



OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

Biometrika Trust

A Test for the Serial Independence of Residuals

Author(s): P. A. P. Moran

Source: *Biometrika*, Vol. 37, No. 1/2 (Jun., 1950), pp. 178-181

Published by: [Oxford University Press](#) on behalf of [Biometrika Trust](#)

Stable URL: <http://www.jstor.org/stable/2332162>

Accessed: 16-12-2015 00:38 UTC

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Biometrika Trust and Oxford University Press are collaborating with JSTOR to digitize, preserve and extend access to *Biometrika*.

<http://www.jstor.org>

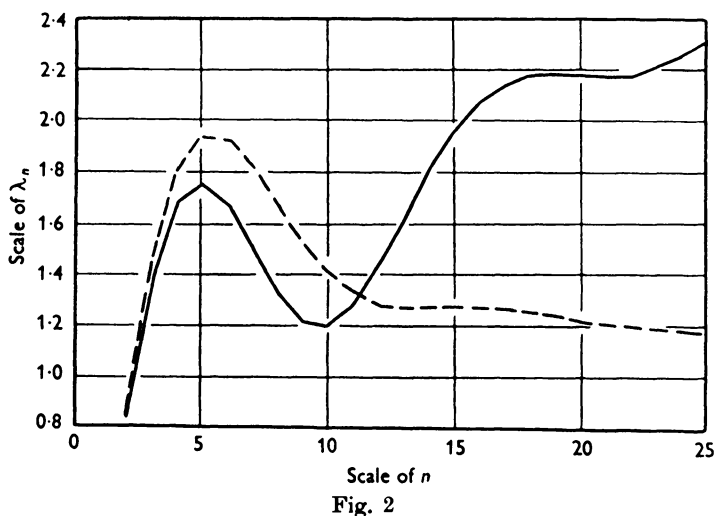


Fig. 2

SUMMARY

A method of estimating the parameters of an autoregressive time-series has been suggested and applied to three series. The method is based on the independence properties of the disturbance function and superposed variation which generate the series, and utilizes a number of serial correlations (or lagged product-sums) of the series for estimation.

The results for the first two series (which are of known construction) are found to be satisfactory. As regards the third (sunspot numbers), it is doubtful whether this series is properly represented by an autoregressive scheme of the type assumed.

REFERENCES

- BARTLETT, M. S. (1946). *J.R. Statist. Soc. Suppl.* **8**, 27.
 KENDALL, M. G. (1944). *Biometrika*, **33**, 105.
 KENDALL, M. G. (1945a). *J.R. Statist. Soc.* **108**, 93.
 KENDALL, M. G. (1945b). *J.R. Statist. Soc.* **108**, 226.
 KENDALL, M. G. (1946a). *Contributions to the Study of Oscillatory Time-Series*. Cambridge University Press.
 KENDALL, M. G. (1946b). *The Advanced Theory of Statistics*, **2**. London: Charles Griffin and Co.
 QUENOUILLE, M. H. (1947). *J.R. Statist. Soc.* **110**, 123.
 SPENCER SMITH, J. L. (1944). *J.R. Statist. Soc.* **107**, 231.
 YULE, G. UDNY (1927). *Philos. Trans. A*, **226**, 267.
 YULE, G. UDNY (1945). *J.R. Statist. Soc.* **108**, 208.

A test for the serial independence of residuals

By P. A. P. MORAN

Institute of Statistics, Oxford University

In economic problems we often wish to test the significance of a regression coefficient of one variate on another. If (x_i, y_i) ($i = 1, \dots, n$) is a sample of n pairs of values we may test the significance of the regression of y on x if we have reason to believe that the values of y are serially independent. But the y 's are often values of a time series which is serially correlated. To test whether this is so we may test the serial correlation coefficient (for references to the relevant theory see Moran (1948)) or, what is practically equivalent, use von Neumann's ratio. However, the serial dependence in the y 's may be due solely to their regression on the x 's if the latter are themselves serially correlated. The purpose of the present note is to set up a test for the serial independence of the residuals from the regression line. If the series are sufficiently long for the test to be reasonably sensitive, and if no serial dependence is indicated we may, with some assurance, use the residuals to provide an estimate of the standard error of the regression coefficient.

The criterion used in the test is the first-order cyclic serial correlation coefficient of the observed residuals, and we determine its first and second sampling moments under the assumption that the true residuals are independently distributed in the same normal distribution.

