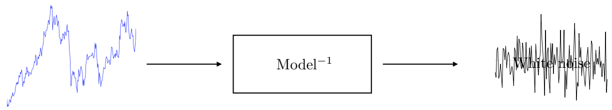
Residual analysis

Andreas Jakobsson

Residual analysis

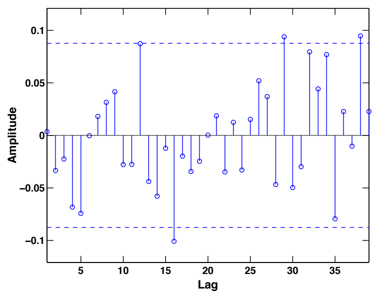


How can one determine if a sequence is "reasonably" white? In this course, we will examine a few different statistical test, namely

- Testing the ACF and the PACF
  - Plot and eyeball!
  - The Box-Pierce test
  - The Ljung-Box-Pierce test
  - The McLeod-Li test
  - The Monti test
- Testing the cumulative periodogram
- Testing sign changes

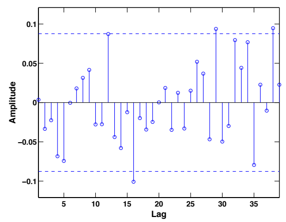
Important: *The first set of tests all assume that the ACF is Gaussian distributed, and may be unreliable if this assumption is violated!*

Testing the ACF and the PACF



Remember that it is 95% confidence intervals; 1 in 20 may be slightly outside.

Testing the ACF and the PACF



The Box-Pierce Q-statistic tests the hypothesis that the initial ACF estimates are not significantly different from zero, forming the Q-statistic as

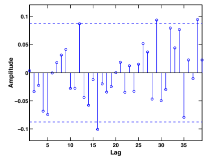
$$Q = N \sum_{\ell=1}^{K} \hat{\rho}_e^2(\ell)$$

where K is the number of considered correlations. Thus, with significance α, one may (asymptotically) reject the hypothesis that the residual is white if

$$Q > \chi^2_{1-\alpha}(K)$$

where $\chi^2_{1-\alpha}(K)$ is the α-quantile of the χ^2-distribution with K degrees of freedom.

Testing the ACF and the PACF



To better handle the small sample case, the Ljung-Box-Pierce Q-statistic tests introduce a scaling, instead forming

$$Q^* = N(N+2)\sum_{\ell=1}^{K} \frac{\hat{\rho}_e^2(\ell)}{N-\ell}$$

One may also examine the squared residual and form the McLeod-Li test,

$$Q_2^* = N(N+2)\sum_{\ell=1}^{K} \frac{\hat{\rho}_{e^2}^2(\ell)}{N-\ell}$$

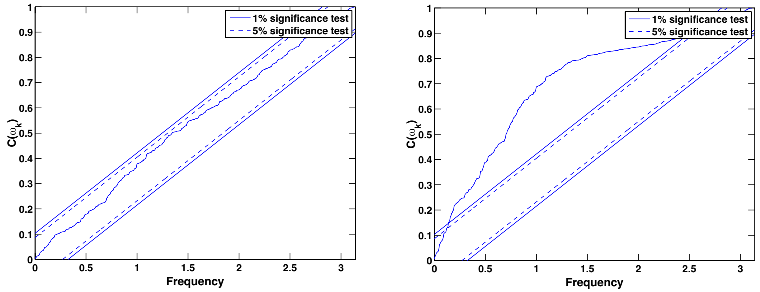
One may instead test the PACF, forming the Monti test

$$Q_M = N(N+2)\sum_{\ell=1}^{K} \frac{\hat{\phi}_{\ell,\ell}^2}{N-\ell}$$

The test statistic Q^*, Q_2^*, and Q_M are all $\chi^2(K)$ for a white signal. In general, the McLeod-Li test depends strongly on the Gaussian assumption and is often unreliable.

These tests are available using the functions `lbpTest`, `mlTest`, and `montiTest`.

Testing the cumulative periodogram



As the (theoretical) spectrum of a white noise is flat, one can compare how well the sum of the periodogram, up to a given frequency ω_k, follows a straight line:

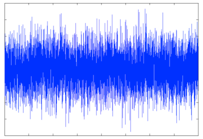
α	0.01	0.05	0.10	0.25
K_α	1.63	1.36	1.22	1.02

$$\phi_c(\omega_k) = \frac{\sum_{\ell=1}^{k} \hat{\phi}_y(\omega_\ell)}{N\hat{\sigma}^2} = \frac{\frac{1}{N}\sum_{\ell=1}^{k} \left|\sum_{t=1}^{N} y_t e^{-i\omega_\ell t}\right|^2}{N\hat{\sigma}_y^2} = \frac{\sum_{\ell=1}^{k} \hat{\phi}_y(\omega_\ell)}{\sum_{\ell=1}^{N} \hat{\phi}_y(\omega_\ell)}$$

The corresponding confidence interval is $\pm K_\alpha/\sqrt{\lfloor N-1/2 \rfloor}$, above and below the theoretical line. The coefficient K_α determines the $(1-\alpha)$ probability that y_t is white noise.

This test is available using the function `plotCumPer`.

Testing sign changes



One can also examine the number of times the signal change sign, P. For a white signal, this should be about every second sample

$$P \in \mathcal{B}\left(N-1, \frac{1}{2}\right)$$

For large data sets, P may be well approximated as being Normal distributed

$$P \in \mathcal{N}\left(\frac{N-1}{2}, \frac{N-1}{4}\right)$$

The test is available in the provided Matlab function `countSignChanges`.

The above discussed tests have all been collected in the function `whitenessTest`.

Is this normal?

One may use the estimated skewness (asymmetry) and kurtosis (peakiness) to determine if an estimated distribution is Gaussian.

S = skewness = 3rd moment / std^3 (This is 0 for a Gaussian).
K = kurtosis = 4th moment / std^4 (This is 3 for a Gaussian).

The Jarque-Berra test is formed as

$$\gamma_{JB} = \frac{N}{6}\left(\hat{S}^2 + \frac{1}{4}(\hat{K}-3)^2\right)$$

which, for a Gaussian, is χ^2 distribution with two degrees of freedom. The test is provided in the function `jbtest`.

For small samples, the D'Agostino-Pearson's K^2 test is often preferable. It is provided in the function `dagostinoK2test`.

To get a visual alternative, one may use the functions `normplot` or `plotNTdist`.