problem of treating the general situation unambiguously remained unsolved.

In 1965, Theil [15] provided the building blocks for another approach towards resolution of the fundamental difficulty in testing for serial correlation in regression problems, and it was this work which inspired the present investigation. The relationship between these earlier efforts and the work of Theil is briefly summarized below.

2.2 *Testing for serial correlation in regression.* Let

$$y = X\beta + u$$

denote the model equation for a univariate multiple regression problem involving an $n \times 1$ vector of dependent variables y, a non-stochastic $n \times k$ design matrix X, a $k \times 1$ vector of parameters β, and an $n \times 1$ vector of disturbances or errors u. For inferences about β, it is usually assumed that $\mathcal{L}(u) = N(0, \sigma^2 \mathrm{I})$ for I the identity matrix. Under these conditions it follows that if $\hat{\beta}$ is the least squares estimator, and if

$$\hat{u} = y - X\hat{\beta}$$

is the vector of residuals,

$$E(\hat{u}\hat{u}') = \sigma^2 M,$$

where $M = I - X(X'X)^{-1}X'$. Hence the elements of $\hat{u}$ are generally correlated (even if $E(uu') = \sigma^2 I$).

To eliminate the correlation under H, and thereby simplify the problem of testing for serial correlation, Theil [15] introduced the $(n - k) \times 1$ vector $x = A^* y$ (see [15] for the computation of A^*), which approximates the last $(n - k)$ elements of u, and under H has the properties:

$$E(x) = 0, \qquad E(xx') = \sigma^2 I.$$

Use of the statistic Q_n, defined in (2.1), with n replaced by $(n - k)$, will yield a test of H vs. A for the quantities $x_1, \cdots, x_{n-k}$. However, we were originally interested in testing H vs. A for the regression disturbances. Fortunately for n large, and much larger than k, this test is approximately a test of H (sphericity) vs. A for the regression errors, because then the x vector approximates the last $(n - k)$ elements of u. However, even for small n, if the y_i's are independent, so are the x_i's. Hence, if the sphericity test for the x_i's requires rejection of this hypothesis, the hypothesis of sphericity for the errors, u_i, must also be rejected.

2.3 *Reduction to canonical form.* The Q_n statistic defined in (2.1) may be placed in the canonical form given in (1.1). Define

$$U^* = \sum_{i=1}^{n-1} (w_{i+1} - w_i)^2,$$

and assume that the w_i's are independent $N(0, \sigma^2)$. For example, if the w_i's are the result of a Theil transformation on the residuals in a regression, the w_i's will satisfy this assumption.

Let $w' \equiv (w_1, \cdots, w_n)$, $y' \equiv (y_1, \cdots, y_{n-1})$, and

$$G = \begin{pmatrix} -1 & 1 & & & 0 \\ 0 & \ddots & \ddots & & \\ & & & & \\ & & & -1 & 1 \end{pmatrix}, \quad G: (n-1) \times n,$$

where all elements not on the two diagonals depicted are zeros (primes denote transpose). Since $U^* = (Gw)'(Gw)$, if $y = Gw$, $U^* = y'y = w'G'Gw = w'Bw$, where $B = G'G$. Note that $|B| = 0$ and rank $(B) = n - 1$. The square matrix $B: n \times n$ is given by

$$B = \begin{bmatrix} 1 & -1 & & & 0 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 2 & -1 \\ 0 & & & -1 & 1 \end{bmatrix},$$

so that all elements off the main, sub, and super diagonals are zeros. It is well known [17] that the ordered, nonzero, eigenvalues of B are given by

$$\lambda_k = 2(1 - \cos \pi k n^{-1}), \quad k = 1, \cdots, n-1. \tag{2.2}$$

Note that the λ_k's are all distinct. Since B is symmetric there exists an orthogonal matrix Γ for which

$$B = \Gamma' D_\lambda \Gamma, \quad \Gamma'\Gamma = I, \quad D_\lambda \equiv \text{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_{n-1}, 0).$$

