Article from:

Risks & Rewards

February 2012 – Issue 59

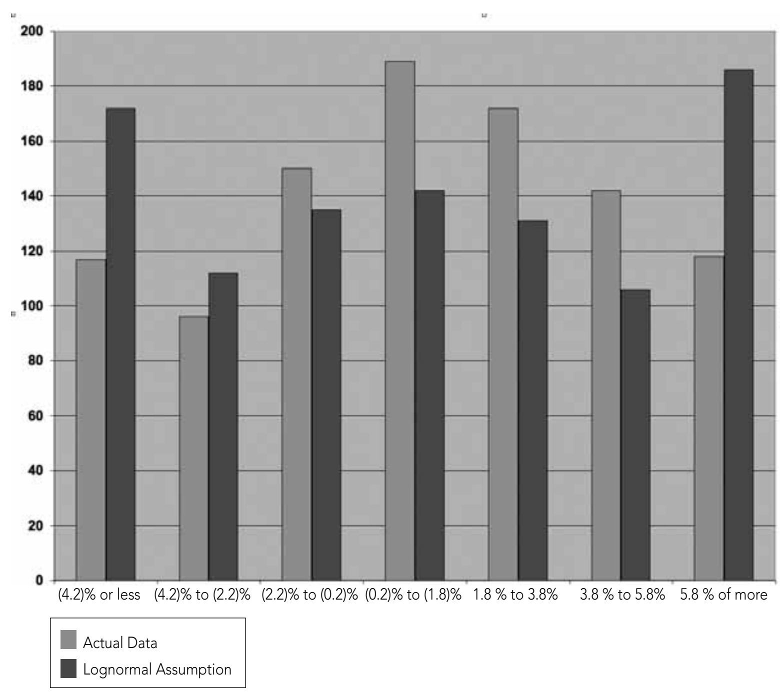
A FRESH LOOK AT LOGNORMAL FORECASTING

By Richard R. Joss

One of the significant contributions of modern academic finance has been to introduce the concept of stochastic investment forecasting. For example, using a Monte Carlo simulation, a forecaster can use actual historical returns to create a whole series of possible future scenarios. Using this technique it is possible not only to provide an expected rate of investment return, but a complete distribution of such returns. In short, using the Monte Carlo technique one could say that the expected return on large company stock investments might be 12 percent, but that there is a 30 percent chance that such an investment could exceed a return of 25 percent for the year. On the down side, it is also possible to say that there is a 30 percent chance that the equity investment could lose money for the year.

Instead of using the Monte Carlo technique, it is possible to create a mathematically derived formula that can be used to create the probability distribution of stock returns.

Actual Data Distribution Compared with Lognormal Assumption



It has been common to assume that this distribution may be described by a lognormal probability density function. (See Appendix 1) Once the parameters mu and sigma are selected for the lognormal distribution, the mathematical approach may be used to provide probabilistic forecasts that are equivalent to forecasts developed using the Monte Carlo techniques.

However, actual experience (such as that exhibited by 401(k) plan participants) has fallen short of expected results. This 401(k) shortfall even made the Nightly News on NBC on Feb. 27, 2011, and was the lead article in the Oct. 19, 2009, issue of TIME Magazine. Both these general news sources cited studies showing that the average near-term 401(k) retiree only had about 25 percent of the funds that he or she was expected to have in order to be able to retire. Thus, the shortfall is really quite significant. While it is easy to blame the markets or poor investment choices on the part of participants as a significant part of the shortfall, perhaps faulty forecasting is also a contributing factor. With that as background, this article takes a fresh look at lognormal forecasting.

LOGNORMAL FORECASTING

As noted above, it has been common to assume that distributions of stock market returns may be modeled using the lognormal probability density function. To select the lognormal probability density function parameters, finance textbooks provide detailed instructions using the arithmetic mean and sample standard deviation from a set of historical returns. What is often missing, however, is a comparison of the actual historical results, and the expected results provided by the lognormal probability density function. This comparison is not as good as one might expect given the widespread use of this particular model. To illustrate this point, the *2008 Ibbotson and Associates SBBI Yearbook* provides of history of 984 months of large company stock return data. The chart to the left compares the distribution of the actual data with the expected distribution provided by the best estimate lognormal density function.

As an example of the difference between the two distributions, the actual distribution shows that for 118 of the 984 months (12 percent of the total) stock returns were 5.8 percent or more for the month. Whereas the best estimate lognormal density function assumes that 186 out of 984 months (19 percent of the total) will have a return that is 5.8 percent or more in the future. This is a substantial difference. It calls into question the use of the basic lognormal probability density function to describe the historical data, and seems to indicate that there may be a fundamental problem with the common lognormal approach.

