

psdensity — Parametric spectral density estimation after arima, arfima, and ucm

[Description](#)

[Options](#)

[Also see](#)

[Quick start](#)

[Remarks and examples](#)

[Menu](#)

[Methods and formulas](#)

[Syntax](#)

[References](#)

Description

`psdensity` estimates the spectral density of a stationary process using the parameters of a previously estimated parametric model.

`psdensity` works after `arfima`, `arima`, and `ucm`; see [\[TS\] arfima](#), [\[TS\] arima](#), and [\[TS\] ucm](#).

Quick start

Obtain spectral density values `spden` and corresponding frequencies `sfreq`

```
psdensity spden sfreq
```

Obtain power spectrum values `pspec` and corresponding frequencies `pfreq`

```
psdensity pspec pfreq, pspectrum
```

As above, but limit the frequency range to between 0 and 1

```
psdensity pspec pfreq, pspectrum range(0 1)
```

After `ucm`, obtain spectral density values for the second stochastic cycle `cpden` and corresponding frequencies `cfreq`

```
psdensity cpden cfreq, cycle(2)
```

Menu

Statistics > Time series > Postestimation > Parametric spectral density

Syntax

```
psdensity [type] newvarsd newvarf [if] [in] [, options]
```

where *newvar_{sd}* is the name of the new variable that will contain the estimated spectral density and *newvar_f* is the name of the new variable that will contain the frequencies at which the spectral density estimate is computed.

<i>options</i>	Description
<u>pspectrum</u>	estimate the power spectrum rather than the spectral density
<u>range(a b)</u>	limit the frequency range to $[a, b]$
<u>cycle(#)</u>	estimate the spectral density from the specified stochastic cycle; only allowed after <code>ucm</code>
<u>smemory</u>	estimate the spectral density of the short-memory component of the ARFIMA process; only allowed after <code>arfima</code>

Options

`pspectrum` causes `psdensity` to estimate the power spectrum rather than the spectral density. The power spectrum is equal to the spectral density times the variance of the process.

`range(a b)` limits the frequency range. By default, the spectral density is computed over $[0, \pi]$. Specifying `range(a b)` causes the spectral density to be computed over $[a, b]$. We require that $0 \leq a < b < \pi$.

`cycle(#)` causes `psdensity` to estimate the spectral density from the specified stochastic cycle after `ucm`. By default, the spectral density from the first stochastic cycle is estimated. `cycle(#)` must specify an integer that corresponds to a cycle in the model fit by `ucm`.

`smemory` causes `psdensity` to ignore the ARFIMA fractional integration parameter. The spectral density computed is for the short-memory ARMA component of the model.

Remarks and examples

[stata.com](http://www.stata.com)

Remarks are presented under the following headings:

The frequency-domain approach to time series
Some ARMA examples

The frequency-domain approach to time series

A stationary process can be decomposed into random components that occur at the frequencies $\omega \in [0, \pi]$. The spectral density of a stationary process describes the relative importance of these random components. `psdensity` uses the estimated parameters of a parametric model to estimate the spectral density of a stationary process.

We need some concepts from the frequency-domain approach to time-series analysis to interpret estimated spectral densities. Here we provide a simple, intuitive explanation. More technical presentations can be found in Priestley (1981), Harvey (1989, 1993), Hamilton (1994), Fuller (1996), and Wei (2006).

In the time domain, the dependent variable evolves over time because of random shocks. The autocovariances $\gamma_j, j \in \{0, 1, \dots, \infty\}$, of a covariance-stationary process y_t specify its variance and dependence structure, and the autocorrelations $\rho_j, j \in \{1, 2, \dots, \infty\}$, provide a scale-free measure of its dependence structure. The autocorrelation at lag j specifies whether realizations at time t and realizations at time $t - j$ are positively related, unrelated, or negatively related.

In the frequency domain, the dependent variable is generated by an infinite number of random components that occur at the frequencies $\omega \in [0, \pi]$. The spectral density specifies the relative importance of these random components. The area under the spectral density in the interval $(\omega, \omega + d\omega)$ is the fraction of the variance of the process that can be attributed to the random components that occur at the frequencies in the interval $(\omega, \omega + d\omega)$.

