AN HISTORICAL OVERVIEW OF STABLE PROBABILITY DISTRIBUTIONS IN SIGNAL PROCESSING

B.W.Stuck

Signal Lake Venture Fund L.P.
578 Post Road East/Suite 667,
Westport CT 06880-4549 USA,
Email: BartStuck@signallake.com

ABSTRACT

An historical, personal and idiosyncratic overview of stable probability distributions in signal processing is presented.

1. INTRODUCTION

From 1972 to 1976 the author worked at Bell Laboratories in Murray Hill NJ in the Mathematics Research Center. Part of that time was devoted to studying stable probability distributions, and then attempting to understand how signal processing might be impacted if additive noise was perturbed away from a Gaussian probability distribution toward a stable probability distribution that was close (in the sense of having a characteristic exponent less than two). This work was motivated by studying noise on analog telephone lines, which number in the hundreds of millions around the world, and which increasingly are being used to send digital information via dialup modems (Stuck and Kleiner, 1974). In addition, there has been a significant amount of statistical data analysis of fluctuations of logarithms of stock prices which suggests that a family of stable probability distributions would fit to within statistical fluctuations (Mandelbrot, 1963 and 1967; Fama, 1965). Finally, a variety of workers had concentrated on numerical aspects of arbitrary stable distributions (DuMouchel, 1971; Cross, 1974) and statistical parameter estimation (Fama and Roll, 1968; Fama and Roll, 1971; DuMouchel, 1971).

This led to a research program into examining the impact of stable probability distributions on classic statistical hypothesis testing for discrete time finite samples (Stuck, 1976a) and for continuous time analogs (Stuck, 1976b; Newman and Stuck, 1979a, 1979b). A natural generalization of the discrete time Kalman filter was obtained for a class of discrete time moving average autoregressive stable processes (Stuck, 1977). A natural generalization of the well-known Gaussian random variable generation to stable random variable generation was obtained (Chambers, Mallows, Stuck, 1977). Finally, an analysis of investment portfolios was carried out assuming a log stable random walk for the underlying security price (Stuck, 1976). Our intent is to discuss each of these in turn.

II TELEPHONE NOISE

The initial work on statistically characterizing the noise from tape recording measurements on five analog telephone lines was carried out in 1972-1973. This work involved two activities, an exploratory data analysis stage where the data are characterized through various non-parametric statistics, and a model-building stage where the data are matched or fit to models to within statistical fluctuations. The exploratory data analysis stage involved examination of noise waveforms, power spectra, and covariance estimates. The results showed that the data consisted of a deterministic component (sinusoids at various frequencies that are audible tones) plus a stochastic component, which was assumed to be independent. After filtering the data to remove the sinusoids, histograms and empirical cumulative distribution functions for the filtered data were examined, as well as central moment estimates. The filtered data appeared to be wide-sense stationary over short periods of time, typically one second. Based largely on quantile-quantile plots, it was concluded that, although close to Gaussian, the filtered data for three of the five samples were distinctly non-Gaussian; the filtered data for the remaining two lines appeared to be Gaussian.

The model building stage involved fitting the filtered data to two classes of models. The first class of models was stable probability distributions. Based on a series of parameter estimation procedures including robust estimation, maximum likelihood estimation, and quantile-quantile plots, and back up by a likelihood ratio test on the goodness-of-fit, the three non-Gaussian samples could be adequately modeled by a stable distribution with characteristic index of roughly 1.95 (the Gaussian distribution has characteristic index of 2.0). The second class of models were a mixture or sum of two processes, a low-variance component from a stationary Gaussian process, and a high-variance component from a filtered Poisson process. The parameters for the Gaussian compnent were estimated using trimmed means and trimmed variances. The parameters specifying the filtered Poisson process were much more complicated to estimate. The instants of time at which noise bursts occurred and the intervals between bursts were first examined; based on power spectra as well as covariance estimates, the intervals appeared to come from a renewal process. Histograms and empirical cumulative distribution functions indicated that the time intervals came from a Poisson process; empirical survivor and hazard function plots showed that a Poisson process with constant rate parameter was not an adequate model, however. Because of the small number of bursts observed, it was quite difficult to fit the time intervals to a Poisson process with varying rate parameter, and for expediency a constant Poisson rate parameter was chosen to model noise burst times of occurrence. The amplitudes of the noise bursts were adequately modeled by a log normal and power Rayleigh, or generalized gamma. The durations of actual noise bursts were used to estimate parameters in the noise burst-shaping filter. A simple indication is presented of how well the filtered data fit the Gaussian-plus-filtered-Poisson-process model.

III HYPOTHESIS TESTING

The problem of classifying a series of observations as coming from one of two or more possible classes or hypotheses has received a great deal of attention in the statistical and engineering literature. In many physical situations, a variety of disturbances corrupt the observations; rather than model each disturbance separately, it is often argued on physical grounds that the disturbances add and are independent, and the central limit theorem of mathematics is invoked to model this sum using a Gaussian distribution. This approach is adequate as long as the sum of disturbances is not dominated by one or a few of the summands; if one or a few of the summands does dominate the sum total disturbance, the disturbances can possibly be modeled as a stable distribution, using the generalized central limit theorem of mathematics.

The Gaussian distribution has enjoyed great popularity in hypothesis testing because it is analytically tractable and because it is the only stable distribution with finite variance. Although it may be argued that mathematical models with infinite variance are physically inappropriate, this view conveniently overlooks the fact that the Gaussian distribution is unbounded, which is also a physically inappropriate mathematical model. The Gaussian model may adequately model disturbances over a narrow range of amplitudes; an infinite variance stable distribution model may adequately model disturbances over a larger range of amplitudes. Both distributions may be physically inappropriate mathematical models, but the infinite variance distribution may, in this sense, be the better model.

