

# The Role of Conditional Probabilities in Risk Assessment

Richard R. Joss

Presented at the:  
2011 Enterprise Risk Management Symposium  
Society of Actuaries  
March 14-16, 2011

Copyright 2011 by the Society of Actuaries.

All rights reserved by the Society of Actuaries. Permission is granted to make brief excerpts for a published review. Permission is also granted to make limited numbers of copies of items in this monograph for personal, internal, classroom or other instructional use, on condition that the foregoing copyright notice is used so as to give reasonable notice of the Society's copyright. This consent for free limited copying without prior consent of the Society does not extend to making copies for general distribution, for advertising or promotional purposes, for inclusion in new collective works or for resale.

## **Abstract**

It has become common to use historical data as a guide in analyzing future risks. However, the statistical tools used often are based on the assumption that the data (regardless of the source) may be treated as independent data for risk analysis purposes. In some cases, the data is conditional in nature, and the proper tool needs to be one that reflects this characteristic of the data. This paper highlights the impact such a change would have on a variety of risk assessment models, with specific emphasis on investment forecasting.

## **Introduction**

It is common to use mathematical techniques in risk assessment. In fact, these techniques are the cornerstone of the entire insurance industry. The probabilities of any particular event happening are developed using highly sophisticated projections of past results. Traditionally, these methods have worked well, and have served the needs of both the insurance industry and the policyholders.

In using historical data, insurance actuaries are careful to segregate appropriate risk classifications, such as male vs. female or smoker vs. non-smoker, to make sure the risks under consideration are based on the data, which does not introduce unintended bias. In situations where bias is potentially likely to enter, the actuary still must deal with the contingency. For example, if a pension plan is planning to provide a significantly improved early retirement benefit, the actual experience-to-date of the early retirements under the plan may be of little use. The new provision is likely to bring about an increase in utilization.

However, in some risk assessment models, it is fairly clear that the data under consideration is conditional in nature, whereas the tool being used does not reflect this important aspect of the data. This introduces an unintended error into the process, and care needs to be exercised to be sure this error does not lead to significantly inaccurate conclusions. This paper will focus on the types of situations where conditional probabilities may play a greater role in risk modeling, with special emphasis on investment forecasting.

## Conditional Probabilities

To begin, let's look at a simple illustration to show how conditional probabilities can affect conclusions. If someone tosses a single six-sided die, and asks you for the probability that the number on the die comes up a 1, your answer is sure to be  $1/6$ . Each of the six possible outcomes has an equal chance of occurring. If the person tosses a second die, and asks the same question, your answer would again be that the probability of seeing a 1 is still  $1/6$ . However, if the person then adds the condition that the sum of the two dice is 8, this condition affects your answer. Once the condition is added, the probability that either one or both of the dice is showing the number 1 is now zero. The new probability space, given the condition that the sum of the two dice is 8, is different than the regular probability space for a single die. Now, each of the numbers 2 through 6 has a probability of appearing of  $1/5$ , whereas the probability of seeing a 1 is zero.

In this simple illustration, if lots of dice were tossed, the probabilities based on the numbers seen would not differ significantly from the  $1/6$  possibility expected. The law of large numbers kicks in. However, in other illustrations, this same dynamic does not occur and the use of conditional probabilities becomes more critical. One of these situations is the use of history-based stock market forecasting.

## Stock Market Forecasting

It has been common in modern academic finance to assume that historically observed stock market investment returns may be treated as independent events in making an investment forecast for the future. It is as if actually observed returns are written on little balls, the balls are then dropped into a giant rotating bin, and then future returns may be modeled by drawing the balls out of the bin. Sometimes the results are taken directly; other times the results are used to calculate an equity risk premium or return in excess of the risk-free rate of return. On the surface, this approach seems very straight forward.

