Chapter 1

Serial Correlation and Durbin–Watson Bounds.

T. W. Anderson, *Department of Economics and Department of Statistics, Stanford University*

The model is $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{u}$, where $\boldsymbol{y}$ is an n-vector of dependent variables, $\boldsymbol{X}$ is a matrix of $n \times k$ independent variables, and $\boldsymbol{u}$ is a n-vector of unobserved disturbance. Let $\boldsymbol{z} = \boldsymbol{y} - \boldsymbol{Xb}$, where $\boldsymbol{b}$ is the least squares estimate of $\boldsymbol{\beta}$. The d-statistic tests the hypothesis that the components of $\boldsymbol{u}$ are independent versus the alternative that the components follow a Markov process. The Durbin–Watson bounds pertain to the distribution of the d-statistics.

1.1 Introduction

A *time series* is composed of a sequence of observations $y_1, \ldots, y_n$, where the index i of the observation y_i represents time. An important feature of a time series is the order of observations: y_i is observed after $y_1, \ldots, y_{i-1}$ are observed. The correlation of successive observations is called a *serial correlation*. Related to each y_i may be a vector of *independent* variables $(x_{1i}, \ldots, x_{ki})$. Many questions of time series analysis relate to the possible dependence of y_i on $x_{1i}, \ldots, x_{ki}$. See Anderson (1971), for example.

A serial correlation (first-order) of a sequence $y_1, \ldots, y_n$ is

$$\frac{\sum_{i=2}^{n} y_i y_{i-1}}{\sum_{i=1}^{n} y_i^2}.$$

This coefficient measures the correlation between $y_1, \ldots, y_{n-1}$ and $y_2, \ldots, y_n$. There are various modifications of this correlation coefficient such as replacing y_i by $y_i - \bar{y}$. See below for the *circular serial coefficient*. The term "auto-correlation" is also used for serial correlation.

I shall discuss two papers coauthored by James Durbin and Geoffrey Watson entitled "Testing for serial correlation in least squares regression I and II," published in 1950 and 1951 respectively. The statistical analysis developed in these papers has proved very useful in econometric research.

The Durbin-Watson papers are based on a model in which there is a set of "independent" variables $x_{1i}, x_{2i}, \ldots, x_{ni}$ associated with each "dependent" variable y_i for $i = 1, \ldots, n$. The dependent variable of y_i is considered as the linear combination

$$y_i = \beta_1 x_{1i} + \ldots + \beta_R x_{Ri} + w_i, \quad i = 1, \ldots, n,$$

where w_i is an unobservable random disturbance. The questions that Durbin and Watson address have to do with the possible dependence in a set of observations $y_1, \ldots, y_n$ beyond what is explained by the independent variables.

1.2 Circular Serial Correlation

R. L. Anderson (1942),who was Watson's thesis advisor, studied the statistic

$$\frac{\sum_{i=1}^{n} (y_i - y_{i-1})^2}{\sum_{i=1}^{n} y_i^2} = 2 - 2\frac{\sum_{i=1}^{n} y_i y_{i-1}}{\sum_{i=1}^{n} y_i^2}$$

where $y_0 = y_n$. The statistic

$$\frac{\sum_{i=1}^{n} y_i y_{i-1}}{\sum_{i=1}^{n} y_i^2}$$

is known as the "circular serial correlation coefficient." Defining $y_0 = y_n$ is a device to make the mathematics simpler. The serial correlation coefficient measures the relationship between the sequence $y_1, \ldots, y_n$ and $y_0, \ldots, y_{n-1}$.

In our exposition we make repeated use of the fact that the distribution of $\mathbf{x}'\mathbf{A}\mathbf{x}$ is the distribution of $\sum_{i=1}^{n} \lambda_i z_i^2$, where $\lambda_1, \ldots, \lambda_n$ are the characteristic roots (latent roots) of $\mathbf{A}$, that is, the roots of

$$|\mathbf{A} - \lambda \mathbf{I}_n| = 0 \qquad (\mathbf{A} = \mathbf{A}') ,$$

and $\boldsymbol{x}$ and $\boldsymbol{z}$ have the density $N(\mathbf{0}, \sigma^2 \boldsymbol{I})$. The numerator of the circular serial correlation is $\mathbf{x}'\mathbf{A}\mathbf{x}$ where

$$\mathbf{A} = \frac{1}{2} \begin{bmatrix} 0 & 1 & 0 & \ldots & 1 \\ 1 & 0 & 1 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 0 & 0 & \ldots & 0 \end{bmatrix} .$$

The characteristic roots are $\lambda_j = \cos 2\pi j/n$ and $\sin 2\pi j/n, j = 1, \ldots, n$. If n is even the roots occur in pairs. The distribution of the circular serial correlation is the distribution of

$$\frac{\sum_{j=1}^{n} \lambda_j z_j^2}{\sum_{j=1}^{n} z_j^2} ;$$

$z_1, \ldots, z_n$ are independent standard normal variables. Anderson studied the distribution of the circular serial correlation, its moments, and other properties.