The spectral density and the autocorrelations provide the same information about the dependence structure, albeit in different domains. The spectral density can be written as a weighted average of the autocorrelations of y_t , and it can be inverted to retrieve the autocorrelations as a function of the spectral density.

Like autocorrelations, the spectral density is normalized by γ_0 , the variance of y_t . Multiplying the spectral density by γ_0 yields the power spectrum of y_t , which changes with the units of y_t .

A peak in the spectral density around frequency ω implies that the random components around ω make an important contribution to the variance of y_t .

A random variable primarily generated by low-frequency components will tend to have more runs above or below its mean than an independent and identically distributed (i.i.d.) random variable, and its plot may look smoother than the plot of the i.i.d. variable. A random variable primarily generated by high-frequency components will tend to have fewer runs above or below its mean than an i.i.d. random variable, and its plot may look more jagged than the plot of the i.i.d. variable.

□ Technical note

A more formal specification of the spectral density allows us to be more specific about how the spectral density specifies the relative importance of the random components.

If y_t is a covariance-stationary process with absolutely summable autocovariances, its spectrum is given by

$$g_y(\omega) = \frac{1}{2\pi} \gamma_0 + \frac{1}{\pi} \sum_{k=1}^{\infty} \gamma_k \cos(\omega k) \tag{1}$$

where $g_y(\omega)$ is the spectrum of y_t at frequency ω and γ_k is the k th autocovariance of y_t . Taking the inverse Fourier transform of each side of (1) yields

$$\gamma_k = \int_{-\pi}^{\pi} g_y(\omega) e^{i\omega k} d\omega \tag{2}$$

where i is the imaginary number $i = \sqrt{-1}$.

Evaluating (2) at $k = 0$ yields

$$\gamma_0 = \int_{-\pi}^{\pi} g_y(\omega) d\omega$$

which means that the variance of y_t can be decomposed in terms of the spectrum $g_y(\omega)$. In particular, $g_y(\omega)d\omega$ is the contribution to the variance of y_t attributable to the random components in the interval $(\omega, \omega + d\omega)$.

The spectrum depends on the units in which y_t is measured, because it depends on the γ_0 . Dividing both sides of (1) by γ_0 gives us the scale-free spectral density of y_t :

$$f_y(\omega) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{k=1}^{\infty} \rho_k \cos(\omega k)$$

By construction,

$$\int_{-\pi}^{\pi} f_y(\omega) d\omega = 1$$

so $f_y(\omega)d\omega$ is the fraction of the variance of y_t attributable to the random components in the interval $(\omega, \omega + d\omega)$. □

Some ARMA examples

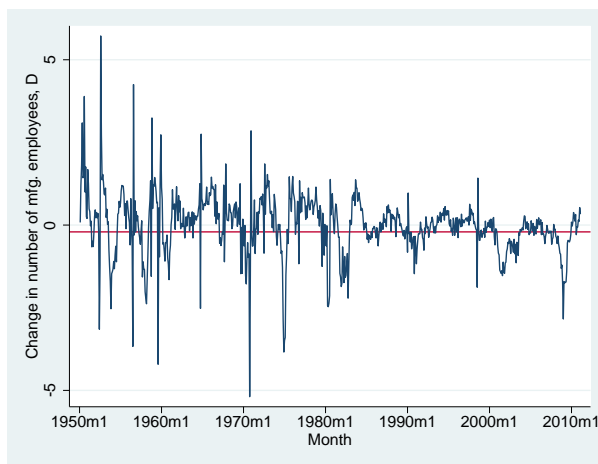
In this section, we estimate and interpret the spectral densities implied by the estimated ARMA parameters. The examples illustrate some of the essential relationships between covariance-stationary processes, the parameters of ARMA models, and the spectral densities implied by the ARMA-model parameters.

See [TS] **ucm** for a discussion of unobserved-components models and the stochastic-cycle model derived by Harvey (1989) for stationary processes. The stochastic-cycle model has a different parameterization of the spectral density, and it tends to produce spectral densities that look more like probability densities than ARMA models. See *Remarks and examples* in [TS] **ucm** for an introduction to these models, some examples, and some comparisons between the stochastic-cycle model and ARMA models.