Therefore, if $x' \equiv (x_1, \cdots, x_n)$, and $x = \Gamma w$,

$$U^* = w'Bw = x'D_\lambda x = \sum_1^{n-1} \lambda_i x_i^2,$$

where $\mathcal{L}(x) = N(0, \sigma^2 I)$. If $V^* \equiv \sum_1^n w_i^2 = w'w = x'x = \sum_1^n x_i^2$, it is clear that $U^* = U$, $V^* = V$, and

$$Q_n = U^*/V^* = U/V = P.$$

Moreover, since σ^2 may be taken equal to unity without loss of generality (because P is independent of the units of the x_i's), Q_n has the canonical form (1.1) for $m \equiv n - 1$, $r \equiv n$. Thus, if in a regression, the y_i's are modified according to a Theil transformation, and the statistic Q_{n-k} is computed, its significance can be tested (test of H vs. A) by using the distribution of P for $m + 1 = r \equiv n - k$, with the λ_k's defined by (2.2). Note that in this case (and for all cases in which $m < r$),

$$0 < P < \lambda_m .$$

For $m = r$, however, $\lambda_1 < P < \lambda_m$. Of course, for cases in which the latent roots are not distinct, the bounds on these inequalities may be achieved.

3. Characteristic function representations for P. In this section we consider the λ_i's as ordered and positive, but otherwise arbitrary. We adopt the canonical form given in (1.1) and find two distinct characteristic function representations for the distribution of P.

First note that V is a chi-square variate with r degrees of freedom (d.f.); i.e., $\mathfrak{L}(V) = \chi^2(r)$. Then note that U is a linear combination of chi-square variates each of which has one d.f.

Define

$$a_k = \lambda_{k+1}\lambda_1^{-1}, \qquad k = 1, \cdots, m - 1,$$

so that $a_k \geqq 1$. Then, if $U_1 = U/\lambda_1$,

$$U_1 = x_1^2 + a_1 x_2^2 + \cdots + a_{m-1} x_m^2.$$

Let $f(U_1)$ denote the density of U_1 and let $f_k(t)$ denote the density of a chi-square variate with k d.f. Then [13], we can write the density of U_1 as a linear combination of chi-square densities:

$$f(U_1) = \sum_{i=0}^{\infty} q_i f_{m+2i}(U_1), \tag{3.1}$$

where the q_i's are constants depending upon m and the λ_i's. (In the special case in which the λ_i's are given by (2.2), the λ_i's are in turn defined in terms of n.) The q_i's are determined from the identity in w:

$$\sum_{k=0}^{\infty} q_k w^k \equiv \prod_{i=1}^{m-1} [a_i - (a_i - 1)w]^{-\frac{1}{2}}, \qquad |w| \leqq 1. \tag{3.2}$$

An explicit representation (recursive) for the q_k's, useful for computational purposes, was given in [13] as follows. Define

$$q_{i,j} = a_i^{-\frac{1}{2}} \cdot (\tfrac{1}{2})(\tfrac{1}{2} + 1) \cdots (\tfrac{1}{2} + j - 1)(j!)^{-1} \cdot (1 - a_i^{-1})^j,$$

for $i = 1, \cdots, m - 1$, $j = 0, 1 \cdots$. Then the q_j's may be computed stepwise from the relations

$$q_j = q_j^{(m-1)}, \qquad q_j^{(s)} = \sum_{i=0}^{j} \{q_{j-i}^{(s-1)} \cdot q_{s,i}\}, \qquad s = 2, \cdots, m - 1,$$

and $q_j^{(1)} = q_{1,j}$. From (3.2) it is not hard to see that $q_i \geqq 0$, and $\sum_0^{\infty} q_i = 1$.

3.1 *Single infinite series representation.* It is well known in the relation $P = U/V$ that P and V are statistically independent (see [12], and [17]). The proof follows immediately after conversion of the x_i's in P to polar coordinates, noting that P depends only upon the angles (and not the modulus), V depends upon the modulus only, and that the modulus and angles are independent.