1.3 Periodic Trends

During World War II R. L. Anderson and I were members of the Princeton Statistical Research Group. We noticed that the jth characteristic vector of $\mathbf{A}$ had the form $\cos 2\pi jh/n$ and/or $\sin 2\pi jh/n$, $h = 1, \ldots, n$. These functions are periodic and hence are suitable to represent seasonal variation. We considered the model

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + u_i$$

where $x_{hi} = \cos 2\pi hi/n$ and/or $\sin 2\pi hi/n$. Then the distribution of

$$r = \frac{\sum_i \left(y_i - \sum \beta_h x_{hi}\right)\left(y_{i-1} - \sum \beta_h x_{h,i-1}\right)}{\sum \left(y_i - \sum \beta_h x_{hi}\right)^2}$$

is the distribution of $\sum_j \lambda_j z_j^2 \Big/ \sum_j z_j^2$, where the sums are over the z's corresponding to the cos and sin terms that did not occur in the trends. The distributions of the serial correlations have the same form as before.

Anderson and Anderson found distributions of r for several cyclical trends as well as moments and approximate distributions.

1.4 Uniformly Most Powerful Tests

Many problems of serial correlation are included in the general model (T. W. Anderson, 1948)

$$K \exp\left\{-\frac{\alpha}{2}\left[(\boldsymbol{y}-\boldsymbol{\mu})'\,\boldsymbol{\Psi}\,(\boldsymbol{y}-\boldsymbol{\mu})+\lambda\,(\boldsymbol{y}-\boldsymbol{\mu})'\,\boldsymbol{\Theta}\,(\boldsymbol{y}-\boldsymbol{\mu})\right]\right\}$$

where K is a constant, $\alpha > 0$, $\boldsymbol{\Psi}$ a given positive definite matrix, $\boldsymbol{\Theta}$ a given symmetric matrix, λ a parameter such that $\boldsymbol{\Psi} - \lambda\boldsymbol{\Theta}$ is positive definite, and $\boldsymbol{\mu}$ is the expectation of $\boldsymbol{y}$,

$$\mathcal{E}\boldsymbol{y} = \boldsymbol{\mu} = \sum \beta_j \boldsymbol{\phi}_j.$$

We shall consider testing the hypothesis

$$H : \lambda = 0.$$

The first theorem characterizes tests such that the probability of the acceptance region when $\lambda = 0$ does not depend on the values of $\beta_1, \ldots, \beta_k$. The second theorem gives conditions for a test being uniformly most powerful when $\lambda > 0$ is the alternative.

These theorems are applicable to the circular serial correlation when $\boldsymbol{\Psi} = \sigma^2\boldsymbol{I}$ and $\boldsymbol{\Theta} = \sigma^2\boldsymbol{A}$ defined above.

The equation

$$\sum (y_i - y_{i-1})^2 = \sum \left(y_i^2 + y_{i-1}^2\right) - 2\sum y_t y_{t-1}$$

suggests that a serial correlation can be studied in terms of $\sum (y_t - y_{t-1})^2$ which may be suitable to test that $y_t, \ldots, y_n$ are independent against the alternative that $y_1, \ldots, y_n$ satisfy an autoregressive process. Durbin and Watson prefer to study

$$d = \frac{\sum (z_i - z_{i-1})^2}{\sum z_i^2},$$

where $\boldsymbol{z}$ is defined below.

