Value-at-Risk and Confidence-Interval Returns Methodology

Morningstar Methodology Paper
January 1, 2012
Methodology

Introduction
For an investor, risk is not merely the volatility of returns, but the possibility of losing money. Value-at-Risk, also known as VaR, measures the potential loss in value of an investment over a specific time period for a given confidence interval. It tries to answer the question: What is the most I can expect to lose in dollar amount or percentage return in this investment over a specific period with a certain level of confidence (for example, over the next month with 95% confidence)? Looking at it in another way: There is a 5% chance (that is, 100% minus 95%) that I may lose at least this much in any given month.

Value-at-Risk is often a misused statistic.

First, VaR is often estimated solely from the past performance of a fund. As with any estimate that is based on history, such VaR estimates rely on the assumption that the future will mirror the past. This is highly unlikely. The fund might have changed. The markets might behave differently. The VaR might not be generated over a long enough period of history to give a representative estimate. There are many reasons why a VaR estimate might be an unreliable guide to the future.

Second, VaR represents just a single point of a probability distribution. The name of Value-at-Risk implies that the VaR statistic indicates a worst-case scenario, below which the fund cannot fall, but that is not so. A fund could perform far worse than its VaR estimate.

The VaR statistic is useful for comparing the relative risks of funds, assuming that the investor is cognizant of the difficulties of using estimates of the past to forecast the future. Morningstar provides the VaR statistic in that spirit — as a comparative tool for gauging the relative risks of funds. Morningstar does not counsel that VaR be taken literally as the true amount of a fund’s "value at risk."

Following are the methods available at Morningstar for measuring VaR:
- Historical
- Log-Normal
- Log-Stable

These methods are described in detail in the following sections, and all of them use historical returns as a starting point. Users can specify the relevant historical period, the frequency of raw returns, and the choice of currency conversion.
Methodology (continued)

While the VaR measure focuses on an investment's potential loss for a given level of confidence, the same concept of confidence interval has other useful applications in illustrating investment risk. The confidence interval can be expressed as a range of potential upside and downside returns. It tries to answer the question: What is the most and the least I can expect to earn in this investment over a specific holding period with a certain level of confidence? For example, over the next year I can expect the return to fall between -37% and +50% with 95% confidence. This methodology is described in the Log-Normal Method section below.

Historical Method

The historical method of VaR calculation sorts actual historical returns in order from worst to best and identifies the return at the specified confidence value cutoff, for example, 5th percentile for a confidence interval of 95%. When the confidence value cutoff is not a multiple of \( 1/(n - 1) \), where \( n \) represents the total number of returns in the historical period, the VaR is interpolated from the return with the nearest percentile rank above the cutoff and the return with the nearest percentile rank below the cutoff.

Log-Normal Method

The normal distribution, also known as the Gaussian distribution, refers to a bell-shaped (symmetrical) distribution curve that is characterized by two parameters—mean and standard deviation. Mean is the probability-weighted arithmetic average of all possible returns. Standard deviation is the square root of variance, and variance is the probability-weighted average of the square of difference between all possible returns and the mean.

The prefix "log" means that the natural logarithmic form of the return relative, \( \ln(1 + R) \), is normally distributed. The lognormal distribution is asymmetrical, skewing to the right. Because the logarithm of 0 is \( -\infty \), the lowest return possible is -100%, which reflects the fact that an unleveraged investment cannot lose more than 100%.
To use the lognormal model, Morningstar first estimates the arithmetic mean and standard deviation from historical returns as follows:

\[ \mu = \frac{1}{N} \sum_{i=1}^{N} R_i \]

\[ \sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (R_i - \mu)^2} \]

Where:
- \( \mu \) = Arithmetic mean return
- \( R_i \) = Return in period \( i \)
- \( N \) = Total number of periods in the history
- \( \sigma \) = Standard deviation of return

The mean and standard deviation of the natural log-relative return—the two parameters underlying the lognormal distribution—can be derived from the arithmetic mean return (\( \mu \)) and standard deviation of return (\( \sigma \)) as follows:

\[ s = \sqrt{\ln \left( 1 + \left( \frac{\sigma}{1+\mu} \right)^2 \right)} \]

\[ m = \ln(1+\mu) - \left( \frac{s^2}{2} \right) \]