Let $\quad P_1 = U_1/V, \quad \tilde{U}_1 = \ln U_1, \quad \tilde{V} = \ln V$, and $\tilde{P}_1 = \ln P_1$.

Then,

$$\tilde{U}_1 = \tilde{P}_1 + \tilde{V}, \tag{3.3}$$

where $\tilde{P}_1$ and $\tilde{V}$ are independent. If we denote the characteristic function of a random variable W by $\phi_W(t) \equiv E \exp(itW)$, it follows that $\phi_{\tilde{U}_1}(t) = \phi_{\tilde{P}_1}(t) \cdot \phi_{\tilde{V}}(t)$, or

$$\phi_{\tilde{P}_1}(t) = \phi_{\tilde{U}_1}(t)/\phi_{\tilde{V}}(t). \tag{3.4}$$

The characteristic function of $\tilde{V}$ is evaluated simply in terms of the density of V.

$$\phi_{\tilde{V}}(t) = E(e^{it\ln V}) = \int_0^\infty e^{it\ln V} c e^{-\frac{1}{2}V} V^{\frac{1}{2}r-1}\, dV,$$

where $c = 2^{-\frac{1}{2}r}\Gamma^{-1}(\frac{1}{2}r)$. Simplification gives

$$\phi_{\tilde{V}}(t) = c(2)^{\frac{1}{2}r+it}\Gamma(\tfrac{1}{2}r + it). \tag{3.5}$$

The characteristic function of $\tilde{U}_1$ is evaluated by using (3.1). Thus,

$$\phi_{\tilde{U}_1}(t) = \int_0^\infty e^{it\ln U_1} \sum_{j=0}^\infty q_j f_{m+2j}(U_1)\, dU_1 .$$

Since we may interchange order of summation and integration [13],

$$\phi_{\tilde{U}_1}(t) = \sum_{j=0}^\infty c_j q_j \int_0^\infty e^{it\ln U_1} U_1^{\frac{1}{2}(m+2j)-1} e^{-\frac{1}{2}U_1}\, dU_1 ,$$

where $c_j = 2^{-\frac{1}{2}(m+2j)}\Gamma^{-1}(\frac{1}{2}(m + 2j))$. Simplifying,

$$\phi_{\tilde{U}_1}(t) = \sum_{j=0}^\infty c_j q_j (2)^{\frac{1}{2}(m+2j)+it}\Gamma(\tfrac{1}{2}(m + 2j) + it). \tag{3.6}$$

Substituting (3.5) and (3.6) into (3.4) and combining terms gives

$$\phi_{\tilde{P}_1}(t) = [\Gamma(\tfrac{1}{2}r)/\Gamma(it + \tfrac{1}{2}r)] \sum_{j=0}^\infty \Gamma[it + \tfrac{1}{2}(m + 2j)]/\Gamma(\tfrac{1}{2}(j + m)). \tag{3.7}$$

Let $F(x)$ denote the cdf of P. Then, since

$$P = U/V = U_1\lambda_1 V^{-1} = \lambda_1 P_1 = \lambda_1 e^{\tilde{P}_1}, \tag{3.8}$$

$$F(x) = P\,\{P \leqq x\} = P\{\lambda_1 e^{\tilde{P}_1} \leqq x\} = P\{\tilde{P}_1 \leqq \ln x\lambda_1^{-1}\}.$$

Hence, to establish the percentage points of $F(x)$, it is sufficient to evaluate the cdf of $\tilde{P}_1$.

3.2. *Doubly infinite series representation.* A doubly infinite series representation for the characteristic function of P may be obtained by using the moments of the distribution.