1.5 Durbin–Watson

The model is

$$\underset{n\times 1}{\boldsymbol{y}} = \underset{n\times k}{\boldsymbol{X}}\,\underset{k\times 1}{\boldsymbol{\beta}} + \underset{n\times 1}{\boldsymbol{u}}.$$

We consider testing the null hypothesis that $\boldsymbol{u}$ has a normal distribution with mean $\mathbf{0}$ and covariance $\sigma^2\boldsymbol{I}_n$ against the alternative that $\boldsymbol{u}$ has a normal distribution with mean $\mathbf{0}$ and covariance $\sigma^2\boldsymbol{A}$, a positive definite matrix. The sample regression of $\boldsymbol{y}$ is $\boldsymbol{b} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$ and the vector of residuals is

$$\begin{aligned}\boldsymbol{z} &= \boldsymbol{y} - \boldsymbol{X}\boldsymbol{b} = \left[\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\right]\boldsymbol{y} \\ &= \left[\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\right](\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}) \\ &= \boldsymbol{M}\boldsymbol{u},\end{aligned}$$

where

$$\boldsymbol{M} = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'.$$

Consider the serial correlation of the residuals

$$r = \frac{\boldsymbol{z}'\boldsymbol{A}\boldsymbol{z}}{\boldsymbol{z}'\boldsymbol{z}} = \frac{\boldsymbol{u}'\boldsymbol{M}'\boldsymbol{A}\boldsymbol{M}\boldsymbol{u}}{\boldsymbol{u}'\boldsymbol{M}'\boldsymbol{M}\boldsymbol{u}}.$$

The matrix $\boldsymbol{M}$ is idempotent, that is, $\boldsymbol{M}^m = \boldsymbol{M}$, and symmetric. Its latent roots are 0 and 1 and it has rank $n-k$. Let the possibly nonzero roots of $\boldsymbol{M}'\boldsymbol{A}\boldsymbol{M}$ be $\boldsymbol{\nu}_1, \ldots, \nu_{n-k}$. There is an $n \times (n-k)$ matrix $\boldsymbol{H}$ such that $\boldsymbol{H}'\boldsymbol{H} = \boldsymbol{I}_{n-k}$ and

$$\boldsymbol{H}'\boldsymbol{M}'\boldsymbol{A}\boldsymbol{M}\boldsymbol{H} = \begin{bmatrix} \nu_1 & 0 & \ldots & 0 \\ 0 & \nu_2 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & \nu_{n-k} \end{bmatrix}.$$

Let $\boldsymbol{w} = \boldsymbol{H}'\boldsymbol{v}$. Then

$$r = \frac{\sum_{j=1}^{n-k} \nu_j w_j^2}{\sum_{j=1}^{n-k} w_j^2}.$$

Durbin and Watson prove that

$$\lambda_j \le \nu_j \le \lambda_{j+k}, \quad j = 1, \ldots, n-k.$$

Define

$$r_L = \frac{\sum_{j=1}^{n-k} \lambda_j w_j^2}{\sum_{j=1}^{n-k} w_j^2}, \qquad r_U = \frac{\sum_{j=k+1}^{n} \lambda_j w_j^2}{\sum_{j=1}^{n-k} w_j^2}.$$

Then $r_L \le r \le r_U$.

The "bounds procedure" is the following. If the observed serial correlation is greater than $r_U^\star$ conclude that the hypothesis of no serial correlation of the disturbances is rejected. If the observed correlation is less than $r_L^\star$, conclude that the hypothesis of no serial correlation of the disturbance is accepted. The interval $(r_L^\star, r_U^\star)$ is called "the zone of indeterminacy." If the observed correlation falls in the interval $(r_L^\star, r_U^\star)$, the data is considered as not leading to a conclusion.

Bibliography

[1] R. L. Anderson. Distribution of the serial correlation coefficient. *The Annals of Mathematical Statistics*, 13(1):1–13, 1942.

[2] R. L. Anderson and T. W. Anderson. Distribution of the circular serial correlation coefficient for residuals from a fitted fourier series. *The Annals of Mathematical Statistics*, 21(1):59–81, 1950.

[3] T. W. Anderson. On the theory of testing serial correlation. *Skandinavisk Aktuarietidskrift*, 31(3-4):88–116, 1948.

[4] T. W. Anderson. *The Statistical Analysis of Time Series*. Wiley, 1971.

[5] T. W. Anderson. *An Introduction to Multivariate Statistical Analysis,* Third Edition. Wiley, 2003.

[6] J. S. Chipman. *Advanced Econometric Theory*. Routledge, 2011.

[7] J. Durbin and G. S. Watson. Testing for serial correlation in least squares regression. I. *Biometrika*, 37:409–428, 1950.

[8] J. Durbin and G. S. Watson. Testing for serial correlation in least squares regression. II. *Biometrika*, 38:159–178, 1951.