Where:
- \( s \) = Standard deviation of log-relative return, \( \ln(1+R) \)
- \( m \) = Mean of log-relative return, \( \ln(1+R) \)
Methodology (continued)

Following are the formulas for the expected return at a given percentile and VaR, based on the lognormal parameters:

\[ ER = e^{(zn + zn \cdot \sqrt{n})} - 1 \]  \[ \text{VaR} = \vert ER \vert \]

Where:

- \( ER \) = Expected return at a given percentile
- \( \text{VaR} \) = Value-at-Risk
- \( z \) = z-score of the specified percentile
- \( n \) = Number of periods in the time horizon

The VaR is the absolute number of the expected return at the percentile in question because by definition VaR is a loss, so the negative sign is not necessary. The z-score of the percentile is that percentile expressed as the number of standard deviations that is above or below the mean of a distribution. For example, the z-score for the 5th percentile is -1.645 because the 5th percentile is 1.645 standard deviations below the mean. When calculating the VaR, which only focused on the left tail (largest losses) of the return distribution, it is the 5th percentile that represents the VaR with a 95% confidence interval because 100% minus 95% is 5%. The z-score for the 1st percentile or the 99% confidence interval for VaR calculation is -2.326. Z-scores can be obtained from a table of cumulative values of the standard normal distribution.

When the confidence interval is expressed as a range of returns that illustrates both the downside and upside potentials, the 2½th percentile and 97½th percentile are the lower and upper bound of a 95% confidence interval because the focus is on both tails of a return distribution. The z-scores for the 2½th and 97½th percentiles are -1.96 and +1.96, correspondingly.


Log-Stable Method

The log-stable distribution is an alternative to and generalization of the lognormal distribution commonly used to model investment return. It was first proposed as a model for returns by the mathematician Benoît Mandelbrot in the early 1960s and implemented by Eugene Fama for his Ph.D. dissertation, which he wrote under Mandelbrot’s direction. (See Mandelbrot and Hudson 2004). When returns follow the log-stable distribution, the log-return relatives, \( \ln(1 + R) \), follow the Stable Pareto distribution.

The main feature of the log-stable model is that it is a fat-tailed distribution. This type of distribution predicts that returns are subject to extreme values (both positive and negative) with significant levels of probability, in contrast to a thin-tailed distribution such as the log-normal model, which assigns trivial probabilities to such events. Kaplan (2012, chapter 18 and 19) presents evidence that the probabilities predicted by the log-stable model are quite close to the frequencies in which extreme returns have occurred historically in monthly equity indexes.

Letting \( R_t \) denote period \( t \) return on the investment in question, \( r_t \) follows a stable distribution and is defined as

\[
\exp(r_t) = \exp(R_t) = \exp(\ln(1 + R_t)) = \exp(\ln(1 + R_t))
\]

Following are the four parameters that define a stable distribution:

- \( \alpha \) determines how fat the tails are. \( \alpha \) must be positive and cannot exceed 2. The normal distribution is a special case of the stable distribution where \( \alpha = 2 \), and the variance is finite in this situation. Otherwise, it is a fat-tailed distribution with infinite variance. When \( \alpha \leq 1 \), the tails are so fat that the mean of the distribution cannot be defined. Therefore, for the purpose of modeling asset returns, Morningstar assumes that \( 1 < \alpha \leq 2 \).

- \( \beta \) determines the skewness of the distribution, except when \( \alpha = 2 \) (that is, normal distribution). \( \beta \) must be between \(-1\) and \(+1\) inclusively. When \( \beta < 0 \), the distribution is left skewed; when \( \beta > 0 \), the distribution is right skewed; and when \( \beta = 0 \) (or \( \alpha = 2 \)), the distribution is symmetric.
Methodology (continued)

- $\gamma$ is the scale parameter, a measure of dispersion. The most commonly used measure of dispersion is standard deviation, which is the square root of variance; however, $\gamma$ is not equivalent to standard deviation. The only situation where it can be defined with standard deviation or variance is when the distribution is normal, in other words, when $\alpha = 2$. In such cases, $\gamma^2$ is half the variance of the random variable in question. In all other cases, when $\alpha < 2$, $\gamma$ cannot be interpreted as standard deviation because variance itself is undefined in this situation.

- $\delta$ is the location parameter. The exact meaning of $\delta$ depends on the parameterization scheme being used. In the parameterization scheme that Morningstar uses, $\delta$ is expected value of the stable random variable in question if $1 < \alpha \leq 2$. (See Nolan [2009a] for details.)

