

Cowles Foundation Paper 21c

Reprinted from

ECONOMETRICA, Journal of the Econometric Society, Vol. 15, No. 3, July, 1947
The University of Chicago, Chicago 37, Illinois, U.S.A.

A NOTE ON A MAXIMUM-LIKELIHOOD ESTIMATE¹

By T. W. ANDERSON²

IN a recent booklet,³ M. G. Kendall has studied the estimation of the "period" of the stochastic model generated by the difference equation,

$$(1) \quad x_t + \alpha x_{t-1} + \beta x_{t-2} = u_t,$$

where u_t is a random term independently distributed with mean zero. Kendall defines the period as the period of the solution of the pure difference equation. This is

$$(2) \quad \gamma = \frac{2\pi}{\arccos \left(\frac{-\alpha}{2\sqrt{\beta}} \right)}.$$

An estimate of γ obtained by applying the method of maximum likelihood under the assumption that u_t is normally distributed is consistent and asymptotically normally distributed. The asymptotic standard deviation is given in this note. Although Kendall considers many estimates of the period in his publication, he does not use the maximum-likelihood estimate although it has desirable properties in large samples that several of the other estimates do not have.⁴ It is interesting to compare the numerical results of using this estimate with those Kendall applies to four artificial series generated by (1), each series with a different pair of coefficients α and β .⁵

If the u_t ($t = 1, 2, \dots, T$) are assumed to be normally distributed and if x_{-1} and x_0 are assumed to be fixed, the estimate defined by the method of maximum likelihood is obtained by substituting in (2) the estimates of α and β found by the method of maximum likelihood under these assumptions [see equations (8)]. H. B. Mann and A. Wald⁶ have

¹ This note will be included in Cowles Commission Papers, New Series, No. 21.

² Research Consultant of the Cowles Commission for Research in Economics.

³ M. G. Kendall, *Contributions to the Study of Oscillatory Times Series*, Cambridge University Press, 1946.

⁴ Kendall considers this estimate applied to a series with superimposed errors in M. G. Kendall, "On Autoregressive Time Series," *Biometrika*, Vol. 33, August, 1944, pp. 105-122, esp. p. 113. His objection to bias due to superimposed errors can be met by modifying the estimate slightly and using higher-order lag moments.

⁵ The values were chosen so the solutions were stable and oscillatory; that is, $\alpha < 4\beta$, $0 < \beta < 1$.

⁶ H. B. Mann and A. Wald, "On the Statistical Treatment of Linear Stochastic Difference Equations," *ECONOMETRICA*, Vol. 11, July-October, 1943, pp. 173-220.

shown that even if the above assumptions about the distribution of u_t are not fulfilled, existence of all moments [proof uses only six] is sufficient for these estimates, a and b , to be consistent and asymptotically normally distributed with covariance matrix

$$(3) \quad \frac{\sigma^2}{T} \begin{pmatrix} m_0 & m_1 \\ m_1 & m_0 \end{pmatrix}^{-1},$$

where

$$(4) \quad \begin{aligned} \sigma^2 &= \mathcal{E}u_t^2, \\ m_0 &= \lim_{t \rightarrow \infty} \mathcal{E}x_t^2, \\ m_1 &= \lim_{t \rightarrow \infty} \mathcal{E}x_t x_{t-1}. \end{aligned}$$

It follows that g , the maximum-likelihood estimate of γ , is consistent. Furthermore, it is asymptotically normally distributed because it is a function (with continuous first derivatives at α, β) of asymptotically normally distributed quantities.⁷ The asymptotic variance of g is given by the formula

$$(5) \quad \mathcal{E}(g-\gamma)^2 = \begin{pmatrix} \frac{\partial \gamma}{\partial \alpha} & \frac{\partial \gamma}{\partial \beta} \end{pmatrix} \begin{pmatrix} \mathcal{E}(a-\alpha)^2 & \mathcal{E}(a-\alpha)(b-\beta) \\ \mathcal{E}(b-\beta)(a-\alpha) & \mathcal{E}(b-\beta)^2 \end{pmatrix} \begin{pmatrix} \frac{\partial \gamma}{\partial \alpha} \\ \frac{\partial \gamma}{\partial \beta} \end{pmatrix}.$$

This variance is

$$(6) \quad \frac{\sigma^2}{T(m_0^2 - m_1^2)} \begin{pmatrix} -\gamma^2 & \alpha\gamma^2 \\ 2\pi\sqrt{4\beta - \alpha^2} & 4\pi\beta\sqrt{4\beta - \alpha^2} \end{pmatrix} \begin{pmatrix} m_0 & -m_1 \\ -m_1 & m_0 \end{pmatrix} \begin{pmatrix} -\gamma^2 \\ \alpha\gamma^2 \\ 2\pi\sqrt{4\beta - \alpha^2} \\ 4\pi\beta\sqrt{4\beta - \alpha^2} \end{pmatrix} \\ = \frac{\gamma^4(1-\beta)(4\beta^3 + 4\beta^2 - 3\beta\alpha^2 + \alpha^2)}{T(2\pi)^2 4\beta^2(4\beta - \alpha^2)},$$

since

$$(7) \quad \begin{aligned} m_0 &= \frac{\sigma^2(1+\beta)}{(1-\beta)[(1+\beta)^2 - \alpha^2]}, \\ m_1 &= \frac{-\sigma^2\alpha}{(1-\beta)[(1+\beta)^2 - \alpha^2]}. \end{aligned}$$

⁷ See, for example, A. Wald, "Setting of Tolerance Limits when the Sample Is Large," *Annals of Mathematical Statistics*, Vol. 13, December, 1942, pp. 389-399, esp. p. 390.

