A test for long-range dependence in a time series

Christopher F. Baum, Boston College, baum@bc.edu
Taüri Room, Boston College, room@bc.edu

Abstract: This insert implements the Hurst–Mandelbrot rescaled range statistic and the Lo (1991) modified rescaled range statistic to test for long-range dependence in a time series.

Keywords: fractional integration, long memory, rescaled range, time series.

Syntax

lomodrs varname [if exp] [in range] [ , maxlag(#) ]

This test is for use with time-series data; you must tsset your data before using lomodrs; see [R] tsset. varname or varlist may contain time-series operators; see [U] Time-series varlists.

Options

maxlag(#) specifies the maximum lag order for the test. By default, maxlag is calculated from the sample size and the first-order autocorrelation coefficient of the varname using the data-dependent rule of Andrews (1991), assuming that the data-generating process is AR(1). If maxlag is set to zero, the test performed is the classical Hurst–Mandelbrot rescaled-range statistic.

Description

The model of an autoregressive fractionally integrated moving average process of a time series of order \( p, d, q \), denoted by ARFIMA \((p, d, q)\), with mean \( \mu \), may be written using operator notation in terms of a white noise series \( \epsilon \) having variance \( \sigma^2_\epsilon \) as

\[
\Phi(L)(1 - L)^d(y_t - \mu) = \Theta(L)\epsilon_t
\]

where \( L \) is the backward-shift operator, \( \Phi(L) = 1 - \phi_1L - \cdots - \phi_pL^p \), \( \Theta(L) = 1 + \theta_1L + \cdots + \theta_qL^q \), and \((1 - L)^d\) is the fractional differencing operator defined by

\[
(1 - L)^d = \sum_{k=0}^{\infty} \frac{\Gamma(k - d)L^k}{\Gamma(-d)\Gamma(k + 1)}
\]

with \( \Gamma(\cdot) \) denoting the gamma (generalized factorial) function. The parameter \( d \) is allowed to assume any real value. The arbitrary restriction of \( d \) to integer values gives rise to the standard autoregressive integrated moving average (ARIMA) model. The stochastic process \( y_t \) is both stationary and invertible if all zeros of \( \Phi(L) \) and \( \Theta(L) \) lie outside the unit circle and \(|d| < 0.5\). The process is nonstationary for \( d \geq 0.5 \), as it possesses infinite variance, for example, see Granger and Joyeux (1980).

Assuming that \( d \in [0, 0.5) \), Hosking (1981) showed that the autocorrelation function, \( \rho(\cdot) \), of an ARFIMA process is proportional to \( k^{2d-1} \) as \( k \to \infty \). Consequently, the autocorrelations of the ARFIMA process decay hyperbolically to zero as \( k \to \infty \) in contrast to the faster, geometric decay of a stationary ARMA process. For \( d \in (0, 0.5) \), \( \sum_{j=-\infty}^{\infty} |\rho(j)| \) diverges as \( n \to \infty \), and the ARFIMA process is said to exhibit long memory, or long-range positive dependence. The process is said to exhibit intermediate memory (anti-persistence), or long-range negative dependence, for \( d \in (-0.5, 0) \).

The importance of long-range dependence in economic and financial time series was first studied by Mandelbrot (1972), who proposed the \( R/S \) (range over standard deviation) statistic, also known as the rescaled-range statistic, originally developed by Hurst (1951) in the context of hydrological studies. The \( R/S \) statistic is the range of the partial sums of deviations of a time series from its mean, rescaled by its standard deviation. For a sample \( x_1, \ldots, x_n \),

\[
Q_n = \frac{1}{s_n} \left[ \max_{1 \leq k \leq n} \sum_{j=1}^{k} (x_j - \bar{x}_n) - \min_{1 \leq k \leq n} \sum_{j=1}^{k} (x_j - \bar{x}_n) \right]
\]

where \( s_n \) is the maximum likelihood estimator of the standard deviation of \( x \). The first bracketed term is the maximum of the partial sums of the first \( k \) deviations of \( x_j \) from the full-sample mean, which is nonnegative. The second bracketed term is the corresponding minimum, which is nonpositive. The difference of these two quantities is thus nonnegative, so that \( Q_n > 0 \). Empirical studies have demonstrated that the \( R/S \) statistic has the ability to detect long-range dependence in the data. Like many other estimators of long-range dependence, though, the \( R/S \) statistic has been shown to be excessively sensitive to “short-range dependence,” or short memory, features of the data. Lo (1991) shows that a sizable AR(1) component in the data generating process will seriously bias the \( R/S \) statistic. He modifies the \( R/S \) statistic to account for the effect of short-range dependence by applying a “Newey–West” correction (using a Bartlett window) to derive a consistent estimate of the long-range variance.
of the time series. For maxlag > 0, the denominator of the statistic is computed as the Newey–West estimate of the long run variance of the series; see [R] newey.

Critical values for the test are taken from Table II of Lo (1991).

**Saved results**

lomodrs saves the following in r():

<table>
<thead>
<tr>
<th>Scalars</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>r(lomodrs)</td>
<td>test statistic</td>
</tr>
<tr>
<td>r(N)</td>
<td>degrees of freedom</td>
</tr>
</tbody>
</table>

**Remarks**

The description of the Hurst–Mandelbrot and Lo statistics draws heavily from Chapter 2 of Campbell et al. (1997).

**Examples**

Data from Terence Mills’ *Econometric Analysis of Financial Time Series* on U.S. S&P 500 stock returns are analyzed.

```
. use http://fmwww.bc.edu/ec-p/data/Mills2d/sp500a.dta
. tsset time variable:  year, 1871 to 1997
. lomodrs sp500ar
Lo Modified R/S test for sp500ar
Critical values for H0: sp500ar is not long-range dependent
90%: [ 0.861, 1.747 ]
95%: [ 0.809, 1.862 ]
99%: [ 0.721, 2.098 ]
Test statistic:  .780838 (1 lags via Andrews criterion) N = 124
. lomodrs sp500ar, max(0)
Hurst-Mandelbrot Classical R/S test for sp500ar
Critical values for H0: sp500ar is not long-range dependent
90%: [ 0.861, 1.747 ]
95%: [ 0.809, 1.862 ]
99%: [ 0.721, 2.098 ]
Test statistic:  .799079 N = 124
. lomodrs sp500ar if tin(1946,)
Lo Modified R/S test for sp500ar
Critical values for H0: sp500ar is not long-range dependent
90%: [ 0.861, 1.747 ]
95%: [ 0.809, 1.862 ]
99%: [ 0.721, 2.098 ]
Test statistic:  1.08705 (0 lags via Andrews criterion) N = 50
```

Applied to the full sample, the Lo modified $R/S$ test rejects the null hypothesis of no long-range dependence at the 95% level. The Hurst–Mandelbrot test yields a similar inference. When the sample is restricted to the postwar era, the Lo test no longer can reject the null hypothesis at any level of significance.

**references**