▷ Example 1

Let's consider the changes in the number of manufacturing employees in the United States, which we plot below.

```
. use https://www.stata-press.com/data/r17/manemp2
(FRED data: Number of manufacturing employees in U.S.)
. tsline D.manemp, yline(-0.206)
```



We added a horizontal line at the sample mean of -0.0206 to highlight that there appear to be more runs above or below the mean than we would expect in data generated by an i.i.d. process.

As a first pass at modeling this dependence, we use `arima` to estimate the parameters of a first-order autoregressive (AR(1)) model. Formally, the AR(1) model is given by

$$y_t = \alpha y_{t-1} + \epsilon_t$$

where y_t is the dependent variable, α is the autoregressive coefficient, and ϵ_t is an i.i.d. error term. See [TS] [arima](#) for an introduction to ARMA modeling and the `arima` command.

```
. arima D.manemp, ar(1) noconstant
(setting optimization to BHHH)
Iteration 0: log likelihood = -870.64844
Iteration 1: log likelihood = -870.64794
Iteration 2: log likelihood = -870.64789
Iteration 3: log likelihood = -870.64787
Iteration 4: log likelihood = -870.64786
(switching optimization to BFGS)
Iteration 5: log likelihood = -870.64786
Iteration 6: log likelihood = -870.64786
ARIMA regression
Sample: 1950m2 thru 2011m2          Number of obs   =       733
Wald chi2(1)                        =       730.51
Log likelihood = -870.6479          Prob > chi2     =       0.0000
```

D.manemp		OPG		z	P> z	[95% conf. interval]	
	Coefficient	std. err.					
ARMA							
ar							
L1.	.5179561	.0191638	27.03	0.000	.4803959	.5555164	
/sigma	.7934554	.0080636	98.40	0.000	.777651	.8092598	

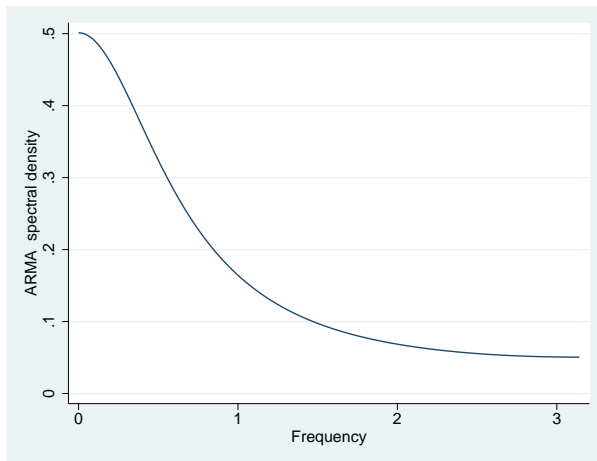
Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at zero.

The statistically significant estimate of 0.518 for the autoregressive coefficient indicates that there is an important amount of positive autocorrelation in this series.

The spectral density of a covariance-stationary process is symmetric around 0. Following convention, `psdensity` estimates the spectral density over the interval $[0, \pi)$ at the points given in [Methods and formulas](#).