The asymptotic theory applies to Kendall's examples since the moments of u_t exist and the properties of x_{-1} and x_0 do not affect the asymptotic properties of g in this case. The random disturbances u_t were obtained from tables of random numbers⁸ in such a way that the probability of u_t taking on any integral value between -49 and $+49$ was $1/99$. The initial values x_{-1} and x_0 were obtained similarly.

In Table 1 are given the characteristics of the four equations [of the form of (1)] from which Kendall's four artificial series were derived. The standard deviations of g have been computed⁹ from the population values of α and β . Then the maximum-likelihood estimates of α , β , and γ are given.

In order to use the formulas of Mann and Wald we need the values of x_{-1} and x_0 . Since Kendall does not give these, we have used the formulas with x_1 and x_2 considered as fixed (i.e., we consider the series as starting with these given). Then Series 1 has 478 terms and each of Series 2, 3, and 4 has 238 terms. Since the expected values are zero, the estimation formulas are

$$(8) \quad a = - \frac{\begin{vmatrix} \sum_{t=3}^T x_t x_{t-1} & \sum_{t=3}^T x_{t-2} x_{t-1} \\ \sum_{t=3}^T x_t x_{t-2} & \sum_{t=3}^T x_{t-2}^2 \end{vmatrix}}{\begin{vmatrix} \sum_{t=3}^T x_{t-1}^2 & \sum_{t=3}^T x_{t-2} x_{t-1} \\ \sum_{t=3}^T x_{t-2} x_{t-1} & \sum_{t=3}^T x_{t-2}^2 \end{vmatrix}},$$

$$b = - \frac{\begin{vmatrix} \sum_{t=3}^T x_{t-1}^2 & \sum_{t=3}^T x_t x_{t-1} \\ \sum_{t=3}^T x_{t-2} x_{t-1} & \sum_{t=3}^T x_t x_{t-2} \end{vmatrix}}{\begin{vmatrix} \sum_{t=3}^T x_{t-1}^2 & \sum_{t=3}^T x_{t-2} x_{t-1} \\ \sum_{t=3}^T x_{t-2} x_{t-1} & \sum_{t=3}^T x_{t-2}^2 \end{vmatrix}}.$$

⁸ M. G. Kendall and B. Babington Smith, *Tables of Random Sampling Numbers*, Tracts for Computers No. 4, Cambridge University Press, 1939.

⁹ The author is indebted to Herman Rubin for checking the computations in Table 1.

TABLE 1

Characteristic \ Series	1	2	3	4
T	478	238	238	238
α	-1.1	-1.2	-1.1	1.0
β	0.5	0.4	0.8	0.5
γ	9.25	19.5	6.92	3.00
Asymptotic Standard Deviation of g	0.47	7.4	0.17	0.066
a	-1.133	-1.210	-1.113	-1.059
b	0.486	0.422	0.795	0.549
g	10.09	16.8	7.01	2.66
Mean distance between peaks	5.05	5.57	5.52	2.69
Mean distance between upcrosses	8.30	12.39	6.38	2.76
Mean distance between correlogram troughs	8.33	12.0	7.67	2.80

Since the series are fairly long, the treatment of the first two and last two observations has little effect on the calculation of a and b .

For Series 1, 2, and 3 the estimate g differs from γ by less than two standard deviations, but for Series 4 the estimate is considerably different from γ . A test of significance (or a confidence-interval statement) based on the asymptotic normal distribution with standard deviation (6) can be applied only if T is large enough so that the asymptotic theory holds well for these values of α and β . At present, however, the size of the error of the normal approximation when $T=238$ is unknown. Kendall suggests that the approach to normality is related to T/γ , the number of periods. However, for the three series with $T=238$ the discrepancy is greatest for the smallest γ . Of course, a study of only four series is nothing more than suggestive.

The statistics Kendall prefers for estimating the period is the mean distance between "upcrosses." By an "upcross" he means a pair of values x_t, x_{t+1} such that $x_t < 0, x_{t+1} > 0$. If $x_t = 0$, then the pair x_{t-1}, x_{t+1} must be such that $x_{t-1} < 0, x_{t+1} > 0$, etc. Kendall also considers the mean distance between peaks of the series and the mean distance between troughs of the correlogram. For Series 1, 2, and 3 the maximum-likelihood estimate seems somewhat better than the other estimates. For Series 4 the other estimates are closer, but all are at least as far from γ as three standard deviations (of g).

Columbia University