From the independence of P and V (see beginning of Section 3.1), the sth moment of P is given by

(3.9) $$p_s \equiv E(P^s) = E(U^s)/E(V^s).$$

Since $U = U_1\lambda_1$, from (3.1),

$$E(U^s) = \lambda_1^s \int_0^\infty \sum_{j=0}^\infty q_j U_1^s \cdot c_j U_1^{\frac{1}{2}(m+2j)-1} e^{-\frac{1}{2}U_1}\, dU_1 .$$

Since the interchange of summation and integration is justified, we find

(3.10) $$E(U^s) = (2\lambda_1)^s \sum_{j=0}^\infty q_j \Gamma(\tfrac{1}{2}(2j+m)+s)[\Gamma(\tfrac{1}{2}(2j+m))]^{-1}.$$

Since $\mathcal{L}(V) = \chi^2(r)$, the sth moment of V is given by

(3.11) $$E(V^s) = 2^r\Gamma(\tfrac{1}{2}(s+r))/\Gamma(\tfrac{1}{2}r).$$

Substitution of (3.10) and (3.11) into (3.9) gives for the sth moment of P,

(3.12) $$p_s = \lambda_1^s \sum_{j=0}^\infty q_j[\Gamma(\tfrac{1}{2}r)/\Gamma(\tfrac{1}{2}r+s)][\Gamma(\tfrac{1}{2}(2j+m)+s)/\Gamma(\tfrac{1}{2}(2j+m)],$$

for $s = 0, 1, 2 \cdots$.

Next recall that P always lies in the finite range $(0, \lambda_m)$, and it is sometimes restricted even further to (λ_1, λ_m). Hence, the moments, p_s, uniquely determine the distribution of P.

Let $\phi_P(t)$ denote the characteristic function of P. We know in general that if all the moments of P exist

(3.13) $$\phi_P(t) = 1 + \sum_{s=1}^\infty (it)^s(s!)^{-1}p_s .$$

Substitution of (3.12) gives the result

(3.14) $$\phi_P(t) = 1 + \sum_{s=1}^\infty \sum_{j=0}^\infty (it\lambda_1)^s q_j \Gamma(\tfrac{1}{2}r)\Gamma(\tfrac{1}{2}(2j+m)+s) \cdot[s!\Gamma(\tfrac{1}{2}r+s)\Gamma(\tfrac{1}{2}(2j+m))]^{-1}.$$

This series representation may be used numerically by precomputing the moments, p_s, and then inverting the resulting function, (3.13).

4. Beta variate representations.

4.1 *Exact distribution.* Bloch and Watson [4] studied the posterior distribution of the cell probabilities in a multinomial distribution under uniform prior probabilities. They noticed that an arbitrary linear combination of the cell probabilities has a posterior distribution of the same form as that of P, defined in (1.1); i.e., a linear combination of correlated beta variates.

In our notation, let $z_k = x_k^2/\sum_1^{m+1} x_i^2$, for $k = 1, \cdots, m$, and define $z' \equiv (z_1, \cdots, z_m)$, $\lambda' \equiv (\lambda_1, \cdots, \lambda_m)$, for the x_i's defined in (1.1). It is well known ([20], pp. 177–182) that the z_i's are identically distributed beta variates with $\frac{1}{2}$ and $\frac{1}{2}m$ d.f., and that z follows a Dirichlet distribution with intraclass covariance matrix

$$\Sigma = (\sigma_{ij}) = \operatorname{Var}(z) = \sigma^2 \begin{pmatrix} 1 & & \rho \\ & \ddots & \\ \rho & & 1 \end{pmatrix},$$

where $\sigma^2 = 2m(m+1)^{-2}(m+3)^{-1}$, $\rho = -m^{-1}$. The result is that

$$\mathcal{L}(P) = \mathcal{L}(\lambda' z). \tag{4.1}$$

4.2. *Approximate distribution.* Since P is distributed as a linear combination of beta variates it is natural to approximate its distribution by that of a single beta variate. Such approximations have been considered by Theil and Nagar [14], and Bloch and Watson [4]. Since the mean and variance of P are easily found from (4.1) as

$$E(P) = \sum_1^m (\lambda_i(m+1)^{-1}), \qquad \mathrm{Var}(P) = \sum_{i=1}^m \sum_{j=1}^m \lambda_i\lambda_j\sigma_{ij},$$

matching these moments to those of a beta variate on the same range as P gives the approximate result

$$\mathcal{L}(P) \cong \lambda_m\beta(a, b),$$

where

$$a = [E(P)/\lambda_m][(E(P)/\mathrm{Var}(P))(\lambda_m - E(P)) - 1],$$
$$b = [1 - E(P)/\lambda_m][(E(P)/\mathrm{Var}(P))(\lambda_m - E(P)) - 1)],$$

which is identical with the result given in [4].

