

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]
```

| <i>corrgram_options</i> | description |
|-------------------------|-------------|
|-------------------------|-------------|

Main

| | |
|-----------------|---|
| <u>lags</u> (#) | calculate # autocorrelations |
| <u>noplot</u> | suppress character-based plots |
| <u>yw</u> | calculate partial autocorrelations by using Yule–Walker equations |

| <i>ac_options</i> | description |
|-------------------|-------------|
|-------------------|-------------|

Main

| | |
|-----------------------------------|--|
| <u>lags</u> (#) | calculate # autocorrelations |
| <u>generate</u> (<i>newvar</i>) | generate a variable to hold the autocorrelations; implies <code>nograph</code> |
| <u>level</u> (#) | set confidence level; default is <code>level(95)</code> |
| <u>fft</u> | calculate autocorrelation by using Fourier transforms |

Plot

| | |
|-----------------------------|--|
| <i>line_options</i> | change look of dropped lines |
| <i>marker_options</i> | change look of markers (color, size, etc.) |
| <i>marker_label_options</i> | add marker labels; change look or position |

CI plot

| | |
|---------------------------------------|--|
| <u>ciopts</u> (<i>area_options</i>) | affect rendition of the confidence bands |
|---------------------------------------|--|

Add plots

| | |
|--------------------------------|--|
| <u>addplot</u> (<i>plot</i>) | add other plots to the generated graph |
|--------------------------------|--|

Y axis, X axis, Titles, Legend, Overall

| | |
|-----------------------|---|
| <i>twoway_options</i> | any options other than by() documented in [G] <i>twoway_options</i> |
|-----------------------|---|

| <i>pac_options</i> | description |
|--|--|
| Main | |
| <u>lags</u> (#) | calculate # partial autocorrelations |
| <u>generate</u> (<i>newvar</i>) | generate a variable to hold the partial autocorrelations; implies nograph |
| <u>yw</u> | calculate partial autocorrelations by using Yule–Walker equations |
| <u>level</u> (#) | set confidence level; default is level(95) |
| Plot | |
| <i>line_options</i> | change look of dropped lines |
| <i>marker_options</i> | change look of markers (color, size, etc.) |
| <i>marker_label_options</i> | add marker labels; change look or position |
| CI plot | |
| <u>ciopts</u> (<i>area_options</i>) | affect rendition of the confidence bands |
| SRV plot | |
| <u>srv</u> | include standardized residual variances in graph |
| <u>srvopts</u> (<i>marker_options</i>) | affect rendition of the plotted standardized residual variances (SRVs) |
| Add plots | |
| <u>addplot</u> (<i>plot</i>) | add other plots to the generated graph |
| Y axis, X axis, Titles, Legend, Overall | |
| <i>twoway_options</i> | any options other than <code>by()</code> documented in [G] <i>twoway_options</i> |

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] 11.4.3 **Time-series varlists**.

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

Main

`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 $n/2$.

`noplots` 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`

Main

`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.

Plot

`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**.

CI plot

`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`.

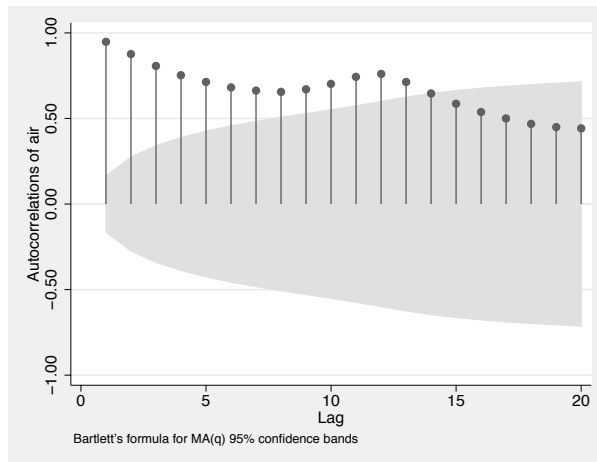
```
. use http://www.stata-press.com/data/r10/air2
(TIMESLAB: Airline passengers)
. corrgram air, lags(20)
```

| LAG | AC | PAC | Q | Prob>Q | -1 | 0 | 1 | -1 | 0 | 1 |
|-----|--------|---------|--------|--------|-------------------|---|---|-------------------|---|---|
| | | | | | [Autocorrelation] | | | [Partial Autocor] | | |
| 1 | 0.9480 | 0.9589 | 132.14 | 0.0000 | | | | | | |
| 2 | 0.8756 | -0.3298 | 245.65 | 0.0000 | | | | | | |
| 3 | 0.8067 | 0.2018 | 342.67 | 0.0000 | | | | | | |
| 4 | 0.7526 | 0.1450 | 427.74 | 0.0000 | | | | | | |
| 5 | 0.7138 | 0.2585 | 504.8 | 0.0000 | | | | | | |
| 6 | 0.6817 | -0.0269 | 575.6 | 0.0000 | | | | | | |
| 7 | 0.6629 | 0.2043 | 643.04 | 0.0000 | | | | | | |
| 8 | 0.6556 | 0.1561 | 709.48 | 0.0000 | | | | | | |
| 9 | 0.6709 | 0.5686 | 779.59 | 0.0000 | | | | | | |
| 10 | 0.7027 | 0.2926 | 857.07 | 0.0000 | | | | | | |
| 11 | 0.7432 | 0.8402 | 944.39 | 0.0000 | | | | | | |
| 12 | 0.7604 | 0.6127 | 1036.5 | 0.0000 | | | | | | |
| 13 | 0.7127 | -0.6660 | 1118 | 0.0000 | | | | | | |
| 14 | 0.6463 | -0.3846 | 1185.6 | 0.0000 | | | | | | |
| 15 | 0.5859 | 0.0787 | 1241.5 | 0.0000 | | | | | | |
| 16 | 0.5380 | -0.0266 | 1289 | 0.0000 | | | | | | |
| 17 | 0.4997 | -0.0581 | 1330.4 | 0.0000 | | | | | | |
| 18 | 0.4687 | -0.0435 | 1367 | 0.0000 | | | | | | |
| 19 | 0.4499 | 0.2773 | 1401.1 | 0.0000 | | | | | | |
| 20 | 0.4416 | -0.0405 | 1434.1 | 0.0000 | | | | | | |

(Continued on next page)

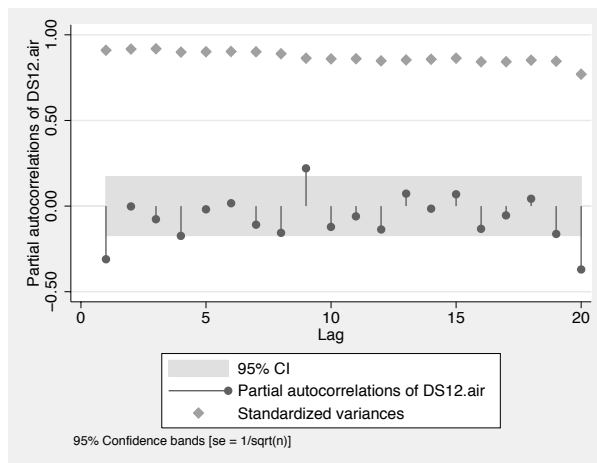
We can use `ac` to produce a graph of the autocorrelations.

