Chapter 7

Empirical Results: Application of Peaks Over Threshold Method (POT) of Extreme Value Theory

7.1. Background of the study

Extreme value theory is a branch of statistics dealing with the extreme deviations from the median of probability distributions. The general theory sets out to assess the type of probability distributions generated by processes. The approaches exist today as mentioned below.

Most common at this moment is the tail-fitting approach based on the second theorem in extreme value theory (Theorem II Pickands\textsuperscript{65} (1975), Balkema and de Haan\textsuperscript{66} (1974)). In general this conforms to the first theorem in extreme value theory (Theorem I Fisher and Tippett\textsuperscript{67} (1928), and Gnedenko\textsuperscript{68} (1943)). The difference between the two theorems is due to the nature of the data generation.

For theorem I the data are generated in full range, while in theorem II data is only generated when it surpasses a certain threshold (POT's models or Peak Over Threshold). The POT approach has been developed largely in the insurance business, where only losses (pay outs) above a certain threshold are accessible to the insurance company. Strangely this approach is often applied to theorem I cases which poses problems with the basic model assumptions. Extreme value distributions are the limiting distributions for the minimum or the maximum of a very large collection of random observations from the same arbitrary distribution. Emil Julius Gumbel (1958) showed that for any well-behaved initial distribution (i.e., F(x) is continuous and has an inverse), only a few models are needed, depending on whether you are interested in the maximum or the minimum, and also if the observations are bounded above or below.

We have applied the peaks over threshold (POT) method for our empirical study.

7.2. Analytical framework of the study

We empirically analyze the extreme behaviors of the tail of the return series from Nifty by applying Extreme Value Theory. We have used Generalized Pareto Distribution (GPD) or Peak Over Threshold (POT) method of Extreme Value Theory. An overview of this method is given herein under.

The generalized Pareto distribution (GPD) is defined as:
Where the tail parameter $\tau = 1/\alpha$ (where $\alpha$ is the shape parameter), the location parameter $\mu$ and the scale parameter $\sigma$.

For any large enough threshold $y$, the conditional excess distribution is always GPD. The GPD is presumed to be the best distribution to model any extreme tail and it has the feature to use all the available data. The GPD requires 3 parameters: (i) The tail parameter $\tau = 1/\alpha$ (where $\alpha$ is the shape parameter), (ii): The location parameter $\mu$ and (iii): The scale parameter $\sigma$.

The shape parameter $\xi$ governs the tail behaviour of the distribution, the sub-families defined by $\xi \to 0$, $\xi > 0$ and $\xi < 0$ correspond, respectively, to the Gumbel, Fréchet and Weibull families, whose cumulative distribution functions are reminded below.

Gumbel or type I extreme value distribution

$$F(x; \mu, \sigma) = e^{-e^{-(x-\mu)/\sigma}} \text{ for } x \in \mathbb{R}$$

Fréchet or type II extreme value distribution

$$F(x; \mu, \sigma, \alpha) = \begin{cases} 0 & x \leq \mu \\ e^{-((x-\mu)/\sigma)^{-\alpha}} & x > \mu \end{cases}$$

Weibull or type III extreme value distribution
\[
F(x; \mu, \sigma, \alpha) = \begin{cases} 
  e^{-(x-\mu)/\sigma} & x < \mu \\
  1 & x \geq \mu 
\end{cases}
\]

Where \( \sigma > 0 \), and \( \alpha > 0 \).

For reliability issues the Weibull distribution is used with the variable \( t = \mu - x \), the time, which is strictly positive. Thus the support is positive - in contrast to the use in extreme value theory. It need be aware of an important distinctive feature of the three extreme value distributions: The support is either unlimited, or it has an upper or lower limit. One can link the type I to types II and III the following way: if the cumulative distribution function of some random variable \( X \) is of type II: \( F(x;0,\sigma,\alpha) \), then the cumulative distribution function of \( \ln X \) is of type I, namely \( F(x;\ln \sigma,1/\alpha) \). Similarly, if the cumulative distribution function of \( X \) is of type III: \( F(x;0,\sigma,\alpha) \), the cumulative distribution function of \( \ln(X) \) is of type I: \( F(x;\ln \sigma, -1/\alpha) \).