One of the difficulties with carrying out statistical hypothesis testing using arbitrary statistical distributions is the underlying numerical analysis can be quite challenging compared with that for Gaussian distributions. Furthermore, closed form analytical expressions that would provide useful rules of thumb can be few and far between. On the other hand, with the advent of the relatively inexpensive personal computers that are capable of performing hundreds of millions of calculations per second, the utility of closed form analytic expressions can in some cases be mitigated by having a graphical display of numerical calculations using parameter estimates, and this by itself can be quite useful. Hence, the bulk of the hypothesis testing work was presented in numerical graphical form.

Two special cases were examined in detail, when the underlying distributions differ only in location, and when they differ only in scale. The probabilities of error of the first and second kind are found for three analytically tractable cases (Gaussian, Cauchy, and Pearson V) by calculating the characteristic function of the log likelihood probability measure induced under each hypothesis; the general case is apparently analytically intractable, and quite challenging from a numerical analysis vantage point. Exponentially sharp upper and lower bounds on both types of probabilities of error, and also the total probability of error, can be simply derived from the Laplace transform of the log likelihood probability measure induced under each hypothesis. These bounds are found analytically in three cases, and relatively inexpensive numerical results are presented for selected other cases.

When the two distributions differ only in location, the likelihood ratio test is shown to be extremely sensitive to whether the distribution is non-Gaussian, when nonlinear soft limiting of large deviations is employed, or Gaussian, when linear processing is used. When the distribution is non-Gaussian stable, performance is found analytically to be quite sensitive to whether a linear (suboptimum) or likelihood (optimum) decision rule is used; asymptotics are developed.

When the two distributions differ only in scale, the likelihood-ratio test is extremely sensitive to whether the distribution is non-Gaussian stable when nonlinear soft limiting of large deviations is used, or Gaussian when a chi-squared test is used. Performance for non-Gaussian stable distributions is extremely sensitive to whether a suboptimum (chi-squared) or optimum (likelihood ratio) test is used; asymptotics are developed.

The analysis of the two remaining cases, distinguishing between one of two characteristic indices and between one of two skewness parameters, closely parallels the analysis that distinguishes between two scale factors.

IV MINIMUM ERROR DISPERSION LINEAR FILTERING OF SCALAR SYMMETRIC STABLE PROCESSES

The well known Kalman-Bucy linear filtering theory for Gaussian Markov processes is generalized (Stuck, 1977) to cover a particular class of non-Gaussian Markov processes, scalar symmetric stable Markov processes. Results are presented for discrete time that are quite analogous to those for Gaussian Markov processes. On the other hand, there is no analogous Wiener-Hopf spectral factorization theory for this class of problems, and the natural extension to the multivariate case is still an open issue.

V STATISTICAL ANALYSIS OF CONTINUOUS TIME PROCESSES

A number of works dealt with continuous time analogs of discrete time processes (Stuck, 1976b; Newman and Stuck, 1979a and 1979b). These works have received limited attention because of the sample path pathologies of continuous time independent increment processes and continuous time Markov processes that are generated from Gaussian distributions or from stable distributions.

BIBLIOGRAPHY:

[1]Cross, M-J, “Tables of Finite Mean Nonsymmetric Stable Distributions as Computed From Their Convergent and Asymptotic Series,” **(1974)**, J. Statistical Computation and Simulation, 1-27.

[2]DuMouchel, W.H., “Stable Distributions in Statistical Inference,” **(1971)**, Department of Statistics, Yale University Ph.D. Dissertation

[3]Fama, E.F., “Behavior of Stock Market Prices,” **(1965)**, J.Business, 34-105.

[4]Fama, E.F. and Roll, R., “Properties of Symmetric Stable Distributions,” **(1968)**, J. American Statistical Association, 817-836.

[5]Fama, E.F. and Roll, R., “Parameter Estimates for Symmetric Stable Distributions,” **(1971)**, J. American Statistical Association, 331-338.

[6]Mandelbrot, B. “The Variation of Certain Speculative Prices,” **(1963)**, J.Business, 394-419.

Mandelbrot, B. “The Variation of Some Other Speculative Prices,” **(1967)**, J.Business, 393-413.

[7]Newman, C.M. and Stuck, B.W. **(1979a)** “Optimal Sampling of Independent Increment Processes,” Stochastics, 213-225.

Newman, C.M. and Stuck, B.W. **(1979b)**, “Chernoff Bounds for Discriminating Between Two Markov Processes,” Stochastics, 139-153.

[8]Stuck, B.W. and Kleiner, B. **(1974)** “A Statistical Analysis of Telephone Noise,” Bell System Technical Journal **(53)**, 1263-1320

[9]Stuck, B.W. **(1976a)** “Distinguishing Stable Probability Measures, Part I: Discrete Time,” Bell System Technical Journal **(55)**, 1125-1182.

[10]Stuck, B.W. **(1976b)** “Distinguishing Stable Probability Measures, Part II: Continuous Time,” Bell System Technical Journal **(55)**, 1183-1196.

[11]Chambers, J.M., Mallows, C.L. and Stuck, B.W. **(1976c)** “A Method for Simulating Stable Random Variables,” Journal of the American Statistical Association **(71)**, 1340-1344.

[12]Stuck, B.W. **(1976d)** “Explicit Solutions to Some Single-Period Investment Problems for risky Log-Stable Stocks,” Journal of Financial Economics **(3)**, 277-294.

[13]Stuck, B.W. **(1977)** “Minimum Error Dispersion Linear Filtering of Scalar Symmetric Stable Processes,” IEEE Transactions on Automatic Control **(AC-23)**, 507-509.