These parameters can be estimated from an historical time series of returns in logarithmic form $(r_t)$ using a number of statistical techniques. Nolan (2009b) provided a list of methods. Morningstar uses the maximum likelihood method, which is the method most commonly recommended by statisticians.

The formula for VaR in the log-stable model is:

$$[8] \quad \text{VaR} = 1 - e^{\frac{\delta n + z(\alpha, \beta) \gamma n^{m}}{n}}$$

Where:

- $\text{VaR} = \text{Value-at-Risk}$
- $z(\alpha, \beta) = z$-score of the specified percentile of a stable distribution with $\gamma = 1$ and $\delta = 0$
- $n = \text{Number of periods in the time horizon}$

Note that when $\alpha = 2$, equation [6] is equation [8] with $\delta = m$ and $\gamma = s/\sqrt{2}$.

Morningstar computes values for $z(\alpha, \beta)$ using routines described in Nolan (2009b). Table 1 shows values of $z(\alpha, \beta)$ at the 5th percentile for selected values of $\alpha$ and $\beta$. 
Table 1: Values of $z(\alpha, \beta)$ for Selected Values of $\alpha$ and $\beta$ at the 5th Percentile

<table>
<thead>
<tr>
<th>$\beta\alpha$</th>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
<th>1.5</th>
<th>1.6</th>
<th>1.7</th>
<th>1.8</th>
<th>1.9</th>
<th>2.0*</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.0</td>
<td>7.6463</td>
<td>4.5017</td>
<td>3.4796</td>
<td>2.9888</td>
<td>2.7117</td>
<td>2.5428</td>
<td>2.4371</td>
<td>2.3727</td>
<td>2.3378</td>
<td>2.3262</td>
</tr>
<tr>
<td>-0.9</td>
<td>7.1240</td>
<td>4.3012</td>
<td>3.3856</td>
<td>2.9458</td>
<td>2.6962</td>
<td>2.5421</td>
<td>2.4433</td>
<td>2.3805</td>
<td>2.3433</td>
<td>2.3262</td>
</tr>
<tr>
<td>-0.8</td>
<td>6.6551</td>
<td>4.1389</td>
<td>3.2183</td>
<td>2.8206</td>
<td>2.5919</td>
<td>2.4578</td>
<td>2.3529</td>
<td>2.2897</td>
<td>2.2490</td>
<td>2.2362</td>
</tr>
<tr>
<td>-0.7</td>
<td>6.2812</td>
<td>4.0340</td>
<td>3.2860</td>
<td>2.8986</td>
<td>2.6988</td>
<td>2.5603</td>
<td>2.4557</td>
<td>2.3800</td>
<td>2.3551</td>
<td>2.3262</td>
</tr>
<tr>
<td>-0.6</td>
<td>6.0145</td>
<td>3.9960</td>
<td>3.2934</td>
<td>2.9585</td>
<td>2.7208</td>
<td>2.5797</td>
<td>2.4817</td>
<td>2.4117</td>
<td>2.3613</td>
<td>2.3262</td>
</tr>
<tr>
<td>-0.5</td>
<td>5.8142</td>
<td>4.0106</td>
<td>3.3366</td>
<td>2.9763</td>
<td>2.7542</td>
<td>2.6056</td>
<td>2.5009</td>
<td>2.4245</td>
<td>2.3679</td>
<td>2.3262</td>
</tr>
<tr>
<td>-0.4</td>
<td>5.6503</td>
<td>4.0562</td>
<td>3.4055</td>
<td>3.0354</td>
<td>2.7988</td>
<td>2.6378</td>
<td>2.5230</td>
<td>2.4384</td>
<td>2.3747</td>
<td>2.3262</td>
</tr>
<tr>
<td>-0.3</td>
<td>5.5052</td>
<td>4.1204</td>
<td>3.4909</td>
<td>3.1080</td>
<td>2.8528</td>
