1 with fixed τ^2 is easily defined either through its covariance function or its spectral density, but we are not aware of studies of possible natural extensions beyond $d = 1$.

ACKNOWLEDGMENTS

This work was partially supported by the National Science Foundation and the Bureau of Labor Statistics through Westat, Inc.

Comment

Emanuel Parzen

All statisticians should be made aware of the message of Jan Beran's comprehensive and stimulating paper, that the practice of statistics cannot be successful without applying awareness of long memory and long tail behavior in data. The data he discusses, especially "NBS precision measurements on the 1-kg check standard weight," demonstrate that statisticians who analyze data must be aware of the trichotomy that should be answered as one of the first steps in a data analysis: should the data be regarded as white noise (independent or zero memory), short memory (weakly dependent) or long memory (long range dependent). I would like to describe some heuristic concepts that I find useful for understanding, diagnosing, and modeling long range dependence.

Given a time series sample $Y(t)$, $t = 1, \ldots, n$, I define the sample spectral density (or periodogram) as a function of ω, $0 \le \omega < 1$:

$$\tilde{f}(\omega) = \frac{\sum_{t=1}^{n} |Y(t)\exp(-2\pi i t\omega)|^2}{\sum_{t=1}^{n} |Y(t)|^2}.$$

The spectral density $f(\omega)$, $0 \le \omega < 1$, is defined (as a descriptor of the hypothetical population of sample paths in the probability model) as the limit as n tends to ∞ of $E[\tilde{f}(\omega)]$.

A time series is called short memory dependent if $f(\omega)$ is bounded above and below (white noise if the spectral density is constant).

A time series is defined to be long memory if $f(0)$ is infinity (more generally if $f(\omega)$ has zeroes or infinities at some frequency). An important role is played by the spectral density $f(1/n)$ at frequency $1/n$. It can be used to express the variability of the sample mean $\overline{Y}_n = (1/n)\sum_{t=1}^{n} Y(t)$; asympotically $VAR[\overline{Y}_n] = (1/n)f(1/n)C$ for a suitable constant C depending on the index δ of the (self-similar) representation

$$f(\omega) = \omega^{-\delta}L(\omega)$$

where $L(\omega)$ is a slowly varying (log-like) function at $\omega = 0$, and $L(0) > 0$. The Hurst exponent H in Beran's formula (3) corresponds to long memory index $\delta = 1 - 2H$.

In addition to spectral techniques, we recommend changepoint and cusum analysis techniques to provide diagnostics of various types of long memory behavior, based on weak convergence theorems for cusum processes such as

$$\tilde{C}(\tau) = n^{5}\sum_{t=1}^{[n\tau]}\frac{Y(t) - \overline{Y}_n}{\tilde{\sigma}}, \qquad 0 \le \tau < 1,$$

where $\tilde{\sigma}$ is the sample standard deviation.

The cusum process is important for applications to quality control problems of identifying changepoints in the series under the null hypothesis that it is white noise.

To identify if there is long memory dependence and to detect changepoints in correlated data, one approach could be to estimate the exponent $\delta = 1 - 2H$, H the Hurst exponent, in the spectral density formula $f(1/n) = n^{\delta}L(1/n)$.

Values of the "fractal dimension" delta have received much public attention since they are used to describe music and how the brain works. *U.S. News and World Report*, June 11, 1990, writes (p. 62): "Surprisingly, the same mathematical formula that characterizes the ebb and flow of music has been discovered to exist widely in nature, from the flow of the Nile to the beating of the human heart to the wobbling of the earth's axis."

Estimation of δ (which can be considered estimating a "fractal dimension") is a central research problem of the analysis of long memory time series. Beran notes that it has analogies with the problem of estimating the tail index of a long tailed distribution. One expects to estimate how $f(1/n)$ depends on n essentially from the values of the sample spectral density in a band of low frequencies (omitting zero frequency) to be selected by suitable criteria.

Emanuel Parzen is Distinguished Professor of Statistics, Department of Statistics, Texas A & M University, College Station, Texas 77843-3143.