Title

corrgram — Tabulate and graph autocorrelations

Syntax

Autocorrelations, partial autocorrelations, and Portmanteau (Q) statistics

corrgram varname [if] [in] [ , corrgram_options ]

Graph autocorrelations with confidence intervals

ac varname [if] [in] [ , ac_options ]

Graph partial autocorrelations with confidence intervals

pac varname [if] [in] [ , pac_options ]

corrgram_options description

Main
lags(#) calculate # autocorrelations
noplot suppress character-based plots
yw calculate partial autocorrelations by using Yule–Walker equations

ac_options description

Main
lags(#) calculate # autocorrelations
generate(newvar) generate a variable to hold the autocorrelations
level(#) set confidence level; default is level(95)
fft calculate autocorrelation by using Fourier transforms

Plot
line_options change look of dropped lines
marker_options change look of markers (color, size, etc.)
marker_label_options add marker labels; change look or position

CI plot
ciopts(area_options) affect rendition of the confidence bands

Add plots
addplot(plot) add other plots to the generated graph

Y axis, X axis, Titles, Legend, Overall
twoway_options any options other than by() documented in [G] twoway_options

77
**corrgram — Tabulate and graph autocorrelations**

**pac_options**

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main</strong></td>
</tr>
<tr>
<td><code>lags(#)</code></td>
</tr>
<tr>
<td><code>generate(newvar)</code></td>
</tr>
<tr>
<td><code>yw</code></td>
</tr>
<tr>
<td><code>level(#)</code></td>
</tr>
<tr>
<td><strong>Plot</strong></td>
</tr>
<tr>
<td><code>line_options</code></td>
</tr>
<tr>
<td><code>marker_options</code></td>
</tr>
<tr>
<td><code>marker_label_options</code></td>
</tr>
<tr>
<td><strong>CI plot</strong></td>
</tr>
<tr>
<td><code>ciopts(area_options)</code></td>
</tr>
<tr>
<td><strong>SRV plot</strong></td>
</tr>
<tr>
<td><code>srv</code></td>
</tr>
<tr>
<td><code>srvopts(marker_options)</code></td>
</tr>
<tr>
<td><strong>Add plots</strong></td>
</tr>
<tr>
<td><code>addplot(plot)</code></td>
</tr>
<tr>
<td><strong>Y axis, X axis, Titles, Legend, Overall</strong></td>
</tr>
<tr>
<td><code>twoway_options</code></td>
</tr>
</tbody>
</table>

You must `tsset` your data before using `corrgram`, `ac`, or `pac`; see [TS] `tsset`. Also, the time series must be dense (nonmissing and no gaps in the time variable) in the sample if you specify the `fft` option. `varname` may contain time-series operators; see [U] 11.4.4 Time-series `varlists`.

**Menu**

`corrgram`

Statistics > Time series > Graphs > Autocorrelations & partial autocorrelations

`ac`

Statistics > Time series > Graphs > Correlogram (ac)

`pac`

Statistics > Time series > Graphs > Partial correlogram (pac)

**Description**

`corrgram` produces a table of the autocorrelations, partial autocorrelations, and Portmanteau (Q) statistics. It also displays a character-based plot of the autocorrelations and partial autocorrelations. See [TS] `wntestq` for more information on the Q statistic.

`ac` produces a correlogram (a graph of autocorrelations) with pointwise confidence intervals that is based on Bartlett’s formula for MA(q) processes.
pac produces a partial correlogram (a graph of partial autocorrelations) with confidence intervals calculated using a standard error of $1/\sqrt{n}$. The residual variances for each lag may optionally be included on the graph.

**Options for corrgram**

- `lags(#)` specifies the number of autocorrelations to calculate. The default is to use \( \min(\lfloor n/2 \rfloor - 2, 40) \), where \( \lfloor n/2 \rfloor \) is the greatest integer less than or equal to \( n/2 \).
- `noplot` prevents the character-based plots from being in the listed table of autocorrelations and partial autocorrelations.
- `yw` specifies that the partial autocorrelations be calculated using the Yule–Walker equations instead of using the default regression-based technique. `yw` cannot be used if `srv` is used.