<td>2.6755</td>
<td>2.5480</td>
<td>2.4535</td>
<td>2.3817</td>
<td>2.3262</td>
</tr>
<tr>
<td>-0.2</td>
<td>5.3838</td>
<td>4.1962</td>
<td>3.5864</td>
<td>3.1988</td>
<td>2.9142</td>
<td>2.7160</td>
<td>2.5755</td>
<td>2.4697</td>
<td>2.3890</td>
<td>2.3262</td>
</tr>
<tr>
<td>-0.1</td>
<td>5.2938</td>
<td>4.2797</td>
<td>3.6865</td>
<td>3.2778</td>
<td>2.9811</td>
<td>2.7645</td>
<td>2.6054</td>
<td>2.4868</td>
<td>2.3965</td>
<td>2.3262</td>
</tr>
<tr>
<td>0.0</td>
<td>5.1646</td>
<td>4.3687</td>
<td>3.7947</td>
<td>3.3699</td>
<td>3.0519</td>
<td>2.8143</td>
<td>2.6373</td>
<td>2.5048</td>
<td>2.4043</td>
<td>2.3262</td>
</tr>
<tr>
<td>0.1</td>
<td>5.0664</td>
<td>4.4616</td>
<td>3.9036</td>
<td>3.4645</td>
<td>3.1256</td>
<td>2.8666</td>
<td>2.6710</td>
<td>2.5238</td>
<td>2.4123</td>
<td>2.3262</td>
</tr>
<tr>
<td>0.2</td>
<td>4.9738</td>
<td>4.5573</td>
<td>4.0145</td>
<td>3.5609</td>
<td>3.2011</td>
<td>2.9210</td>
<td>2.7062</td>
<td>2.5436</td>
<td>2.4205</td>
<td>2.3262</td>
</tr>
<tr>
<td>0.3</td>
<td>4.8858</td>
<td>4.6553</td>
<td>4.1265</td>
<td>3.6583</td>
<td>3.2779</td>
<td>2.9768</td>
<td>2.7428</td>
<td>2.5641</td>
<td>2.4288</td>
<td>2.3262</td>
</tr>
<tr>
<td>0.4</td>
<td>4.8018</td>
<td>4.7549</td>
<td>4.2392</td>
<td>3.7562</td>
<td>3.3556</td>
<td>3.0337</td>
<td>2.7904</td>
<td>2.5865</td>
<td>2.4375</td>
<td>2.3262</td>
</tr>
<tr>
<td>0.5</td>
<td>4.7211</td>
<td>4.8558</td>
<td>4.3523</td>
<td>3.8544</td>
<td>3.4337</td>
<td>3.0914</td>
<td>2.8199</td>
<td>2.6071</td>
<td>2.4633</td>
<td>2.3262</td>
</tr>
<tr>
<td>0.6</td>
<td>4.6434</td>
<td>4.9576</td>
<td>4.4657</td>
<td>3.9526</td>
<td>3.5120</td>
<td>3.1496</td>
<td>2.8851</td>
<td>2.6295</td>
<td>2.4653</td>
<td>2.3262</td>
</tr>
<tr>
<td>0.7</td>
<td>4.5662</td>
<td>5.0601</td>
<td>4.5790</td>
<td>4.0507</td>
<td>3.5903</td>
<td>3.2081</td>
<td>2.8897</td>
<td>2.6523</td>
<td>2.4645</td>
<td>2.3262</td>
</tr>
<tr>
<td>0.8</td>
<td>4.4952</td>
<td>5.1632</td>
<td>4.6923</td>
<td>4.1466</td>
<td>3.6685</td>
<td>3.2663</td>
<td>2.9031</td>
<td>2.6756</td>
<td>2.4739</td>
<td>2.3262</td>
</tr>
<tr>
<td>0.9</td>
<td>4.4242</td>
<td>5.2667</td>
<td>4.8053</td>
<td>4.2460</td>
<td>3.7465</td>
<td>3.3257</td>
<td>2.9757</td>
<td>2.6993</td>
<td>2.4835</td>
<td>2.3262</td>
</tr>
<tr>
<td>1.0</td>
<td>4.3550</td>
<td>5.3706</td>
<td>4.9181</td>
<td>4.3431</td>
<td>3.8242</td>
<td>3.3845</td>
<td>3.0195</td>
<td>2.7234</td>
<td>2.4832</td>
<td>2.3262</td>
</tr>
</tbody>
</table>

* $\alpha=2$ is the normal distribution with variance $\Delta^2$; therefore $z_{\alpha\beta} = 1.645\sqrt{2} = 2.3262$.

Source: Kaplan (2012), Table 19A.1
Methodology (continued)

References


