Financial Data Analysis

GARCH: Application to Value-at-Risk

Summer 2014

July 15, 2014
Application to Risk Management: Value–at–Risk (VaR)

• Value–at–Risk (VaR) is a widely employed risk measure to characterize the downside risk of a financial position.\(^1\)

• The (daily) VaR with confidence level \(1 - \alpha\) answers the following question:
  \(\text{What loss is such that it will only be exceeded with probability } \alpha \text{ in the next trading day?}\)

• If VaR is defined in terms of currency units, e.g., in dollars, denoted \(\$\text{VaR}\), then, for a continuous loss distribution, \(\$\text{VaR}\) is implicitly defined by

\[
\Pr(\text{Loss} > \$\text{VaR}) = \alpha. \tag{1}
\]

• With probability \(1 - \alpha\), the Loss will be \textit{smaller} than the \(\$\text{VaR}\).

• \(\alpha\) (the \textit{shortfall probability}) is typically a small number, such as 0.01 (1%).

Application to Risk Management: Value–at–Risk

• Equivalently, we can define (and will use in the sequel) a VaR based on log–returns $r_t$ as

\[
\Pr(-r_t > \text{VaR}) = \alpha
\]

\[
\iff \quad \Pr(r_t < -\text{VaR}) = \alpha.
\]

(2)

• Now the (daily) VaR with confidence level $1 - \alpha$ is defined such that the next day’s log–return is smaller than $-\text{VaR}$ only with probability $\alpha$.

• That is, in statistical terms, the (daily) VaR at level $1 - \alpha$ is just the (negative of the) $\alpha$–quantile of the next day’s return distribution.
The different definitions of VaR can easily be recovered from one another.

The loss at day $t$ is

$$\text{Loss} = -(P_t - P_{t-1}) = -\Delta P_t,$$  \hspace{1cm} (3)

where $P_t$ is the portfolio’s value (price) at time $t$.

The log–return is

$$r_t = \log P_t - \log P_{t-1} = \log(P_t/P_{t-1}).$$  \hspace{1cm} (4)
Application to Risk Management: Value–at–Risk

• Thus

\[
\text{Pr}(-\Delta P_t > \$\text{VaR})
\]

\[
= \text{Pr}\left(\frac{-\Delta P_t}{P_{t-1}} > \frac{\$\text{VaR}}{P_{t-1}}\right)
\]

\[
= \text{Pr}\left(\frac{P_t}{P_{t-1}} < 1 - \frac{\$\text{VaR}}{P_{t-1}}\right)
\]

\[
= \text{Pr}\left(\log\frac{P_t}{P_{t-1}} < \log\left(1 - \frac{\$\text{VaR}}{P_{t-1}}\right)\right)
\]

\[
= \text{Pr}(r_t < \log\left(1 - \frac{\$\text{VaR}}{P_{t-1}}\right)),
\]

i.e.,

\[
\text{VaR} = -\log(1 - \frac{\$\text{VaR}}{P_{t-1}}), \quad \text{and} \quad \$\text{VaR} = P_{t-1}(1 - \exp\{-\text{VaR}\}).
\]
Application to Risk Management: Value–at–Risk

• VaR is just a single quantile of the return distribution and provides no information about the losses that we may actually expect when a shortfall (or hit or violation) occurs, i.e., when $r_t < -\text{VaR}_t$.

• Thus, to get a better understanding of the risk inherent in the position,

“[r]eporting the entire tail of the return distribution corresponds to reporting VaRs for many different coverage rates, say $[\alpha]$ ranging from 0.01% to 2.5% in increments.”

Backtesting VaR measures

• If we assume that $r_t \sim N(\mu, \sigma^2)$, then the VaR at *nominal* level $1 - \alpha$ or *shortfall probability* $\alpha$ is given by

$$-\text{VaR} = \mu + z_\alpha \sigma,$$

where $z_\alpha$ is the $\alpha$–quantile of the standard normal ($N(0, 1)$) distribution.

• However, if our assumption is wrong, then the *actual shortfall probability* (say $p$) of the VaR calculated according to (10) may be considerably different.

• E.g., if the tails are fat, then, for small $\alpha$, we will have $p > \alpha$, and our model tends to *underestimate* the risk in our position (which is undesirable).
Backtesting VaR measures

• E.g., if our risk model is $r_t \sim \mathcal{N}(0, 1)$, then our 0.5% VaR is\(^3\)

$$-\Phi^{-1}(0.005) = -(−2.5758) = 2.5758.$$  \hspace{1cm} (11)

• However, if the true return distribution is a unit–variance $t$ distribution with 3 degrees of freedom, then the actual shortfall probability $p$ for the VaR at nominal level $\alpha = 0.005$ is

$$F_t(−2.5758; \nu = 3) = 0.0105,$$  \hspace{1cm} (12)

twice as large.

• For a given risk model, we will want to test whether $p = \alpha$ (i.e., we want to backtest the model).