Suppose that the true regression of the variate y on x is given by

$$Y = \alpha + \beta X,$$

and suppose that in the sample $n\bar{x} = \sum_1^n x_i = 0$. By choosing the origin of measurements for y we can arrange that $\alpha = 0$. The true sample residuals, which we assume for simplicity to have unit variance, are then

$$\eta_i = y_i - \beta x_i,$$

and the observed residuals are

$$\epsilon_i = y_i - \hat{a} - \hat{b}x_i,$$

where

$$\hat{a} = \bar{y} = \frac{1}{n} \sum_1^n y_i,$$

$$\hat{b} = \frac{\sum_1^n x_i y_i}{\sum_1^n x_i^2}.$$

Also $\bar{y} = \beta\bar{x} + \bar{\eta} = \bar{\eta}$ because $\bar{x} = 0$. Then

$$\epsilon_i = \eta_i - \bar{\eta} - x_i \left(\frac{\sum_1^n x_i \eta_i}{\sum_1^n x_i^2} \right).$$

We now consider the moments of the cyclic serial correlation coefficient

$$R_1 = \frac{\sum_1^n \epsilon_i \epsilon_{i+1}}{\sum_1^n \epsilon_i^2}, \quad (1)$$

where ϵ_{n+1} is defined to be equal to ϵ_1 . Now it is known that $\sum_1^n \epsilon_i^2$ is distributed as χ^2 with $n-2$ degrees of freedom. R_1 is the ratio of a quadratic form in the ϵ 's to another quadratic form which is distributed as χ^2 . It follows that (compare Koopmans, 1942) R_1 and $\sum_1^n \epsilon_i^2$ are distributed independently of each other and so

$$E(R_1^p) = \frac{E\left\{\sum_1^n \epsilon_i \epsilon_{i+1}\right\}^p}{E\left\{\sum_1^n \epsilon_i^2\right\}^p} \quad (2)$$

for any integer p . From this we can obtain the moments of R_1 . For the denominator we have

$$E\left\{\sum_1^n \epsilon_i^2\right\} = n-2, \quad (3)$$

$$E\left\{\sum_1^n \epsilon_i^2\right\}^2 = n(n-2). \quad (4)$$

Write

$$B = \frac{\sum x_i \eta_i}{\sum x_i^2},$$

where (and subsequently) the symbol Σ indicates a sum taken from 1 to n unless otherwise indicated. We then have

$$\begin{aligned} \Sigma \epsilon_i \epsilon_{i+1} &= \Sigma \{\eta_i - \bar{\eta} - x_i B\} \{\eta_{i+1} - \bar{\eta} - x_{i+1} B\} \\ &= \Sigma (\eta_i - \bar{\eta}) (\eta_{i+1} - \bar{\eta}) - B \Sigma x_i (\eta_{i+1} - \bar{\eta}) \\ &\quad - B \Sigma x_{i+1} (\eta_i - \bar{\eta}) + B^2 \Sigma x_i x_{i+1}. \end{aligned}$$

Taking expectations we find that

$$E(\Sigma \epsilon_i \epsilon_{i+1}) = -1 - \frac{\Sigma x_i x_{i+1}}{\Sigma x_i^2},$$

since it may be easily verified that

$$E(\eta_i - \bar{\eta})(\eta_{i+1} - \bar{\eta}) = -\frac{1}{n}, \quad (5)$$

$$E(B^2) = (\Sigma x_i^2)^{-1}, \quad (6)$$

$$E\{B(\eta_i - \bar{\eta})\} = E(B\eta_i) = x_i(\Sigma x_i^2)^{-1}, \quad (7)$$

and therefore

$$E(\Sigma \epsilon_i \epsilon_{i+1}) = -1 - r_{1(e)},$$

where $r_{1(e)}$ is the first-order cyclic serial correlation coefficient of the observed x 's, namely,

$$r_{1(e)} = \frac{\Sigma x_i x_{i+1}}{\Sigma x_i^2}.$$

We thus have

$$E(R_1) = -\frac{1 + r_{1(e)}}{n - 2}. \quad (8)$$

We now find the variance of R_1 . The awkwardness of the following calculations arises from the fact that unlike the situation in the corresponding argument used in a previous paper (Moran, 1948) the ϵ_i are not all on the same footing but have variances and covariances which depend on their suffixes. It is convenient to write

$$z_i = \eta_i - \bar{\eta}.$$

Then the z 's and B are distributed normally and we have

$$\text{var}(z_i) = \frac{n-1}{n}, \quad (9)$$

$$\text{cov}(z_i, z_j) = -\frac{1}{n}, \quad \text{if } i \neq j, \quad (10)$$

$$\text{cov}(Bz_i) = x_i(\Sigma x_i^2)^{-1}. \quad (11)$$

Using (6), (9), (10) and (11), and remembering that B and the z_i are normally distributed about zero we have, from the consideration of the characteristic function of their joint distribution

$$\begin{aligned} E(B^4) &= \frac{3}{(\Sigma x_i^2)^2}, \\ E(B^2 z_i) &= \frac{3x_i}{(\Sigma x_i^2)^2}, \\ E(B^2 z_i^2) &= \frac{n-1}{n \Sigma x_i^2} + \frac{2x_i^2}{(\Sigma x_i^2)^2}, \\ E(B^2 z_i z_j) &= -\frac{1}{n \Sigma x_i^2} + \frac{2x_i x_j}{(\Sigma x_i^2)^2}, \quad \text{if } i \neq j, \\ E(Bz_i^2) &= \frac{3(n-1)x_i}{n \Sigma x_i^2}, \\ E(Bz_i^2 z_j) &= \frac{(n-1)x_j}{n \Sigma x_i^2} - \frac{2x_i x_j}{n \Sigma x_i^2}, \quad \text{if } i \neq j, \\ E(Bz_i z_j z_k) &= -\frac{x_i + x_j + x_k}{n \Sigma x_i^2} \quad \text{if } i, j, k \text{ are unequal.} \end{aligned}$$