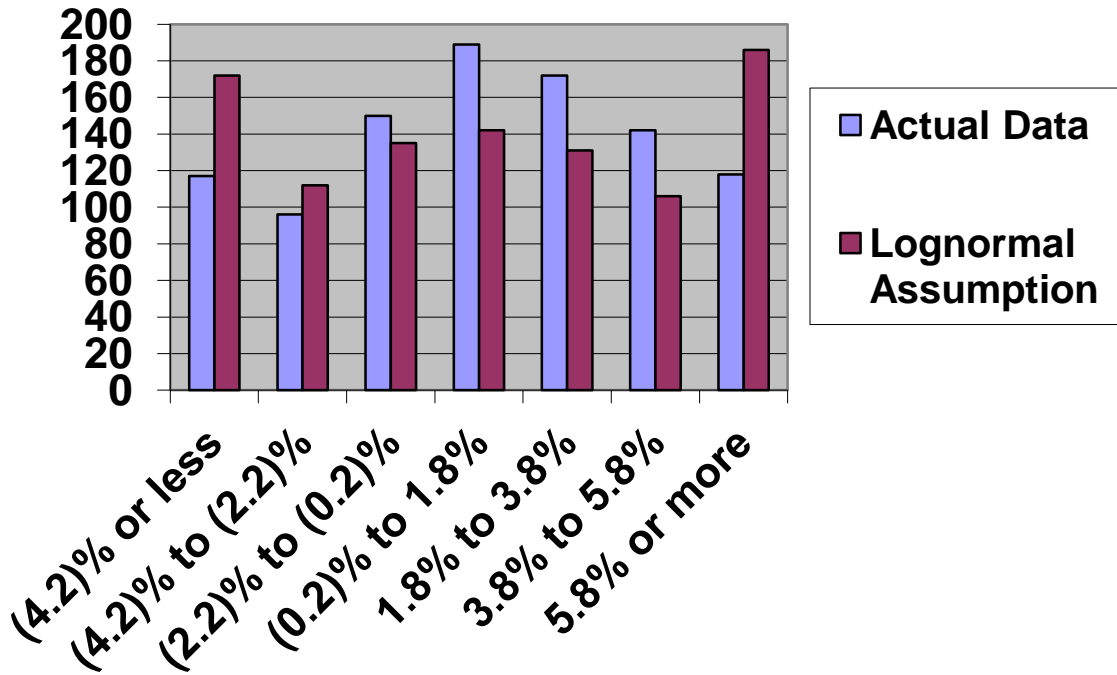
Furthermore, using the actual data in this approach leads to a distribution of future returns that can be described by the lognormal probability density function. Hence, the mechanical process may be replaced with a mathematical formula approach. The outcome of either approach is substantially the same.

Using the lognormal density function tool (or actually drawing results out of a bin), it is possible not only to provide an expected rate of investment return but a complete distribution of such returns. In short, using the tool one could say the expected return on stock investments might be 12 percent but that there is a 30 percent chance your equity investments could exceed a return of 25 percent for the year. On the down side, it is also possible to say there is a 30 percent chance your equity investments could lose money for the year.

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 a history of 984 months of stock return data. The chart below compares the distribution of the actual data with the expected distribution provided by the “best estimate” lognormal density function.

## Actual Data Distribution Compared with Lognormal Assumption



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 189 out of 984 months (19 percent of the total) will have a return of 5.8 percent or more. This is a substantial difference. It calls into question the use of the basic lognormal probability density function to describe the historical data. Keep in mind that the use of the lognormal density function does not produce results that are any different than if the data had been used directly.

### The Conditional Nature of Historical Investment Return Data

The problem with the above model is that all of the data is conditional in nature, not independent. The only way one gets a historical rate of return is to observe actual market behavior. If an investment grows from \$100 to \$110 over a one-year period, it is possible to calculate the rate of return from the actual change in wealth. In this case, the rate of return is 10 percent.

But as one contemplates the source of all historical investment return data, it is clear the data are periodic observations of a single long-term historical asset growth. For example, using the Ibbotson data described above, \$1 invested on Jan. 1, 1926, grew to \$3,246 as of Dec. 31, 2007. Rates of return for any given historical period are then calculated by looking at the change in wealth over the given period. But the basic nature of each of these data sets is conditional.

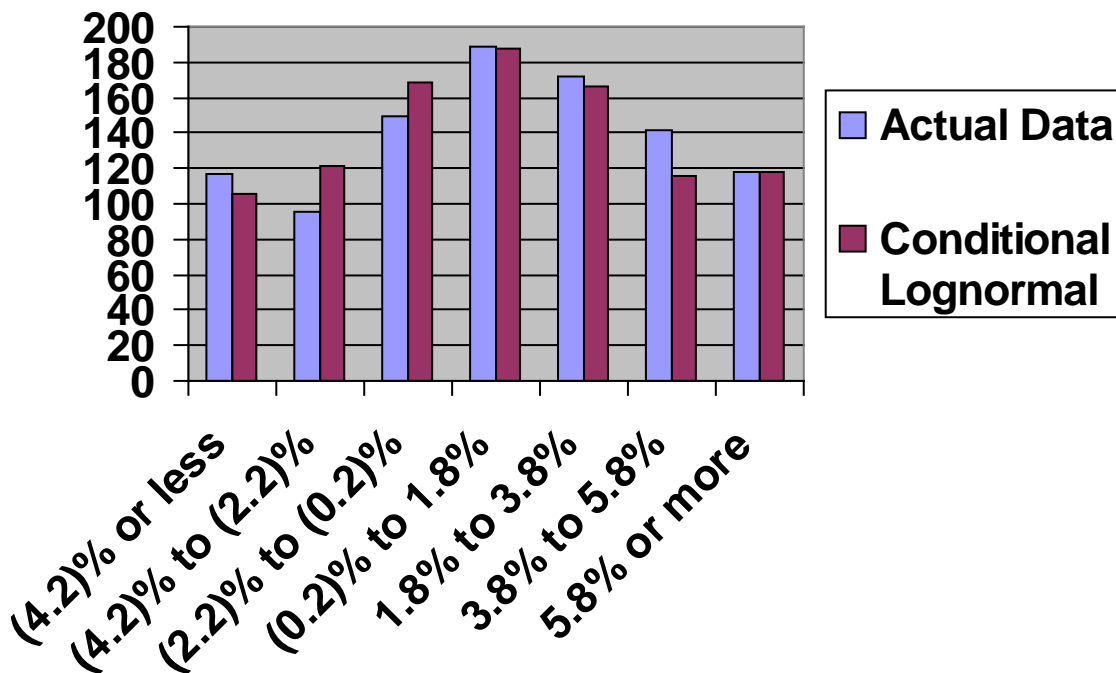
If one is looking at monthly data, there is a data set of 984 monthly returns, given the condition that \$1 invested on Jan. 1, 1926, grew to \$3,246 by Dec. 31, 2007. If one is looking at quarterly returns, there is a data set of 328 quarterly returns, given the condition that \$1 invested on Jan. 1, 1926, grew to \$3,246 by Dec. 31, 2007. Finally, if one is looking at annual returns, there is a data set of 82 annual returns, given the condition that \$1 invested on Jan. 1, 1926, grew to \$3,246 by Dec. 31, 2007.