Acknowledgment. I am grateful to Professor H. Theil for suggesting this problem.

REFERENCES

[1] ANDERSON, R. L. (1942). Distribution of the Serial Correlation Coefficient. *Ann. Math. Statist.* **13** 1–13.

[2] ANDERSON, R. L. and ANDERSON, T. W. (1950). Distribution of the Circular Serial Correlation Coefficient for Residuals from a Fitted Fourier Series. *Ann. Math. Statist.* **21** 59–81.

[3] ANDERSON, T. W. (1948). On the Theory of Testing Serial Correlation. *Skand. Aktuarietidskr.* **31** 88.

[4] BLOCH, DANIEL A. and WATSON, GEOFFREY S. (1967). A Bayesian Study of the Multinomial Distribution. *Ann. Math. Statist.* **38** 1423–1435.

[5] COCHRANE, D. and ORCUTT, G. H. (1949). Application of Least Squares Regressions to Relationships Containing Auto-Correlated Error Terms. *J. Amer. Statist. Assoc.* **44** 32–61.

[6] DIXON, WILFRID J. (1944). Further Contributions to the Problem of Serial Correlation. *Ann. Math. Statist.* **15** 119–144.

[7] DURBIN, J. and WATSON, G. S. (1950). Testing for Serial Correlation in Least Squares Regression I. *Biometrika* **37** 409–428.

[8] DURBIN, J. and WATSON, G. S. (1951). Testing for Serial Correlation in Least Squares Regression II. *Biometrika* **38** 159–178.

[9] KOOPMANS, T. C. (1942). Serial Correlation and Quadratic Forms in Normal Variables. *Ann. Math. Statist.* **13** 14–33.

[10] MAUCHLY, J. W. (1940). Significance Test for Sphericity of a Normal n-Variate Distribution. *Ann. Math. Statist.* **11** 204–209.

[11] MORAN, P. A. P. (1950). A test for the serial independence of residuals. *Biometrika* **37** 178.

[12] Pitman, E. J. G. (1937). The "closest" estimates of statistical parameters. *Proc. Cambridge Philos. Soc.* **33** 212.

[13] Robbins, H. and Pitman, E. J. G. (1949). Application of the Method of Mixtures to Quadratic Forms in Normal Variates. *Ann. Math. Statist.* **20** 552–560.

[14] Theil, H. and Nagar, A. L. (1961). Testing the Independence of Regression Disturbances. *J. Amer. Statist. Assoc.* **56** 793–806.

[15] Theil, H. (1965). The Analysis of Disturbances in Regression Analysis. *J. Amer. Statist. Assoc.* **60** 1067–1079.

[16] von Neumann, J., Kent, R. H., Bellinson, H. R. and Hart, B. I. (1941). The Mean Square Successive Difference. *Ann. Math. Statist.* **12** 153–162.

[17] von Neumann, J. (1941). Distribution of the Ratio of the Mean Square Successive Difference to the Variance. *Ann. Math. Statist.* **12** 367–395.

[18] von Neumann, J. (1942). A Further Remark Concerning the Distribution of the Ratio of the Mean Square Successive Difference to the Variance. *Ann. Math. Statist.* **13** 86–88.

[19] von Neumann, J. and Hart, B. I. (1942). Tabulation of the Probabilities for the Ratio of the Mean Square Successive Difference to the Variance. *Ann. Math. Statist.* **13** 207–214.

[20] Wilks, S. S. (1962). *Mathematical Statistics*. Wiley, New York.

[21] Yule, G. U. (1921). On the Time-Correlation Problem. *J. Roy. Statist. Soc.* **84** 496–537.