**Options for ac and pac**

- `lags(#)` specifies the number of autocorrelations to calculate. The default is to use \( \min(\lfloor n/2 \rfloor - 2, 40) \), where \( \lfloor n/2 \rfloor \) is the greatest integer less than or equal to \( n/2 \).
- `generate(newvar)` specifies a new variable to contain the autocorrelation (ac command) or partial autocorrelation (pac command) values. This option is required if the `nograph` option is used.
- `nograph` (implied when using `generate()` in the dialog box) prevents ac and pac from constructing a graph. This option requires the `generate()` option.
- `yw` (pac only) specifies that the partial autocorrelations be calculated using the Yule–Walker equations instead of using the default regression-based technique. `yw` cannot be used if `srv` is used.
- `level(#)` specifies the confidence level, as a percentage, for the confidence bands in the ac or pac graph. The default is `level(95)` or as set by `set level`; see [R] `level`.
- `fft` (ac only) specifies that the autocorrelations be calculated using two Fourier transforms. This technique can be faster than simply iterating over the requested number of lags.

`line_options`, `marker_options`, and `marker_label_options` affect the rendition of the plotted autocorrelations (with ac) or partial autocorrelations (with pac).

`line_options` specify the look of the dropped lines, including pattern, width, and color; see [G] `line_options`.

`marker_options` specify the look of markers. This look includes the marker symbol, the marker size, and its color and outline; see [G] `marker_options`.

`marker_label_options` specify if and how the markers are to be labeled; see [G] `marker_label_options`.

`ciopts(area_options)` affects the rendition of the confidence bands; see [G] `area_options`. 
SRV plot

srv (pac only) specifies that the standardized residual variances be plotted with the partial autocorrelations. srv cannot be used if yw is used.

srvopts(marker_options) (pac only) affects the rendition of the plotted standardized residual variances; see [G] marker_options. This option implies the srv option.

Add plots

addplot(plot) adds specified plots to the generated graph; see [G] addplot_option.

Y axis, X axis, Titles, Legend, Overall

twoway_options are any of the options documented in [G] twoway_options, excluding by(). These include options for titling the graph (see [G] title_options) and for saving the graph to disk (see [G] saving_option).

Remarks

corrgram tabulates autocorrelations, partial autocorrelations, and Portmanteau (Q) statistics and plots the autocorrelations and partial autocorrelations. The Q statistics are the same as those produced by [TS] wntestq. ac produces graphs of the autocorrelations, and pac produces graphs of the partial autocorrelations.

Example 1

Here we use the international airline passengers dataset (Box, Jenkins, and Reinsel 1994, Series G). This dataset has 144 observations on the monthly number of international airline passengers from 1949 through 1960. We can list the autocorrelations and partial autocorrelations by using corrgram.

```
use http://www.stata-press.com/data/r11/air2
(TIMESLAB: Airline passengers)
corrgram air, lags(20)
```

<table>
<thead>
<tr>
<th>LAG</th>
<th>AC</th>
<th>PAC</th>
<th>Q</th>
<th>Prob&gt;Q [Autocorrelation]</th>
<th>Prob&gt;Q [Partial Autocor]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9480</td>
<td>0.9589</td>
<td>132.14</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.8756</td>
<td>-0.3298</td>
<td>245.65</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.8067</td>
<td>0.2018</td>
<td>342.67</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.7526</td>
<td>0.1450</td>
<td>427.74</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.7138</td>
<td>0.2585</td>
<td>504.8</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.6817</td>
<td>-0.0269</td>
<td>575.6</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.6629</td>
<td>0.2043</td>
<td>643.04</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.6556</td>
<td>0.1561</td>
<td>709.48</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.6709</td>
<td>0.5686</td>
<td>779.59</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.7027</td>
<td>0.2926</td>
<td>857.07</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.7432</td>
<td>0.8402</td>
<td>944.39</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.7604</td>
<td>0.6127</td>
<td>1036.5</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.7127</td>
<td>-0.6660</td>
<td>1118.0</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.6663</td>
<td>-0.3846</td>
<td>1185.6</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.5859</td>
<td>0.0787</td>
<td>1241.5</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.5380</td>
<td>-0.2066</td>
<td>1289.0</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.4997</td>
<td>-0.0581</td>
<td>1330.4</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.4687</td>
<td>-0.0435</td>
<td>1367.0</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>0.4499</td>
<td>0.2773</td>
<td>1401.1</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.4416</td>
<td>-0.0405</td>
<td>1434.1</td>
<td>0.0000</td>
<td></td>
</tr>
</tbody>
</table>
We can use `ac` to produce a graph of the autocorrelations.

```
    . ac air, lags(20)
```