\(^3F_t\) is the unit–variance Student’s $t$ cdf.
Backtesting VaR measures

- Consider a sequence of returns $r_t$, $t = 1, \ldots, T$, and an associated series of \textit{ex-ante} Value-at-Risk measures $\text{VaR}_t$ with nominal shortfall probability $\alpha$, as derived from a given risk model.

- A \textit{violation} or \textit{hit} is said to occur at time $t$ if

$$r_t < -\text{VaR}_t.$$ 

- For a $\text{VaR}_t$ sequence with nominal shortfall probability $\alpha$ and derived from a correctly specified risk model, we expect $(100 \cdot \alpha)\%$ of the observed return values to be violations.
Backtesting VaR measures

• To test the models’ suitability for calculating accurate ex-ante VaR measures, define the binary sequence (*hit sequence*)

\[ I_t = \begin{cases} 
1 & \text{if } r_t < -\text{VaR}_t \\
0 & \text{if } r_t \geq -\text{VaR}_t 
\end{cases} \quad t = 1, \ldots, T. \quad (13) \]

• The hit sequence is a sample of size \(T\) from the Bernoulli distribution with parameter \(p\) (the actual shortfall prob., which we hope is \(\alpha\)), with mass function

\[ f(I_t; p) = p^{I_t}(1-p)^{1-I_t}, \quad I_t \in \{0, 1\}. \]

• For a random sample, the likelihood function is, with \(x = \sum_{t=1}^{T} I_t\),

\[ L(p) = \prod_{t=1}^{T} p^{I_t}(1-p)^{1-I_t} = p^{\sum_t I_t}(1-p)^{T-\sum_t I_t} = p^x(1-p)^{T-x}. \]
Backtesting VaR measures

- The log–likelihood is

\[
\log L(p) = x \log p + (T - x) \log(1 - p),
\]

and maximizing (14) with respect to \( p \) produces the MLE,

\[
\hat{p} = x/T = T^{-1} \sum_t I_t.
\]

- The log–likelihood evaluated at \( \hat{p} \) is

\[
\log L(\hat{p}) = x \log(x/T) + (T - x) \log(1 - x/T).
\]
Backtesting VaR measures

• Under the null hypothesis of a correctly specified risk model,

\[ H_0 : p = \alpha, \]  

the log–likelihood is

\[ \log L(\alpha) = x \log \alpha + (T - x) \log(1 - \alpha). \]  

• A likelihood ratio test (LRT) tests whether the improvement of evaluating \( \log L(p) \) at the MLE \( \hat{p} \) rather than \( \alpha \) is so large that we have to reject (17).
Backtesting VaR measures

• The LRT test statistic and its limiting distribution under $H_0$ are

\[
    \text{LRT} = -2(\log L(\alpha) - \log L(\hat{p})) \overset{d}{\to} \chi^2(1). \tag{19}
\]

• The null hypothesis is rejected at the $(100 \cdot \xi)$% level (not to be confused with the VaR level!) if

\[
    \text{LRT} > F^{-1}_{\chi^2(1)}(1 - \xi), \tag{20}
\]

where $F^{-1}_{\chi^2(1)}(1 - \xi)$ is the $(1 - \xi)$–quantile of the $\chi^2(1)$ distribution.
Example: Out–of–Sample VaR measures for the S& P 500

- Daily data from January 2000 to October 2011, 2970 observations (after zeros have been deleted).
- We first estimate models using the first 1000 observations.
- Then use these estimates to calculate VaR measures for the next day \((t = 1001)\), i.e., genuine out–of–sample VaR measures.
- Then update the parameters using the first 1001 observations.\(^4\)
- Calculate \textit{ex–ante} VaR measures for day \(t = 1002\).
- Proceeding this way until we arrive at the end of the sample, we obtain \(T = 1970\) out–of–sample one–step ahead VaR measures for each model and each nominal VaR level, \(\alpha\).

\(^4\)That is, an expanding window of data is used. Alternatively, a \textit{moving data window} may be used (e.g., to account for slowly changing parameters). In the latter case, the e.g. most recent 1000 observations in the sample would be used.
Table 1: GARCH(1,1) Value–at–Risk measures for the S&P 500, reported is $100 \times \hat{p} = 100 \times \frac{1}{T} \sum_{t} I_t$, i.e., the percentage empirical shortfall frequency

<table>
<thead>
<tr>
<th>nominal VaR level $\alpha$:</th>
<th>0.005</th>
<th>0.01</th>
<th>0.025</th>
<th>0.05</th>
<th>0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$#$ violations</td>
<td>25</td>
<td>44</td>
<td>84</td>
<td>118</td>
<td>202</td>
</tr>
<tr>
<td>$100 \times \hat{p}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gaussian GARCH(1,1)</td>
<td>1.269***</td>
<td>2.234***</td>
<td>4.264***</td>
<td>5.990*</td>
<td>10.25</td>
</tr>
<tr>
<td>Student's $t$ GARCH(1,1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$#$ violations</td>
<td>16</td>
<td>32</td>
<td>73</td>
<td>124</td>
<td>215</td>
</tr>
<tr>
<td>$100 \times \hat{p}$</td>
<td>0.812*</td>
<td>1.624**</td>
<td>3.706***</td>
<td>6.294**</td>
<td>10.91</td>
</tr>
</tbody>
</table>

Asterisks *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively, based on the test (19).