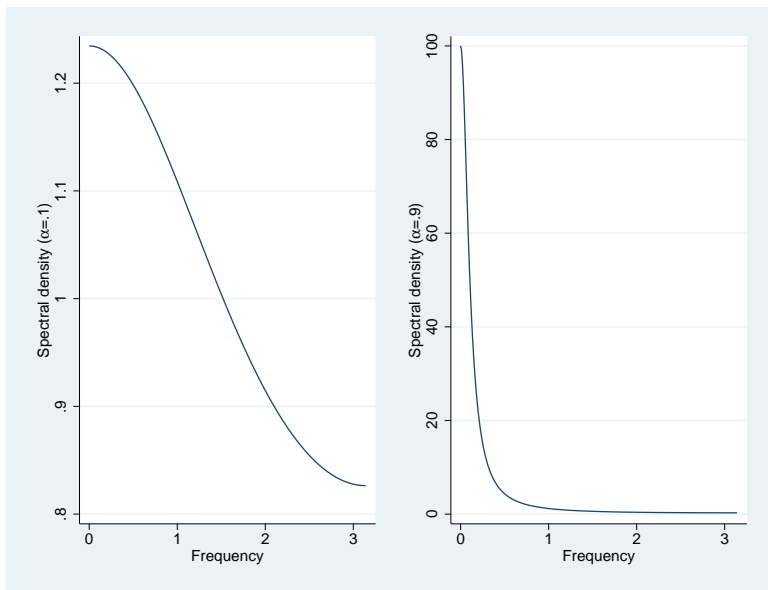
Now we use `psdensity` to estimate the spectral density of the process implied by the estimated ARMA parameters. We specify the names of two new variables in the call to `psdensity`. The first new variable will contain the estimated spectral density. The second new variable will contain the frequencies at which the spectral density is estimated.

```
. psdensity psden1 omega
. line psden1 omega
```



The above graph is typical of a spectral density of an AR(1) process with a positive coefficient. The curve is highest at frequency 0, and it tapers off toward zero or a positive asymptote. The estimated spectral density is telling us that the low-frequency random components are the most important random components of an AR(1) process with a positive autoregressive coefficient.

The closer the α is to 1, the more important are the low-frequency components relative to the high-frequency components. To illustrate this point, we plot the spectral densities implied by AR(1) models with $\alpha = 0.1$ and $\alpha = 0.9$.



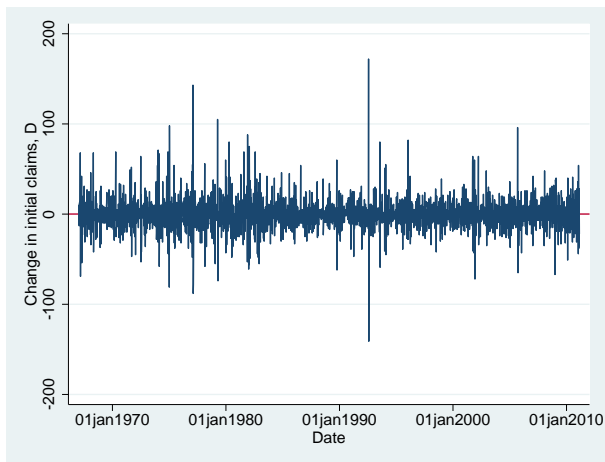
As α gets closer to 1, the plot of the spectral density gets closer to being a spike at frequency 0, implying that only the lowest-frequency components are important.

◀

▷ Example 2

Now let's consider a dataset for which the estimated coefficient from an AR(1) model is negative. Below we plot the changes in initial claims for unemployment insurance in the United States.

```
. use https://www.stata-press.com/data/r17/icsa1, clear
. tsline D.icsa, yline(0.08)
```



The plot looks a little more jagged than we would expect from an i.i.d. process, but it is hard to tell. Below we estimate the AR(1) coefficient.

```
. arima D.icsa, ar(1) noconstant
(setting optimization to BHHH)
Iteration 0: log likelihood = -9934.0659
Iteration 1: log likelihood = -9934.0657
Iteration 2: log likelihood = -9934.0657
```

