19. Simulation

In an age of fast computers, and as a powerful test of one's understanding, it is both useful and interesting to be able to generate example time series with known statistical properties. Such simulations can be done both in the frequency and time domains (next Chapter). Suppose we have a power spectral density, $\Phi(s)$—either a theoretical or an estimated one, and we would like to generate a time series having that spectral density.

Consider first a simpler problem. We wish to generate a time series of length N, having a given mean, m, and variance σ^2. There is a trap here. We could generate a time series having exactly this sample mean, and exactly this sample variance. But if our goal is to generate a time series which would be typical of an actual physical realization of a real process having this mean and variance, we must ask whether it is likely any such realization would have these precise sample values. A true coin will have a true (theoretical) mean of 0 (assigning heads as +1, and tails as -1). If we flip a true coin 10 times, the probability that there will be exactly 5 heads and 5 tails is finite. If we flip it 1000 times, the probability of 500 heads and tails is very small, and the probability of "break-even" (being at zero) diminishes with growing data length. A real simulation would have a sample mean which differs from the true mean according to the probability density for sample means for records of that duration. As we have seen above, sample means for Gaussian processes have a probability density which is normal $G\left(0, \sigma^2/N\right)$. If we select each element of our time series from a population which is normal $G(0, \sigma)$, the result will have a statistically sensible sample mean and variance. If we generated 1000 such time series, we would expect the sample means to scatter about the true mean with probability density, $G\left(0, \sigma^2/N\right)$.

So in generating a time series with a given spectral density, we should *not* give it a sample spectral density exactly equal to the one required. Again, if we generated 1000 such time series, and computed their estimated spectral densities, we could expect that their average spectral density would be very close to the required one, with a scattering in a χ^2_ν distribution. How might one do this? One way is to employ our results for the periodogram. Using

$$y_q = \sum_{n=1}^{[T/2]} a_n \cos\left(\frac{2\pi n q}{T}\right) + \sum_{n=1}^{[T/2]-1} b_n \sin\left(\frac{2\pi n q}{T}\right). \tag{19.1}$$

a_n, b_n are generated by a Gaussian

-random number generator $G(0, \Phi(s_n))$ such that $< a_n >=< b_n >= 0$, $< a_n^2 >=< b_n^2 >= \Phi(s = n/T)/2$, $< a_n a_m >=< b_n b_m >= 0, m \neq n$, $< a_n b_m >= 0$. The requirements on the a_n, b_n
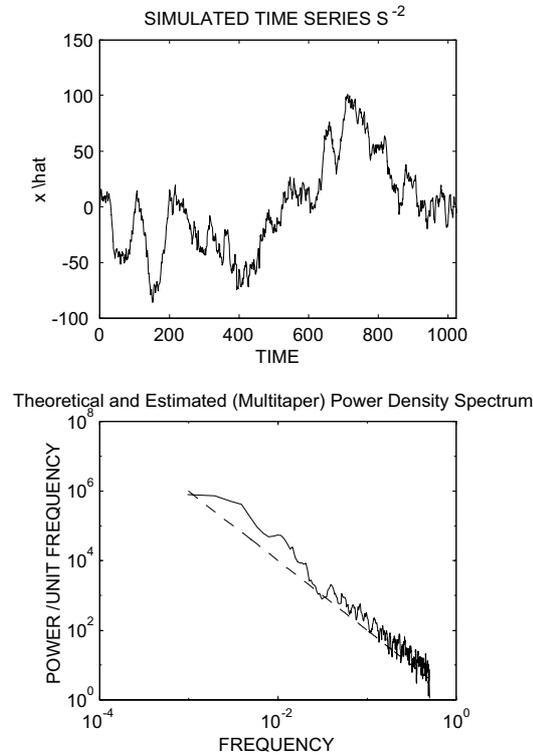
FIGURE 28. A simulated time series with power density proportional to s^{-2}. Although just a "red-noise" process, the eye seems to see oscillatory patterns that are however, ephemeral.

assure wide-sense stationarity. Confirmation of stationarity, and that the appropriate spectral density is reproduced can be simply obtained by considering the behavior of the autocovariance $< y_t y_q >$ (see Percival and Walden, 1993). A unit time step, $t = 0, 1, 2, ...$ is used, and the result is shown in Figure 28.

(This result is rigorously correct only asymptotically as $T \to \infty$, or for white noise. The reason is that for finite record lengths of strongly colored processes, the assumption that $\langle a_n a_{n'} \rangle = \langle b_n b_{n'} \rangle = 0$, etc., is correct only in the limit (e.g., Davenport and Root, 1958)).