cumsp — Graph cumulative spectral distribution

Description

cumsp plots the cumulative sample spectral-distribution function evaluated at the natural frequencies for a (dense) time series.

Quick start

Plot cumulative sample spectral-distribution function for y using tsset data

cumsp y

As above, and create newv containing the cumulative distribution estimates

cumsp y, generate(newv)

Menu

Statistics > Time series > Graphs > Cumulative spectral distribution
Syntax

cumsp  varname  [if]  [in]  [ , options ]

options  Description

Main

generate(newvar)  create newvar holding distribution values

Plot

cline_options  affect rendition of the plotted points connected by lines
marker_options  change look of markers (color, size, etc.)
marker_label_options  add marker labels; change look or position

Add plots

ddplot(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-3] twoway_options

You must tsset your data before using cumsp; see [TS] tsset. Also, the time series must be dense (nonmissing with no gaps in the time variable) in the sample specified.

varname may contain time-series operators; see [U] 11.4.4 Time-series varlists.

Options

--- Main ---
generate(newvar)  specifies a new variable to contain the estimated cumulative spectral-distribution values.

--- Plot ---
cline_options  affect the rendition of the plotted points connected by lines; see [G-3] cline_options.
marker_options  specify the look of markers. This look includes the marker symbol, the marker size, and its color and outline; see [G-3] marker_options.
marker_label_options  specify if and how the markers are to be labeled; see [G-3] marker_label_options.

--- Add plots ---
ddplot(plot)  provides a way to add other plots to the generated graph; see [G-3] addplot_option.

--- Y axis, X axis, Titles, Legend, Overall ---
twoway_options  are any of the options documented in [G-3] twoway_options, excluding by(). These include options for titling the graph (see [G-3] title_options) and for saving the graph to disk (see [G-3] saving_option).
Remarks and examples

Example 1

Here we use the international airline passengers dataset (Box, Jenkins, and Reinsel 2008, Series G). This dataset has 144 observations on the monthly number of international airline passengers from 1949 through 1960. In the cumulative sample spectral distribution function for these data, we also request a vertical line at frequency $1/12$. Because the data are monthly, there will be a pronounced jump in the cumulative sample spectral-distribution plot at the $1/12$ value if there is an annual cycle in the data.

```stata
use http://www.stata-press.com/data/r15/air2
(TIMESLAB: Airline passengers)
cumsp air, xline(.083333333)
```

![Sample spectral distribution function](image)

The cumulative sample spectral-distribution function clearly illustrates the annual cycle.

Methods and formulas

A time series of interest is decomposed into a unique set of sinusoids of various frequencies and amplitudes.

A plot of the sinusoidal amplitudes versus the frequencies for the sinusoidal decomposition of a time series gives us the spectral density of the time series. If we calculate the sinusoidal amplitudes for a discrete set of “natural” frequencies $(1/n, 2/n, \ldots, q/n)$, we obtain the periodogram.

Let $x(1), \ldots, x(n)$ be a time series, and let $\omega_k = (k - 1)/n$ denote the natural frequencies for $k = 1, \ldots, \lfloor n/2 \rfloor + 1$ where $\lfloor \rfloor$ indicates the greatest integer function. Define

$$C_k^2 = \frac{1}{n^2} \left| \sum_{t=1}^{n} x(t)e^{2\pi i(t-1)\omega_k} \right|^2$$

A plot of $nC_k^2$ versus $\omega_k$ is then called the periodogram.
The sample spectral density may then be defined as $\hat{f}(\omega_k) = nC_k^2$.

If we let $\hat{f}(\omega_1), \ldots, \hat{f}(\omega_Q)$ be the sample spectral density function of the time series evaluated at the frequencies $\omega_j = (j - 1)/Q$ for $j = 1, \ldots, Q$ and we let $q = \lfloor Q/2 \rfloor + 1$, then

$$\hat{F}(\omega_k) = \frac{\sum_{i=1}^{k} \hat{f}(\omega_j)}{q \sum_{i=1}^{q} \hat{f}(\omega_j)}$$

is the sample spectral-distribution function of the time series.

References


Also see

[TS] tssset — Declare data to be time-series data
[TS] corrgram — Tabulate and graph autocorrelations
[TS] pergram — Periodogram