We must now evaluate

$$\begin{aligned} E(\Sigma \epsilon_i \epsilon_{i+1})^2 &= E \sum_i \sum_j (z_i - Bx_i)(z_{i+1} - Bx_{i+1})(z_j - Bx_j)(z_{j+1} - Bx_{j+1}) \\ &= E \sum_i \sum_j \{z_i z_{i+1} z_j z_{j+1} + B^4 x_i x_{i+1} x_j x_{j+1} + B^2 z_i z_{i+1} x_j x_{j+1} \\ &\quad + B^2 z_{i+1} z_{j+1} x_i x_j + 2B^2 z_j z_{j+1} x_i x_{i+1} - 2Bz_i z_{i+1} z_j x_{j+1} \\ &\quad - 2Bz_i z_{i+1} z_{j+1} x_j - 2B^2 z_j x_i x_{i+1} x_{j+1} \\ &\quad - 2B^2 z_{j+1} x_i x_{i+1} x_j + 2B^2 z_i z_{j+1} x_{i+1} x_j\}. \\ &= A_1 + A_2 + A_3 + A_4 + A_5 + A_6 + A_7 + A_8 + A_9 + A_{10}, \quad \text{say.} \end{aligned}$$

We evaluate each of these terms separately.

$$(A_1) \quad A_1 = \sum_1^n z_i^2 z_{i+1}^2 + 2 \sum_1^n z_i z_{i+1}^2 z_{i+2} + \sum' z_i z_{i+1} z_j z_{j+1},$$

where the last sum is over all values of i and j such that the suffixes are unequal. Now

$$E(z_i^2 z_{i+1}^2) = \frac{n^2 + 1}{n^2},$$

$$E(z_i z_{i+1}^2 z_{i+2}) = -\frac{n-3}{n^2},$$

$$E(z_i z_{i+1} z_j z_{j+1}) = \frac{3}{n^2}.$$

There are $n(n-3)$ terms of the type $z_i z_{i+1} z_j z_{j+1}$, and adding the various terms together we get

$$E(A_1) = \frac{(n-1)(n+2)}{n}.$$

In a similar manner we can prove by straightforward algebra the following results:

$$E(A_2) = 3r_{1(x)}^2,$$

$$E(A_3) = 1 + 2r_{1(x)}^2,$$

$$E(A_4) = 1 + 2r_{1(x)}^2,$$

$$E(A_5) = -2r_{1(x)} + 4r_{1(x)}^2,$$

$$E(A_6) = -2 + 2r_{1(x)} - 2r_{2(x)},$$

where $r_{2(x)}$ is the second-order serial correlation coefficient of the x 's, namely,

$$\frac{\sum_1^n x_i x_{i+2}}{\sum_1^n x_i^2},$$

$$E(A_7) = -2 + 2r_{1(x)} - 2r_{2(x)},$$

$$E(A_8) = -6r_{1(x)}^2,$$

$$E(A_9) = -6r_{1(x)}^2,$$

$$E(A_{10}) = 2r_{2(x)} + 4r_{1(x)}^2.$$

Adding the above we obtain finally

$$E(\text{numerator})^2 = \frac{(n-2)(n+1)}{n} + 2r_{1(x)} + 3r_{1(x)}^2 - 2r_{2(x)}.$$

$$\begin{aligned} \text{We therefore have} \quad E(R_1^2) &= \frac{1}{n(n-2)} \left\{ \frac{(n-2)(n+1)}{n} + 2r_{1(x)} + 3r_{1(x)}^2 - 2r_{2(x)} \right\} \\ &= \frac{n+1}{n^2} + \frac{1}{n(n-2)} \{ 2r_{1(x)} + 3r_{1(x)}^2 - 2r_{2(x)} \}, \end{aligned} \quad (12)$$

$$\text{and} \quad \text{var}(R_1) = E(R_1^2) - \{E(R_1)\}^2, \quad (13)$$

and is more easily calculated using this formula since $E(R_1)$ will have to be found in any case. By lengthy but elementary methods, the distribution of R_1 can be shown to tend to normality.

The same method could be applied to the non-circular serial correlation coefficient, but it appears that this would result in rather more complicated formulae.

REFERENCES

- KOOPMANS, T. (1942). *Ann. Math. Statist.* **13**, 14.
MORAN, P. A. P. (1948). *Biometrika*, **35**, 255.