```
. 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] **11.4.3 Time-series varlists** for more information about time-series operators.

Saved Results

`corrgram` saves the following in `r()`:

| | |
|----------------------|------------------------------------|
| Scalars | |
| <code>r(lags)</code> | number of lags |
| Matrices | |
| <code>r(AC)</code> | vector of autocorrelations |
| <code>r(PAC)</code> | vector of partial autocorrelations |
| <code>r(Q)</code> | vector of Q statistics |
| <code>r(ac#)</code> | AC for lag # |
| <code>r(pac#)</code> | PAC for lag # |
| <code>r(q#)</code> | 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, \dots, x_n is defined for $|v| < n$ as

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

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

$$\widehat{\rho}_v = \frac{\widehat{R}(v)}{\widehat{R}(0)}$$

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

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

The partial autocorrelation at lag v measures the correlation between x_t and x_{t+v} after the effects of $x_{t+1}, \dots, x_{t+v-1}$ 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}, \dots, 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 $\widehat{R}(0)$.

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

$$\widehat{\phi}_{11} = \widehat{\rho}_1$$

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

$$\hat{\phi}_{vj} = \hat{\phi}_{v-1,j} - \hat{\phi}_{vv} \hat{\phi}_{v-1,v-j} \quad j = 1, 2, \dots, v-1$$

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 (2003, 618).

McCullough (1998) discusses other methods of estimating ϕ_{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 Beckett (1992), a past editor of the *Stata Technical Bulletin*.

References

- Beckett, S. 1992. `stsl`: Autocorrelation and partial autocorrelation graphs. *Stata Technical Bulletin* 5: 27–28. Reprinted in *Stata Technical Bulletin Reprints*, vol. 1, pp. 221–223.
- Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. 1994. *Time Series Analysis: Forecasting and Control*. 3rd ed. Englewood Cliffs, NJ: Prentice Hall.
- Brockwell, P. J., and R. A. Davis. 2002. *Introduction to Time Series and Forecasting*. 2nd ed. New York: Springer.
- Chatfield, C. 2004. *The Analysis of Time Series: An Introduction*. 6th ed. London: Chapman & Hall/CRC.
- Enders, W. 2004. *Applied Econometric Time Series*. 2nd ed. New York: Wiley.
- Greene, W. H. 2003. *Econometric Analysis*. 5th ed. Upper Saddle River, NJ: Prentice Hall.
- Hamilton, J. D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.
- McCullough, B. D. 1998. Algorithm choice for (partial) autocorrelation functions. *Journal of Economic and Social Measurement* 24: 265–278.
- Newton, H. J. 1988. *TIMESLAB: A Time Series Laboratory*. Pacific Grove, CA: Wadsworth & Brooks/Cole.

Also See

- [TS] **tsset** — Declare data to be time-series data
- [TS] **wntestq** — Portmanteau (Q) test for white noise
- [TS] **pergram** — Periodogram