never forget the law of statistical entropy[1]

by **David** M. Rowe, Ph.D.

At the heart of most risk analysis techniques developed starting in the mid-1980s lies the methods and assumptions of classical statistics. Central to these methods is the assumption of a "stable random process". In such a process, the value of any single draw is unknowable in advance but sizable samples will exhibit broadly similar characteristics such as:

- the mean,
- the dispersion (standard deviation),
- the degree of symmetry or lack thereof (skewness),
- the tendency for probability in the tails to dwindle rapidly or slowly (kurtosis).

Given the assumption of stable randomness, the larger the sample size, the more nearly identical will such characteristics be across randomly selected sets.

Classical statistical analysis recognizes that sampling techniques can never produce fully exact values for these characteristics and has developed measures for the uncertainty of such estimates. The most common of these is the "standard error of estimate," which is simply the standard deviation of the implied distribution of possible values for the true underlying parameter.

What is vital for general business executives to remember, however, is that these errors of estimate assume stability of the underlying stochastic process. This is often a realistic assumption when dealing with physical processes. It is virtually never the case, however, in a social scientific setting. Structural change is the constant bane of econometric forecasters. Such changes are driven by a wide variety of influences including technological advances, demographic shifts, political upheavals[2], natural disasters and, perhaps most importantly, behavioral feedback loops.

Structural change creates a fundamental dilemma for socio-statistical analysis. Classical statistics argues that the more data the better since, assuming stochastic stability, this results in smaller estimation errors. For analysis based on time series, however, a larger data set implies incorporation of a greater variety of structural changes that undermine the practical relevance of the classical assumption of stochastic stability.

1 / This essay is a slightly edited excerpt from the author's recently published book An Insider's Guide to Risk Management – Relearning the Lessons of the Global Financial Crisis. The printed book is available from both www.amazon.com and www.barnesandnoble.com. It also is available as an iBook from the Apple App Store.

2 / Sometime in the early 1980s I came across an old working paper entitled An Econometric Model of Iran. Unfortunately it had been written in 1978, a year prior to the Iranian Revolution! This is one of the most dramatic instances of being blindsided by structural change that I can recall.

This makes it all the more important for risk managers to focus obsessively on the law of "statistical entropy". Like water, information can never rise higher than its source. In the case of information, that source is the set of data on which an analysis is based. In assessing the reliability of any risk estimate, including such things as credit ratings, always start with a review of the volume and quality of the available data. No amount of complex mathematical/statistical analysis can possibly squeeze more information from a data set than it contains initially. Indeed, in complex settings it is virtually impossible to extract 100% of the information that does exist. Something is always lost in the process of aggregating and summarizing. This is why I refer to the "Law of Statistical Entropy" rather than the "Law of the Conservation of Information," drawing an analogy to the Second rather than the First Law of Thermodynamics.

A glaring example of failure to focus on the weakness of the available data was the way many banks and investors blindly accepted the AAA rating for senior tranches of subprime mortgage portfolios in the years preceding the onset of the Global Financial Crisis. Before the crisis, such holdings were often treated as equivalent to AAA corporate bonds. Rating agencies have about a century of experience in rating such bonds. This provides a wealth of experience and data to support the effort. Subprime mortgages were a fairly recent phenomenon, and their default experience had been dominated by a period of comparatively benign housing markets with stable to rising prices. Determining how much subordination was necessary to bring the chance of any failure of timely payment of principal or interest down to a target level required making an estimate of behavior deep into the tail of the default distribution.

A casual look at the available data for conducting this analysis should have made one thing clear. Any estimate of the required level of subordination would necessarily be surrounded by significant uncertainty. We know that this market was undermined by serious erosion in underwriting standards to meet the apparently insatiable appetite for these securities in 2005 and 2006. Even before consideration of this type of structural change, however, the limited volume of data supporting the original AAA rating alone should have made banks wary of building up uncontrolled volumes of such securities.

The fundamental lesson to take from this experience is always to ask how much uncertainty surrounds risk estimates given the volume and applicability of the available data. When such uncertainty is clearly excessive, be especially cautious in taking on corresponding exposures.

author

David M. Rowe, PhD

David M. Rowe wrote the monthly Risk Analysis column in Risk magazine from 1999 through late 2015. He has over 40 years of experience at the interface between economic forecasting, finance, and risk management with the rapidly changing world of information technology. His professional career included years spent at Wharton Econometric Forecasting Associates, Townsend-Greenspan & Co., Security Pacific Bank, Bank of America, SunGard and Misys as well as his own small consulting firm. Dr. Rowe is also a former board member of PRMIA.