Progress Report: DMT Gaussianity Test Monitor

Steve Penn, Syracuse University, sdpenn@syr.edu

The Gaussianity Test Monitor can perform three levels of tests on the Gaussianity of a given channel: statistical characterization, gaussian fit, and the Hinich Test for Gaussianity. I will describe each of the tests below. The Gaussianity Test Monitor also has two modes of operation: diagnostic mode, and "background" mode. In "background" mode, the monitor writes a summary of the test results to the screen/file and can run in approximately real time. In diagnostic mode the monitor outputs a series of graphs associated with the above tests that help the user determine the nature of the nonGaussian component of the channel. These graphs cannot currently be generated in real time with the machine/algorithm combination I am using. On the other hand, the graphs are not very interesting unless one of the test summaries indicate that the channel has a strong nonGaussian component.

Statistical Characterization: To characterize the statistical nature of the channel, *x*, we look the cumulants up to 4th order. The cumulants are defined as:

$$C_{1x} = E\{x(n)\} \text{ == mean.}$$
$$C_{2x}(k) = E\{x^*(n)\ x(n+k)\}$$
$$C_{3x}(k,l) = E\{x^*(n)\ x(n+k)\ x(n+l)\}$$
$$C_{4x}(k,l,m) = E\{x^*(n)\ x(n+k)\ x(n+l)\ x^*(n+m)\}$$
$$-C_{2x}(k)\ C_{2x}(l-m) - C_{2x}(l)\ C_{2x}(k-m) - C_{2x}(m)\ C_{2x}(k-l)$$

where *E{…}* is the expectation value. The "zero-lag" cumulants are more commonly referred to as the variance ($C_{2x}(0)$), the skewness ($C_{3x}(0,0)/(C_{2x}(0))^{1.5}$), and the kurtosis ($C_{4x}(0,0,0)/(C_{2x}(0))^2$). If the mean and variance are constant, the signal obeys the weak criteria for stationarity. If the skewness is zero, then *x* is distributed symmetrically about the mean. Finally if the Kurtosis is zero, then *x* has a Gaussian distribution.

Gaussian Fit Test: A quick, intuitive, but not very sensitive method for readily recognizing the Gaussianity of a channel is to histogram the channel and then fit the result to a normal distribution. Often the deviation from an ideal Gaussian occurs far out in the tails where these deviations may be difficult to recognize. Therefore the data is fit twice – once over the entire range and a second time over the central region ($< 2\sigma$) – and the fit parameters are compared. The histogram with fit is also plotted on a "log vs x^2" plot so that the Gaussian is projected like an inverted "V" and it is easy to distinguish the deviations in the tails of the distribution. This test is a nice heuristic tool if the nonGaussian nature of the signal is obvious, but it is easy to produce a nonGaussian signal which appears Gaussian under this test.

Hinich Test for Gaussianity: This final test uses the higher order statistics of a given time series to calculate the probability that the time series contains a nonGaussian component. First we must calculate the bispectrum and the bicoherence. The bispectrum, $S_3(\omega_1, \omega_2)$ is the FT of $C_{3x}(k,l)$ and will be zero for a Gaussian distribution. (Actually $C_{3x}(k,l)$ will be zero for a distribution of zero skewness, and it is the trispectrum that is zero for Gaussian processes. However we use the bispectrum because it is easier). The bicoherence is defined as:

$\mathrm{bic}(\omega_1, \omega_2) = \frac{S_3(\omega_1, \omega_2)}{\sqrt{S_2(\omega_1) S_2(\omega_2) S_2(\omega_1 + \omega_2)}}$. It turns out that for a Gaussian process, the bicoherence will have a chi-squared distribution, thus from the bicoherence distribution we calculate the probability that the the time series is Gaussian.