ARIMA regression

Sample: 14jan1967 thru 19feb2011

Number of obs = 2302
Wald chi2(1) = 666.06
Prob > chi2 = 0.0000

Log likelihood = -9934.066

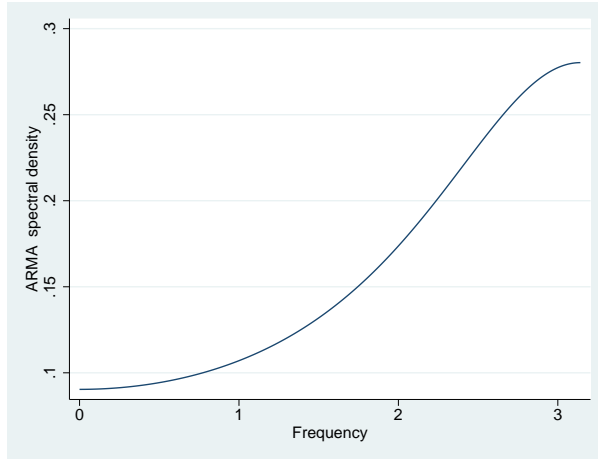
D.icsa	OPG		z	P> z	[95% conf. interval]	
	Coefficient	std. err.				
ARMA						
ar						
L1.	-.2756024	.0106789	-25.81	0.000	-.2965326	-.2546722
/sigma	18.10988	.1176556	153.92	0.000	17.87928	18.34048

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at zero.

The estimated coefficient is negative and statistically significant.

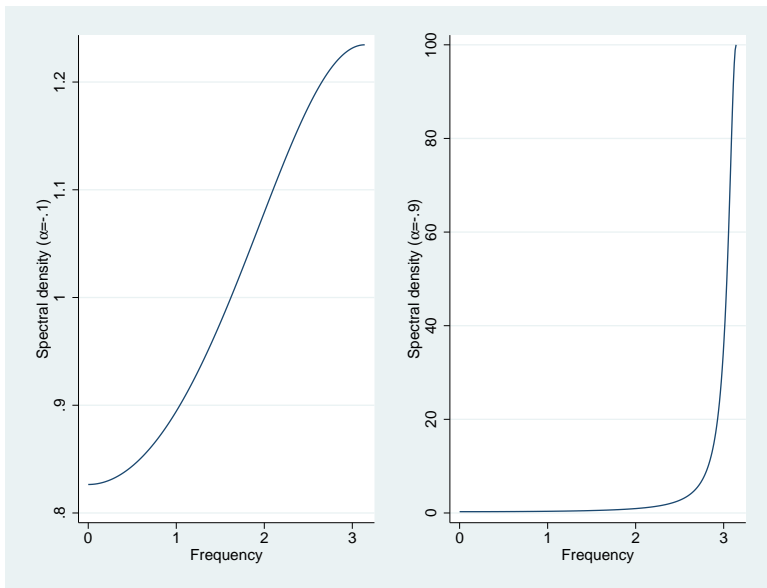
The spectral density implied by the estimated parameters is

```
. psdensity psden2 omega2
. line psden2 omega2
```



The above graph is typical of a spectral density of an AR(1) process with a negative coefficient. The curve is lowest at frequency 0, and it monotonically increases to its highest point, which occurs when the frequency is π .

When the coefficient of an AR(1) model is negative, the high-frequency random components are the most important random components of the process. The closer the α is to -1 , the more important are the high-frequency components relative to the low-frequency components. To illustrate this point, we plot the spectral densities implied by AR(1) models with $\alpha = -0.1$, and $\alpha = -0.9$.



As α gets closer to -1 , the plot of the spectral density shifts toward becoming a spike at frequency π , implying that only the highest-frequency components are important.

◀

For examples of `psdensity` after `arfima` and `ucm`, see [TS] [arfima](#) and [TS] [ucm](#).

Methods and formulas

Methods and formulas are presented under the following headings:

[Introduction](#)

[Spectral density after arima or arfima](#)

[Spectral density after ucm](#)

Introduction

The spectral density $f(\omega)$ is estimated at the values $\omega \in \{\omega_1, \omega_2, \dots, \omega_N\}$ using one of the formulas given below. Given a sample of size N , after accounting for any `if` or `in` restrictions, the N values of ω are given by $\omega_i = \pi(i - 1)/(N - 1)$ for $i \in \{1, 2, \dots, N\}$.

In the rare case in which the dataset in memory has insufficient observations for the desired resolution of the estimated spectral density, you may use `tsappend` or `set obs` (see [TS] [tsappend](#) or [D] [obs](#)) to increase the number of observations in the current dataset.

You may use an `if` restriction or an `in` restriction to restrict the observations to handle panel data or to compute the estimates for a subset of the observations.

