Assessing VaR Accuracy

Kevin Dowd

Kevin Dowd
is a professor of economics at the University of Sheffield in the U.K.

If we are to make practical use of VaR estimates and bet real money on them, we would like to know whether they are precise enough to rely on. But how can we assess the accuracy of our VaR estimates?

THEORY

The answer depends, in part, on the approach we use to estimate the VaR itself.[1] Perhaps the most popular approach is parametric — we specify the distribution of profit/loss (or returns) and use this to infer the VaR. Of the distributions we can choose, the most widely used is the normal distribution.[2]

Yet we know surprisingly little even about the distribution of normal VaR estimates, let alone about the distribution of nonnormal ones. Going back to first principles, we know that normal VaR is given by:

$$\text{VaR} = -\alpha\sigma - \mu \tag{1}$$

where α is a parameter reflecting the confidence level on which our VaR is predicated (e.g., so $\alpha = -1.645$ if our VaR is predicated on the 95% confidence level, and so on), μ is the mean profit/loss, and σ is the standard deviation of profit/loss. Our estimate of normal VaR is then obtained by replacing μ and σ — which are unknown — with their estimates $\bar{x}$ and s:

$$\text{VaR}^e = -\alpha s - \bar{x} \tag{2}$$

In the special case where μ is taken to be some particular deterministic value (e.g., zero), we can invoke statistical theory to construct a confidence interval for s, and then derive the VaR confidence interval by multiplying our s-confidence interval by $-\alpha$ and subtracting a constant μ, which is assumed to be known.[3]

However, μ is seldom known, if ever, so our VaR estimate will usually depend on two stochastic variables (i.e., $\bar{x}$ and s) rather than one. Unfortunately, in their classic statistical reference, Kendall and Stuart [1973, p. 132] describe the construction of analytic confidence intervals for this more general case as involving “very considerable difficulty” (i.e., nearly impossible).

If we cannot derive analytic expressions for VaR confidence intervals, we have little alternative but to resort to simulation. Fortunately, this is very easy to carry out. Our VaR is the sum of two unknown constant terms (i.e., σ times $-\alpha$, and $-\mu$), and we construct a confidence distribution for VaR by simulating each of these terms using the information we have. This information consists of our estimates s and $\bar{x}$, and our knowledge of their distributions: If we have a sample size of n, statistical theory tells us that:

$$(n-1)s^2/\sigma^2 \sim \chi^2_{n-1}$$

and

$$\bar{x} \sim N(\mu, \sigma^2/n) \tag{3}$$

After a little rearranging, we can therefore treat σ^2, σ and μ "as if" they were distributed as:

$$\sigma^2 \sim (n-1)s^2 / \chi^2_{n-1}$$

$$\sigma \sim \sqrt{(n-1)s^2 / \chi^2_{n-1}}$$

and

$$\mu \sim N[\bar{x}, (n-1)s^2 / (n\chi^2_{n-1})] \tag{4}$$

Substituting into (1), we then obtain our "confidence distribution" for VaR:

VaR confidence distribution =

$$-\alpha\sqrt{(n-1)s^2 / \chi^2_{n-1}} - N[\bar{x}, (n-1)s^2 / (n\chi^2_{n-1})] \tag{5}$$

We can now simulate this distribution using the appropriate software and read off the confidence intervals from the quantiles of the simulated distribution.

SIMULATION FINDINGS

For the sake of illustration, I choose a 95% confidence interval, so the lower bound of the confidence interval is the 2.5% quantile of the simulated distribution and the upper bound is the 97.5% quantile. The actual simulation exercise consists of 10,000 simulations of the Latin hypercube routine in @Risk. (Latin hypercube routines are more accurate than conventional Monte Carlo simulation (Palisade [1997, pp. 257-258]).)

This exercise produced a series of histograms that simulate Equation (5), and we are particularly interested in the mean of this histogram and its 2.5% and 97.5% quantiles. The mean can be regarded as an estimate of the unknown "true" VaR, and the two quantiles can be regarded as estimates of the lower and upper bounds of the 95% confidence interval. Simulations were run for a range of (annualized) values for sample mean and sample standard deviation, with the former ranging from 0.1 to 0.3 and the latter from 0.1 to 0.4. These ranges are plausible empirically, and should give some indication of the sensitivity of results to sample values. The simulations are also carried out using sample size (n) values ranging from 10 (i.e., extremely low) to 10,000 (extremely high). This range enables us to gauge the effect of sample size on our results; however, in practice, we would usually deal with n values in the smaller range from perhaps 100 to 1,000.