USING CONDITIONAL PROBABILITIES

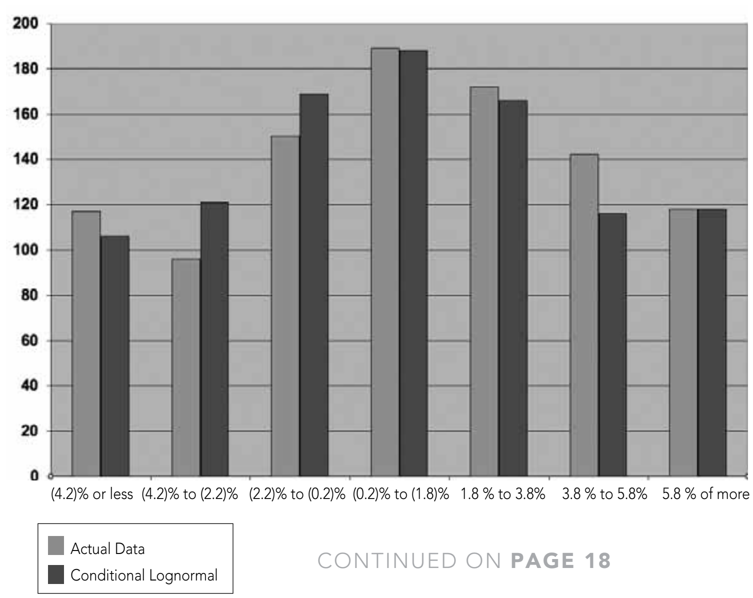
It is interesting to note that the traditional method of selecting lognormal parameters involves the use of the arithmetic mean of a set of historical data. The arithmetic mean of a set of historical data must, of mathematical necessity, always exceed the actual rate of wealth growth. If a fund is to grow at a 5 percent annual rate for a given day, the arithmetic mean of the hourly returns (when expressed as annualized values) must exceed this number. If an investment is to grow at a 0 percent rate for a given month, the arithmetic mean of the annualized daily values must be positive. In each case these higher arithmetic means are just a natural byproduct of the wealth accumulation process. The higher arithmetic means add nothing to the ending wealth.

It has been common in academic finance to say that the best estimate of next year's return will be 12 percent. But the only way that this can occur is if the monthly returns for the year exceed their long-term average of 1 percent. The monthly returns would have to have an average of 1.2 percent or so in order for the end of year wealth to be at the 12 percent rate, assuming stocks exhibit their normal volatility. If it is assumed that the arithmetic mean of the returns for the next 12 months will be the 1 percent historical average, the annual return for the year must be a number that is less than the historically observed 12 percent return, perhaps a number like the geometric mean of 10 percent. It is mathematically impossible to have both the 1 percent monthly rate and the 12 percent annual rate occur simultaneously, if one assumes stocks exhibit normal volatility.

As one contemplates the source of historical investment return data, it is clear that they are periodic observations of a single long-term historical asset growth. As such, the mathematical theory of probability and statistics would place this single observation at the mean of long-term results, with each of the periodic returns being described by a conditional lognormal probability density function. (See Appendix 1) When this one change is made, the comparison between the actual historical results and those described by the probability density function improves dramatically, as is shown in the following chart.

Not only is this large company stock return comparison improved, but the same level of improvement is seen if one does a similar comparison with other data sources, such as *SBBI Yearbook* data for stock returns in small companies or for the Dow Jones Industrial Average. The concept of using a conditional probability to match historical data is not only well-grounded based on the underlying mathematics coupled with the source of actual historical data, tests using actual data confirm the improvement.

Actual Data Distribution Compared with Conditional Lognormal Assumption



CONTINUED ON **PAGE 18**

"NOT ONLY IS THE COMPARISON SIGNIFICANTLY IMPROVED, BUT THIS ONE CHANGE HELPS EXPLAIN THE DISASTROUS 401(K) PLAN RESULTS THAT HAVE BEEN SEEN."

FINANCIAL IMPACT

Not only is the comparison significantly improved, but this one change helps explain the disastrous 401(k) plan results that have been seen. This change in density function causes the best estimate rate of a future return to change from an arithmetic mean of historical returns to the lower geometric mean of historical returns. Given that employee participants have been led to believe that they would receive the higher arithmetic mean returns, it is not surprising that they are disappointed with the actual geometric mean results.

In addition, this one conceptual change helps explain some of the turmoil that has been seen recently in the financial services industry. When followed through completely, the concept that historical data is conditional data, not uncorrelated data, helps explain the collapse of Long-Term Capital Management about 15 years ago, and the more recent collapses of Bear Stearns and Lehman Brothers Holdings.

SUMMARY

Actuaries are used to dealing with data. And it is common for them to consider the appropriateness of historical data when using the data to make forecasts. For example, using data from smokers to make estimates of general population mortality is clearly unwarranted. In this article, actuaries are asked to take a second look at investment return data. The data seems to be conditional in nature, and to the extent that it is, treating it as if it is determined independently may also be unwarranted. This one concept could be critical in dealing with the recent financial crises that has created so much concern.

Richard R. Joss, FSA, is retired. He can be contacted at rrjoss@comcast.net

The formula below is for the traditional lognormal probability density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \frac{1}{x} \cdot e^{\frac{-(\ln x - \mu)^2}{2\sigma^2}}$$

The formula below is for the conditional lognormal probability density function given the assumption that the one observed result is at the mean of the expected distribution of long-term investment results:

$$c(x) = \frac{1}{x} \cdot \frac{\sqrt{n}}{\sqrt{n-1}} \cdot \frac{1}{\sigma\sqrt{2\pi}}\, e^{\left[\frac{-n(\ln x - (\mu + \frac{\sigma^2}{2}))^2}{2(n-1)\sigma^2}\right]}$$