Spectral density after arima or arfima

Let ϕ_k and θ_k denote the p autoregressive and q moving-average parameters of an ARMA model, respectively. [Box, Jenkins, and Reinsel \(2008\)](#) show that the spectral density implied by the ARMA parameters is

$$f_{\text{ARMA}}(\omega; \phi, \theta, \sigma_\epsilon^2, \gamma_0) = \frac{\sigma_\epsilon^2}{2\pi\gamma_0} \frac{|1 + \theta_1 e^{-i\omega} + \theta_2 e^{-i2\omega} + \dots + \theta_q e^{-iq\omega}|^2}{|1 - \phi_1 e^{-i\omega} - \phi_2 e^{-i2\omega} - \dots - \phi_p e^{-ip\omega}|^2}$$

where $\omega \in [0, \pi]$ and σ_ϵ^2 is the variance of the idiosyncratic error and γ_0 is the variance of the dependent variable. We estimate γ_0 using the `arima` parameter estimates.

The spectral density for the ARFIMA model is

$$f_{\text{ARFIMA}}(\omega; \phi, \theta, d, \sigma_\epsilon^2, \gamma_0) = |1 - e^{i\omega}|^{-2d} f_{\text{ARMA}}(\omega; \phi, \theta, \sigma_\epsilon^2)$$

where d , $-1/2 < d < 1/2$, is the fractional integration parameter. The spectral density goes to infinity as the frequency approaches 0 for $0 < d < 1/2$, and it is zero at frequency 0 for $-1/2 < d < 0$.

The `smemory` option causes `psdensity` to perform the estimation with $d = 0$, which is equivalent to estimating the spectral density of the fractionally differenced series.

The power spectrum omits scaling by γ_0 .

Spectral density after ucm

The spectral density of an order- k stochastic cycle with frequency λ and damping ρ is (Trimbur 2006)

$$f(\omega; \rho, \lambda, \sigma_{\kappa}^2) = \left\{ \frac{(1 - \rho^2)^{2k-1}}{\sigma_{\kappa}^2 \sum_{i=0}^{k-1} \binom{k-1}{i}^2 \rho^{2i}} \right\} \times \frac{\sum_{j=0}^k \sum_{i=0}^k (-1)^{j+i} \binom{k}{j} \binom{k}{i} \rho^{j+i} \cos \lambda(j-i) \cos \omega(j-i)}{2\pi \{1 + 4\rho^2 \cos^2 \lambda + \rho^4 - 4\rho(1 + \rho^2) \cos \lambda \cos \omega + 2\rho^2 \cos 2\omega\}^k}$$

where σ_{κ}^2 is the variance of the cycle error term.

The variance of the cycle is

$$\sigma_{\omega}^2 = \sigma_{\kappa}^2 \frac{\sum_{i=0}^{k-1} \binom{k-1}{i}^2 \rho^{2i}}{(1 - \rho^2)^{2k-1}}$$

and the power spectrum omits scaling by σ_{ω}^2 .

References

- Box, G. E. P., G. M. Jenkins, and G. C. Reinsel. 2008. *Time Series Analysis: Forecasting and Control*. 4th ed. Hoboken, NJ: Wiley.
- Fuller, W. A. 1996. *Introduction to Statistical Time Series*. 2nd ed. New York: Wiley.
- Hamilton, J. D. 1994. *Time Series Analysis*. Princeton, NJ: Princeton University Press.
- Harvey, A. C. 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- . 1993. *Time Series Models*. 2nd ed. Cambridge, MA: MIT Press.
- Priestley, M. B. 1981. *Spectral Analysis and Time Series*. London: Academic Press.
- Trimbur, T. M. 2006. Properties of higher order stochastic cycles. *Journal of Time Series Analysis* 27: 1–17. <http://doi.org/10.1111/j.0143-9782.2005.00462.x>.
- Wei, W. W. S. 2006. *Time Series Analysis: Univariate and Multivariate Methods*. 2nd ed. Boston: Pearson.

Also see

- [TS] **arfima** — Autoregressive fractionally integrated moving-average models
- [TS] **arima** — ARIMA, ARMAX, and other dynamic regression models
- [TS] **ucm** — Unobserved-components model