The data probably have a trend component as well as a seasonal component. First-differencing will mitigate the effects of the trend, and seasonal differencing will help control for seasonality. To accomplish this goal, we can use Stata’s time-series operators. Here we graph the partial autocorrelations after controlling for trends and seasonality. We also use `srv` to include the standardized residual variances.

```
    . pac DS12.air, lags(20) srv
```

See [U] 11.4.4 Time-series varlists for more information about time-series operators.
Saved results

corrgram saves the following in r():

Scalars

\( r(\text{lags}) \) number of lags

Matrices

\( r(AC) \) vector of autocorrelations
\( r(PAC) \) vector of partial autocorrelations
\( r(Q) \) vector of Q statistics
\( r(ac#) \) AC for lag #
\( r(pac#) \) PAC for lag #
\( r(q#) \) Q for lag #

Methods and formulas

corrgram, ac, and pac are implemented as ado-files.

Box, Jenkins, and Reinsel (1994); Newton (1988); Chatfield (2004); and Hamilton (1994) provide excellent descriptions of correlograms. Newton (1988) also discusses the calculation of the various quantities.

The autocovariance function for a time series \( x_1, x_2, \ldots, x_n \) is defined for \( |v| < n \) as

\[
\hat{R}(v) = \frac{1}{n} \sum_{i=1}^{n-|v|} (x_i - \bar{x})(x_{i+v} - \bar{x})
\]

where \( \bar{x} \) is the sample mean, and the autocorrelation function is then defined as

\[
\hat{\rho}_v = \frac{\hat{R}(v)}{\hat{R}(0)}
\]

The variance of \( \hat{\rho}_v \) is given by Bartlett’s formula for MA(q) processes. From Brockwell and Davis (2002, 94), we have

\[
\text{Var}(\hat{\rho}_v) = \begin{cases} 
\frac{1}{n} & v = 1 \\
\frac{1}{n} \left( 1 + 2 \sum_{i=1}^{v-1} \hat{\rho}^2(i) \right) & v > 1 
\end{cases}
\]

The partial autocorrelation at lag \( v \) measures the correlation between \( x_t \) and \( x_{t+v} \) after the effects of \( x_{t+1}, \ldots, x_{t+v-1} \) have been removed. By default, corrgram and pac use a regression-based method to estimate it. We run an OLS regression of \( x_t \) on \( x_{t-1}, \ldots, x_{t-v} \) and a constant term. The estimated coefficient on \( x_{t-v} \) is our estimate of the \( v \)th partial autocorrelation. The residual variance is the estimated variance of that regression, which we then standardize by dividing by \( \hat{R}(0) \).

If the \( yw \) option is specified, corrgram and pac use the Yule–Walker equations to estimate the partial autocorrelations. Per Enders (2004, 65), let \( \phi_{vv} \) denote the \( v \)th partial autocorrelation coefficient. We then have

\[
\hat{\phi}_{11} = \hat{\rho}_1
\]
and for \( v > 1 \)

\[
\hat{\phi}_{vv} = \frac{\hat{\rho}_v - \sum_{j=1}^{v-1} \hat{\phi}_{v-1,j} \hat{\rho}_{v-j}}{1 - \sum_{j=1}^{v-1} \hat{\phi}_{v-1,j} \hat{\rho}_j}
\]

and

\[
\hat{\phi}_{vj} = \hat{\phi}_{v-1,j} - \hat{\phi}_{vv} \hat{\phi}_{v-1,v-j}, \quad j = 1, 2, \ldots, v - 1
\]

Unlike the regression-based method, the Yule–Walker equations-based method ensures that the first-sample partial autocorrelation equal the first-sample autocorrelation coefficient, as must be true in the population; see Greene (2008, 725).

McCullough (1998) discusses other methods of estimating \( \phi_{vv} \); he finds that relative to other methods, such as linear regression, the Yule–Walker equations-based method performs poorly, in part because it is susceptible to numerical error. Box, Jenkins, and Reinsel (1994, 68) also caution against using the Yule–Walker equations-based method, especially with data that are nearly nonstationary.

**Acknowledgment**

The `ac` and `pac` commands are based on the `ac` and `pac` commands written by Sean Becketti (1992), a past editor of the *Stata Technical Bulletin*.

**References**


**Also see**

[Ts] `tsset` — Declare data to be time-series data

[Ts] `wntestq` — Portmanteau (Q) test for white noise

[Ts] `pergram` — Periodogram