- Gaussian GARCH underestimates risk in the tail.
- Student’s $t$–GARCH better but still accuracy of its VaR measures tends to be rejected.
- Skewness may be an issue here (in addition to fat tails).
Conditional Skewness

• Asymmetric versions of the GED and the \( t \) distribution have been proposed.

• Regarding the GED, the skewed exponential power (SEP) distribution of Fernandez, Osiewalski, and Steel (1995) has density

\[
f(z; p, \theta) = \frac{\theta p}{1 + \theta^2 2^{1/p} \Gamma(1/p)} \begin{cases} 
\exp \left\{ -\frac{1}{2} (|z| \theta)^p \right\} & \text{if } z < 0 \\
\exp \left\{ -\frac{1}{2} \left( \frac{z}{\theta} \right)^p \right\} & \text{if } z \geq 0,
\end{cases}
\]

where \( \theta, p > 0 \).

• This distribution nests the normal for \( \theta = 1 \) and \( p = 2 \). For \( \theta < 1 (\theta > 1) \), the density is skewed to the left (right), and is fat–tailed for \( p < 2 \).

• \( \theta < 1 \) stretches the left tail and shrinks the right tail and thus leads to left–skewness (and \( \theta > 1 \) produces right–skewness).
Skewed Exponential Power Distribution with $p = 1.5$

- $\theta = 1$
- $\theta = 0.75$
- $\theta = 0.5$
Conditional Skewness

• Various skewed versions of the Student’s t also exist.

• A t version of (21) is the skewed t distribution proposed by Mittnik and Paolella (2000), which has density

\[
f(z; \nu, p, \theta) = \frac{\theta}{1 + \theta^2 \nu^{1/p} B(\nu, 1/p)} \begin{cases} 
\left(1 + \frac{|z|\theta}{\nu}\right)^{-(\nu+1/p)} & \text{if } z < 0 \\
\left(1 + \frac{z/\theta}{\nu}\right)^{-(\nu+1/p)} & \text{if } z \geq 0,
\end{cases}
\]

where \( \nu, p, \theta > 0 \), and \( B(\cdot, \cdot) \) is the beta function.

• Repeat the Value–at–Risk forecast calculations with the GARCH(1,1) with skewed t innovations (22).
Expanding window estimates of skewness parameter $\theta$ of the skewed t density
Table 2: GARCH(1,1) Value–at–Risk measures for the S&P 500, reported is $100 \times \hat{p} = 100 \times \frac{1}{T} \sum_{t} I_t$, i.e., the percentage empirical shortfall frequency

<table>
<thead>
<tr>
<th>nominal VaR level $\alpha$:</th>
<th>0.005</th>
<th>0.01</th>
<th>0.025</th>
<th>0.05</th>
<th>0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td># violations</td>
<td>25</td>
<td>44</td>
<td>84</td>
<td>118</td>
<td>202</td>
</tr>
<tr>
<td>$100 \times \hat{p}$</td>
<td>1.269***</td>
<td>2.234***</td>
<td>4.264***</td>
<td>5.990*</td>
<td>10.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>nominal VaR level $\alpha$:</th>
<th>0.005</th>
<th>0.01</th>
<th>0.025</th>
<th>0.05</th>
<th>0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td># violations</td>
<td>16</td>
<td>32</td>
<td>73</td>
<td>124</td>
<td>215</td>
</tr>
<tr>
<td>$100 \times \hat{p}$</td>
<td>0.812*</td>
<td>1.624**</td>
<td>3.706***</td>
<td>6.294**</td>
<td>10.91</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>nominal VaR level $\alpha$:</th>
<th>0.005</th>
<th>0.01</th>
<th>0.025</th>
<th>0.05</th>
<th>0.1</th>
</tr>
</thead>
<tbody>
<tr>
<td># violations</td>
<td>13</td>
<td>21</td>
<td>61</td>
<td>107</td>
<td>198</td>
</tr>
<tr>
<td>$100 \times \hat{p}$</td>
<td>0.660</td>
<td>1.066</td>
<td>3.096</td>
<td>5.431</td>
<td>10.05</td>
</tr>
</tbody>
</table>

Asterisks *, **, and *** indicate significance at the 10%, 5% and 1% levels, respectively, based on the test (19).
Conditional Skewness

• It appears that both conditional skewness and kurtosis may be important and can considerably improve conditional predictive densities (which are used to derive risk measures such as VaR).
Ex-ante 0.5% Value-at-Risk measures implied by various GARCH models
Ex-ante 0.5% Value-at-Risk measures implied by various GARCH models

- Normal
- Student t
- Skewed t