In probability theory and statistics the Gumbel distribution (named after Emil Julius Gumbel (1891–1966)) is used to find the minimum (or the maximum) of a number of samples of various distributions. For example we would use it to find the maximum level of a river in a particular year if we had the list of maximum values for the past ten years. It is therefore useful in predicting the chance that an extreme situation like earthquake, flood or other natural disaster will occur.

The distribution of the samples could be of the normal or exponential type. The Gumbel distribution, and similar distributions, is
used in theory. In particular, the Gumbel distribution is a special case of the Fisher-Tippett distribution (named after Sir Ronald Aylmer Fisher (1890–1962) and Leonard Henry Caleb Tippett (1902–1985)), also known as the log-Weibull distribution.

In statistics, the Weibull distribution (named after Waloddi Weibull) is a continuous probability distribution with the probability density function

\[
f(x; k, \lambda) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k}
\]

For \( x \geq 0 \) and \( f(x; k, \lambda) = 0 \) for \( x < 0 \), where \( k > 0 \) is the shape parameter and \( \lambda > 0 \) is the scale parameter of the distribution.

The Weibull distribution is often used in the field of data analysis of life insurance due to its flexibility - it can mimic the behavior of other statistical distributions such as the normal and the exponential. If the failure rate decreases over time, then \( k < 1 \). If the failure rate is constant over time, then \( k = 1 \). If the failure rate increases over time, then \( k > 1 \).

When \( k = 3.4 \), then the Weibull distribution appears similar to the normal distribution. When \( k = 1 \), then the Weibull distribution reduces to the exponential distribution.

Given a random variate \( U \) drawn from the uniform distribution in the interval (0, 1), then the variate
\[ X = \lambda \left( -\ln(U) \right)^{1/k} \]

It has a Weibull distribution with parameters \( k \) and \( \lambda \). This follows from the form of the cumulative distribution function.

An interesting point to be noted that generally I am most interested in the case \( \tau > 0 \) (the Frechet distribution) as this corresponds with fat-tails. In the Frechet distribution the shape parameter \( \alpha = 1/\tau \) represents the maximal order of finite moments. For example if \( \alpha > 1 \) - then the mean exist, if \( \alpha > 2 \) - then the variance is finite, if \( \alpha > 3 \) - then the skewness is ill defined and so on. The shape parameter is an intrinsic parameter of the distribution of returns and does not depend on the number of returns \( n \) from which the minimal return is selected. The shape parameter corresponds to the degrees of freedom for the t distribution and to the characteristics exponent of a stable Pareto distribution.

The estimator for the tail is

\[ \tilde{G}_{\tau,\mu,\sigma}(x + y) \]

for \( x > 0 \) and a “high” threshold \( y \) takes the following form:

\[ \tilde{G}_{\tau,\mu,\sigma}(x + y) = \frac{n}{N} \left( 1 + \tau \frac{x - \mu}{\sigma} \right)^{\frac{1}{\tau}} \]

In the Frechet and Gumbel case - that is \( \tau \geq 0 \), the domain restriction is \( x \geq 0 \) - it clearly stresses that the upper tail is estimated.

An estimator for the quantile \( x_p \) follows immediately by inverting the tail estimation formula as above (for \( \tau \geq 0 \)): 
\[ x_p = \mu + \frac{\sigma}{\tau} \left( \frac{N}{n} (1 - p) \right)^{-\tau} - 1 \]

where \( p \) is the probability.

It also be noted that \( X_p \) can be recognized as \( \text{VaR}_p \) if the probability satisfies the following condition:

\[ p > G_{\tau;\mu,\sigma}(y) \]

7.3 Data Sources

The data have been downloaded from daily closing value of Infosys from National Stock Exchange of India (NSE)’s website www.nseindia.com. These are the daily value of Infosys as officially published by NSE at its website. The data range from 1st January 1997 to 29th December 2006.

7.4 Methodology & Derivation of Empirical Results

*Graphical Data Exploration Tools*

Quantile plots (QQ–plots) is used as an important tool among the graphical data exploration techniques which can be used to distinguish visually different distribution functions. Whereas the sample data come from the family of distributions \( F \), the plot will be close to a straight line. If the deviation from the straight line is too strong is concluded that the sample comes from a different distribution. The figure mentioned below depicts the sample quantile plotted against GPD quantiles.
Figure 7.4.1: QQ–plot of sample quantiles against quantiles of GPD distribution.

