Flaw of averages

Statistical intuition is an essential skill for risk managers but, argues David Rowe, don't assume that other business professionals share such intuition, and beware of overconfidence in our own

Most risk managers have considerable statistical training and experience. This merely reflects a normal prior-selection process, since such skills are important prerequisites for filling the role effectively. Such skills are generally accompanied by reasonably sound intuition concerning the inevitable uncertainty that surrounds any point estimate of an unknown value.

Generally, we know the key percentiles of the normal distribution by heart and we also know then to beware of leptokurtosis, or "fat tails". Beyond that, we generally recognise the pervasive importance of diversification in mitigating volatility and reducing the likelihood of outliers.

But all of this can lead to a dangerous misconception – namely that everyone else views the world through similar eyes. A small but growing body of evidence indicates a surprisingly low level of statistical intuition among non-technical business professionals.

David Rowe is president of the Infinity business unit at SunGard Trading & Risk Systems
e-mail: david.rowe@risk.sungard.com

Spinners and averages

For several years, Sam Savage (who originated the term The Flaw of Averages), of Stanford University's department of management science and engineering, has been conducting seminars designed to develop practical statistical insights (see www.analycorp.com, and www.stanford.edu~savage). His work clearly indicates that knowledge of statistical theory does not necessarily imply sound statistical intuition. He regularly asks his audiences to describe the distribution of the results (or the probability density function, although he would never describe it as such) for two different random processes.

The first is a simple spinner that twirls around a circle marked from zero to one. The result of any given spin is the number to which the head of the arrow points when it comes to rest. After those answers are recorded, he shows the correct distribution using Monte Carlo simulation in Excel. The answer, of course, abstracting for some small sample noise, is a uniform distribution with equal likelihood at all points from zero to one. He then asks participants to describe the distribution that occurs when the result is determined as the simple average of two such spins. The precise result is a triangular distribution, but he is willing to accept any shape that goes up in the middle and falls in both tails.

In almost all cases, the audience's answers are one of these two alternatives. But surprisingly, only about 25% get both queries correct. About 25% think both distributions are uniform, failing to see the implications of this classic illustration of diversification. Another 25% say both distributions are higher in the middle, apparently thinking some kind of normal shape is the most likely answer to any statistical question. More remarkable still, about 25% think the first distribution is higher in the middle but, when told it is uniform, decide the second distribution is also uniform.

It is no wonder then that more subtle issues prove even more elusive. For example, say a project depends on successful completion of seven tasks. To keep it very simple, assume there are no forward and backward dependencies, but that all tasks can be performed simultaneously. Nevertheless, all tasks must be accomplished for the project to be complete.

If the expected time to complete each task is one month with a standard deviation of two weeks, it is subtly tempting to assume that the expected time to complete the project is also one month. Here we see the "flaw of averages" at work. Stated simply, the average of a function of random variables is not equal to the function of the averages of those same variables unless the function is linear (and such functions are less often linear than we think). In this case, the time to complete the project is the maximum realisation of seven random durations for each task, which is not a linear function.

In many cases, undesirable results (such as cost or risk exposure) occur on both sides of the mean value for uncertain variables. A common example is the potential credit exposure for a pair of matched interest rate swaps. If rates follow the implied forward path through the life of the swap, total exposure will be small. If, however, rates deviate significantly from that path in either direction, total exposure will be large. Quite obviously, exposure on the average path is not the average exposure. This is, however, a mistake I have seen made frequently over the years.

Even more subtle cases can arise when dealing with results that are linear but where the volatility of the result matters. Return on investment is an obvious case. Most business people understand the importance of diversification in their personal investment portfolios. When it comes to making decisions about funding alternative investment projects in their business, however, this insight often disappears.

The standard approach is to require return analysis based on mean expectations for cost and revenue, then fund those projects in the order of highest to lowest expected return. This can easily lead to an excessive concentration of projects with common economic drivers and high correlation of returns. Consideration of likely correlations would often indicate that choosing some lower return projects with greater statistical independence would reduce volatility of return with only a modestly lower level of expected return.

Distributions

A common reaction to volatility-based analysis is that there is no data to estimate the required distributions. I believe this situation is little different from the basis for expected returns. It is always necessary to apply reasoned judgment concerning market size, potential penetration and estimated production costs. Many factors enter these calculations, and a significant portion of them will relate to external factors beyond the planner's immediate control.

Rarely is there adequate information to ground the projections in a formal statistical estimation process. Rather, some reasonable judgment must be applied and defended at each step until the final conclusion. The same is true when specifying the distributions of unknown variables and the relationships among them. Obviously it is good to experiment with a range of values for such things as correlations among random variables. Even reasonable point estimates of such correlations, determined on a purely judgmental basis, combined with corresponding Monte Carlo analysis, will provide a huge advance over simple point estimates of the means of the variables.

Remember, even if the point estimates of the means are exactly correct, the resulting estimate of net income is only correct if all relationships are linear, and that is rarely the case.

The bottom line of all of this is that volatility and uncertainty are all-pervasive phenomena. Risk analysts need to be constantly on guard against the misleading implications of analysis based only on estimates of the averages of unknown variables. In the end, Monte Carlo analysis based on purely hypothetical distributions will at least frame the problem correctly, and highlight dangerously misleading implications resulting from the flaw of averages. ■