In these illustrations, the condition on investment growth is similar to the sum of two dice being 8 in the simple conditional probability illustration. It has a bearing on the appropriate probability density function that describes the data. And given the poor “match” between the actual historical data and the common lognormal assumption, consideration of a different density function appears to be warranted.

### The Conditional Lognormal Density Function

The mathematical theory of probability and statistics would place the single observation of investment growth from \$1 in 1926 to \$3,246 in 2007 at the mean of long-term results, with each of the periodic returns being described by a conditional lognormal probability density function. (For completeness, the formula for the conditional lognormal density function is attached as 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.

**Actual Data Distribution Compared with Conditional Lognormal Assumption**



Of course, the match is never perfect. In one case, the data is the actually observed historical returns, and in the other it is data created by a mathematical formula specifically selected to describe the data. However, the improvement in the match between the data and the formula is significant, especially when it is compared with the previous chart.

### **Impact of Making the Change**

Not only is the comparison significantly improved, but this one change helps explain several problems recently seen in the financial services industry. The change provides additional insight into the collapse of Long-Term Capital Management about a decade ago, and the more recent collapses of Bear Stearns and Lehman Brothers Holdings. In addition, the experience to date in 401(k) savings plans has shown actual accounts of near retirees to be significantly less than what the participants had been anticipating. The hardship caused by this situation was described in the lead article of the Oct. 19, 2009, issue of Time Magazine.

The change in density function from an independent lognormal function to a conditional lognormal function causes the best estimate rate of return to change from an arithmetic mean of historical returns to the lower geometric mean of historical returns. For large company stocks that have been the focus of the illustration used for this paper, the expected rate of return is lowered from about 12 percent to about 10 percent. For data that is more volatile, the degree of change will be even larger. Given that employee participants have been led to believe they would receive the higher arithmetic mean returns, it is not surprising they are disappointed with the lower actual geometric mean results.

Table 1 below shows the complete impact this change from an arithmetic mean-based forecasting philosophy to a geometric mean-based forecasting philosophy would have on the information provided to an average investor. The same impact affects any risk assessment models that have an investment-based component. In all cases, the anticipated wealth decreases significantly from former expectations based on the traditional lognormal approach.

**TABLE 1**  
**A Comparison of Projected Large Company Stock Returns**

<b>Percentile</b>	<b>Geometric Annual Return</b>	<b>Arithmetic Annual Return</b>
90%	34.23%	38.64%
75	21.57	24.53
60	13.49	15.59
50	8.89	10.52
40	4.48	5.67
25	(2.46)	(1.92)
10	(11.66)	(11.90)
Expected Value	10.36%	12.26%



## Summary

Although the current approach of treating historical investment returns as independent statistical events has been in place for about a half a century, the results from using these models has not been great. The actual experience in the past two years in the financial services industry bears witness to these problems. As shown above, the conversion to an approach based on conditional probabilities not only makes sense from a purely mathematical perspective, it achieves a better match with historical data and helps explain some of the problems seen in the financial services industry.

While this paper has focused primarily on investment forecasting, the basic conditional approach has a huge impact on traditional stock option pricing and affects any risk assessment model where the basic nature of the data is conditional rather than independent. Any risk assessment professional needs to be aware of conditional probability density functions and the role they can play in making sure risk-assessment models are as accurate as they can be.

## Appendix 1

For completeness, the conditional probability density function used in this paper is shown below. The formula assumes that the underlying investment return data is lognormally distributed but that the single observed long-term result lies at the mean of the probability space of long-term results. The periodic shorter term observations are then anticipated to be distributed in accordance with the conditional lognormal density function shown below. The parameters mu and sigma are the classic lognormal distribution parameters; n is the number of periods included in the single long-term result; x is the value of the return expressed as a return relative.

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