The picture as above strongly suggests that our data follow a GPD distribution is acceptable as the plotted line tends to a straight line.

Another graphical tool which is helpful for the selection of the threshold \( u \), which defines the exceedances in our data set, is the sample mean excess plot. The figure, mentioned below, shows the sample mean excess plot corresponding to our data.
By above abrupt upward turn of slopes in plot I suggest trying the values $u = 1.43$ and $u = 2.10$ the threshold. The reason is that at these values I observe a change from a horizontal line to a line with a positive slope.
Above figure further strengthens our estimation that the distribution is a GPD as the slope tends to a straight-line.

Maximum Likelihood Estimation

We have used the maximum likelihood method to compute the estimates. Further, According to our interpretation of the sample mean excess plot, we computed the values for \( u = 1.43 \) and \( u = 2.10 \) which maximize the log-likelihood function for the two samples corresponding to a threshold of \( u = 1.43 \) and \( u = 2.10 \), the shape parameter is \( \xi \). We obtained the estimates \( \xi = 0.232 \) for \( u = 1.43 \) and \( \xi = 0.358 \) for \( u = 2.10 \). We observe that the shape parameter varies very little between the two values of \( u \) and I therefore choose the threshold \( u = 1.43 \) which leaves 563 observations in the tail instead of
The GPD fitted to the $\text{Nu} = 563$ exceedances above the threshold $u = 1.43$ is plotted in the figure above. High quantiles may now be directly read in the plot or computed from equations. Choosing $p = 0.01$, we have VaR as 2.43.
We have obtained different values of sigma and xi which are enumerated in the table 7.4.1 below:

**Table 7.4.1: Point estimates and Bounds of Xi and Sigma**

<table>
<thead>
<tr>
<th></th>
<th>Upper Bound</th>
<th>Point Estimate</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Point Estimate</th>
<th>Lower Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>Xi</td>
<td>-0.100</td>
<td>0.232</td>
<td>0.351</td>
<td>0.463</td>
<td>0.512</td>
<td>0.566</td>
</tr>
<tr>
<td>ML Method</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Result of Value at Risk Computation is stated in the table 7.4.2 below.

**Table 7.4.2: Results of EVT Value at Risk**

<table>
<thead>
<tr>
<th></th>
<th>Upper Bound</th>
<th>Point Estimate</th>
<th>Lower Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>VaR</td>
<td>2.336</td>
<td>2.430</td>
<td>2.536</td>
</tr>
</tbody>
</table>

The results, as shown in Table 7.4.2, indicate that with probability 0.01 the tomorrow’s loss will exceed the value 2.43% i.e. VaR by EVT method. These point estimates are completed with 95% confidence intervals with upper and lower bounds. Thus the expected loss will, in 95 out of 100 cases, lie between 2.336 and 2.536%.

We analyze VaR estimated with the POT method as it arrests the movement in the tails. Assuming the normal distribution for the observations, the calculated convention VaR with Basle multiplication factors may usually fall beyond the upper bound for the expected shortfall i.e. 2.336 as made in our calculation above. Technically the POT method provides more accurate information. The reason for the improved performance of the methods based on Extreme Value Theory is that it models the tails movement and evaluates the risk left out in the tails as it is empirically seen by many researchers that most of the cases the financial return data are heavy tailed. As such, in the emerging markets like India, the methods based on Extreme Value Theory is a better performer in modeling extreme behaviors of financial instruments like stock as in our case Infosys.
7.5. Summary

The Results of our study that Value at Risk as per Peaks Over Threshold method of Extreme Value Theory is 2.43. Further, expected loss is found to be within the upper and lower bounds of 2.336 and 2.536. The VaR is 2.43 is evidently shows the extreme movement in the tails along with the bounds, which may be ignored by the conventional VaR estimate, even the pass the regulatory backtesting procedures. Therefore, it signals prudently the probable extreme risk. In fact, the Peak Over Threshold (POT) method is an important model to estimate local measures of risk when financial asset values are likely to follow a nonstationary process or show volatility clustering. Backtesting and comparing the new method to existing ones on real financial recordings show that the proposed method provides a rather realistic model for the extremal behavior of financial processes, and therefore a precise estimation of risk measures. Through the GPD, it provides a way of estimating the tail behaviour of the random variables without knowledge of the true distribution and as such it is a good candidate for Vale at Risk computation.