Finally, simulations were carried out with VaRs predicated on the 95% and 99% confidence levels, and on one-day and ten-day holding periods.[4]

These simulation exercises give us results for the mean VaR, and for the lower and upper bounds of the VaR confidence interval:

- The means of the simulated distributions suggest that the conventional VaR estimate (1) suffers from a slight downward bias, but this bias gets smaller as the sample size increases. Thus, when n = 100, the bias is almost always less than 1% of the "true" VaR; when n = 500, the bias is around 0.2% of true VaR, and when n = 1,000, the bias is around 0.1% of VaR. For reasonable sample sizes, we can therefore regard the standard estimator (2) as having a negligible bias.
- The lower boundary for the confidence interval primarily depends on the sample size n, and gradually moves toward 1 as n gets bigger. To give a typical example, if $\bar{x} = 0.2$ and s = 0.2, the lower confidence level for the VaR (at the 95% confidence level and the one-day holding period) is 0.829 when n = 100, 0.917 when n = 500, and 0.943 when n = 1,000. Other things (e.g., n) being equal, the lower bounds for the confidence level usually move further away from the mean VaR as $\bar{x}$ gets bigger, but do not seem to depend in any noticeable way on s. In addition, the lower bound usually moves further away from the mean VaR as the holding period rises. However, these effects are fairly small and, even where most pronounced, are still small relative to the effects of sample size.
- The results for the upper bound of the confidence interval broadly reflect those for the lower bound: The upper bound depends mainly on n, and moves toward the mean VaR as n rises and, other things being equal, the upper bound usually moves away from the mean as $\bar{x}$ and the holding period each rise. However, these latter effects are again still small relative to the effects of sample size.

Thus, the width of the confidence interval depends mostly on the sample size, n, and gets smaller as n gets larger. In the earlier example with $\overline{x}$ and s both equal to 0.2, the confidence interval for the VaR at the 95% confidence level and one-day holding period is [0.765, 1.329] if n = 50, [0.829, 1.218] if n = 100, [0.917, 1.09] if n = 500, and [0.943, 1.061] if n = 1,000.[5]

Put another way, a sample size of 50 produces a confidence interval that is roughly our VaR estimate plus or minus 30%; a sample size of 100 produces a confidence interval roughly equal to our estimated VaR plus or minus 20%; n = 500 produces a confidence interval equal to our VaR estimate plus or minus 9%; and n = 1,000 produces a confidence interval equal to our estimated VaR plus or minus 6%. The degree of accuracy of our estimate of VaR depends critically on the size of our sample.

CONCLUSIONS

It is very easy to use simulation methods to construct confidence intervals for normal VaRs, and the approach suggested here is particularly useful when closed-form solutions for these confidence intervals are lacking (i.e., when we wish to take account of random estimates of both the mean and standard deviation of profit/loss or returns). The simulation results suggest that normal estimates of VaR have fairly wide confidence intervals, but also suggest that the key factor behind the width of these confidence intervals is sample size. The clear implication is that practitioners need a lot of data if they are to get tight confidence intervals around their VaR estimates.

This approach is straightforward to implement, so those who want more specific estimates of confidence intervals can easily generate them for themselves. Moreover, and more importantly, the same methodology can also be adapted to derive confidence intervals for non-normal VaRs as well.

ENDNOTES

The author thanks Dave Chappell of Sheffield University for helpful advice.

[1]When VaR is estimated by historical simulation, we should resort to historical kernel methods to derive confidence intervals (Butler and Schachter [1996]); when VaR is estimated by Monte Carlo or related methods, we can derive estimates of confidence levels directly from our simulation results, as the level of precision of estimates can be inferred from the number of simulations and (possibly) other features of the simulation setup (see, e.g., Pritsker [1996, pp. 45-47]) or Dowd [1998, Chapter 5]).

[2]It is of course also the most widely criticized, but I do not wish to go into those criticisms here. Instead, I wish to outline an approach that enables us to derive confidence intervals for normal VaR, and which is extendible to other, non-normal, VaRs as well.

[3]For details, see Chappell and Dowd [1999, pp. 1-2].

[4]More details of the simulation results and the software used to generate them will be made available from the author's website, www.shef.ac.uk/~var.

[5]For low values of n, the confidence intervals also show some asymmetry, but this asymmetry disappears as n gets bigger.

REFERENCES

Butler, J.S., and B. Schachter. "Improving Value-at-Risk Estimates by Combining Kernal Estimation with Historical Simulation." Vanderbilt University and Office of the Comptroller of the Currency, Washington, 1996.

Chappell, D., and K. Dowd. "Confidence Intervals for VaR." *Financial Engineering News*, March 1999, pp. 1-2.

Dowd, K. *Beyond Value at Risk: The New Science of Risk Management*. Chichester and New York: John Wiley and Sons, 1999.

Kendall, M.G., and A. Stuart. *The Advanced Theory of Statistics. Volume 2: Inference and Relationship*. Third edition. London: Griffin, 1973.

Palisade Corporation. @Risk: Advanced Risk Analysis for Spreadsheets. Newfield, NY: Palisade, 1996.

Pritsker, M. "Evaluating Value at Risk Methodologies: Accuracy versus Computational Time." Board of Governors of the Federal Reserve System, Washington, 1996.