INVESTMENT
MANAGEMENT
SOCIETY

EXTREME VALUE THEORY: How to predict extreme losses

by Mohammad Pourmohammadi
MSc in Finance Student, Bocconi
Risk Analyst, Minerva IMS

OVERVIEW

Everyone wants to make money from the stock market. People think it is the answer to their prayers, “I want to get rich, while doing nothing!”. Looking at the level of the S&P 500 over the years it seems that some of these prayers have indeed been answered.

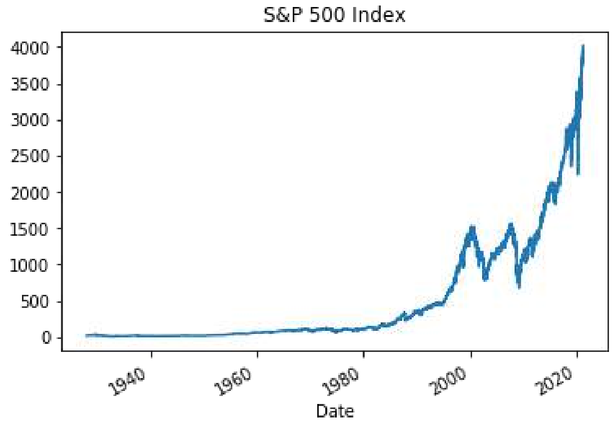

Figure 1 - S&P 500 index 1927 -2021

Any amount we would have invested in the S&P 500 in the 1930’s would have grown by 22,500% percent. However, this is one of quadrillions of possibilities that could have occurred and, as intelligent investors we must ask ourselves what sort of probabilistic distribution could have created such a plot. We can get a sense of this by plotting the histogram of log returns.

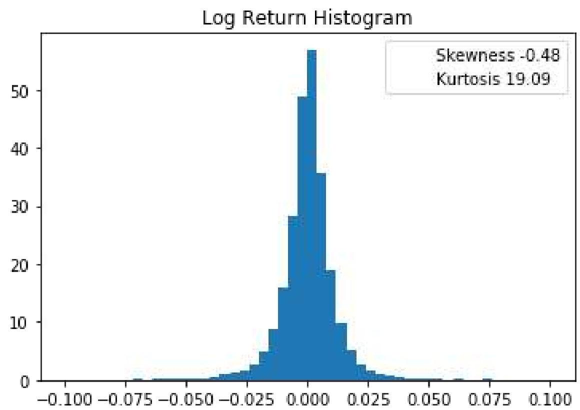

Figure 2 - Historical log-return of S&P 500

Positive kurtosis and negative skewness indicate large negative returns that inflate the left tail, far exceeding that of the normal distribution (the Gaussian distribution has a skewness of zero and kurtosis of 3). In practice, these events are what dictate the survival of the average portfolio manager. To hedge ourselves from these cases, we need distributions that are robust and stable, contingent on the occurrence of such shocks. This leads to a possible solution by Extreme Value theory (EVT). For further explanation on the drawbacks of traditional financial modelling and motivation about EVT refer to (Taleb 2020).

EXTREME VALUE THEORY

EVT is a branch of Statistics that deals with extreme deviations from the median of probability distributions. This theory tries to give a description for heavy tailed distributions like the ones of financial markets. We use two separate, but similar methods to illustrate the power of this theory. First, the block Maxima method which consists in inferring the extreme observations of a certain time-series, using extreme data in each period. For example, in our case we took into consideration the highest loss per two months. Second, we use peaks over threshold. In this method the distribution for the tail is inferred using data over a certain quantity, in our analysis we use data over the 90% quantile of losses. For a formal introduction to both methods refer to (Haan 2006).

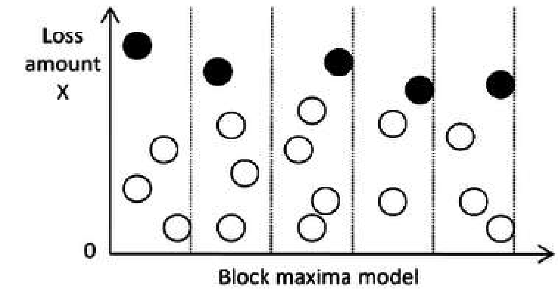

Figure 3 - Block Maxima (Bhattacharyya 2008)

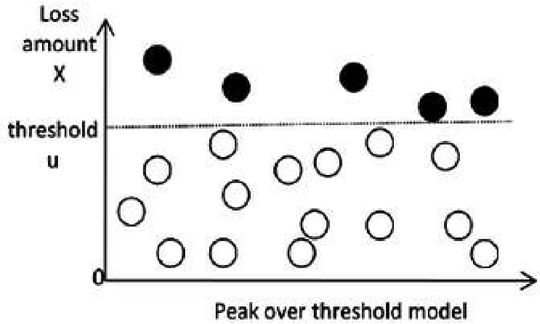

Figure 4 - Peak Over Threshold (Bhattacharyya 2008)

BLOCK MAXIMA

In the first case, we attempt to find the best heavy tailed distribution that fit the maxima, the Generalized Extreme Value Distribution (GEVD); essentially this is a central limit theorem for maxima, and we can prove that, under suitable conditions, the probability distribution function (PDF) of the maxima converge to

$$f(x;c) = \exp\left(-(1-cx)^{\frac{1}{c}}\right)(1-cx)^{\frac{1}{c}-1}$$

if $c > 0$ we have $-\infty \leq x < 1/c$, and if $c < 0$ we have $1/c \leq x < \infty$. Here c is the shape parameter, controlling the tail attribute of the distribution.

We now fit our training set, taking the daily data of negative log returns (representing losses) from the S&P 500 from 1927 up to 2019. For our test set we use data for 1927 up to the current day April 4th, 2021. The fitting is done by an adjusted Maximum Likelihood argument, implemented in the SciPy package of python (more details in the Jupyter notebook in the source code). Here is the result in the two periods:

```
Train Parameters = [-0.2833  0.016   0.0087]
Test Parameters = [-0.2908  0.0161  0.0087]
VaR pre and post are -0.098 and -0.101
```

As can be seen the parameters post and pre 2019 are strikingly similar. This was to be expected, as there is considerable overlap between the datasets. However, what makes these results exceptional is that in the parameter estimation more weight is given to the extreme market events, events that did indeed transpire in March 2020 due to pandemic. The small decrease in the index c (first parameter), is arguably caused by this downturn.

The second and third parameters control the location and scale of the distribution respectively, under the transformation $x := \frac{x-loc}{scale}$.

The value at risk (VaR) 99% is computed using the 99% percentile of the distribution implied by the parameters. Note that this is not equivalent to a VaR implied by the distribution of returns, this is the 99% percentile of the maxima over a 2-month period. Therefore, a much more extreme, measure of risk.

By taking the integral, this PDF implies the following cumulative distribution function (CDF)

$$F(x;c) = \exp\left(-(1-cx)^{\frac{1}{c}}\right)$$

in our case c is negative. As x increases this function will converge to one and the exponent will go to zero. Using the first order Taylor expansion we can estimate this function as

$$F(x;c) \approx 1-(1-cx)^{\frac{1}{c}}$$

This means that the tail of the distribution has a mass that goes to zero at a polynomial rate. In our empirical example $c \approx -0.3$, implying that statistical moments higher than the third (skewness) are not well defined mathematically. This means that their empirical counter parts will not converge. Implying that law of large numbers will require more data (slower convergence in practice), details in (Taleb 2020). Some literature suggests that this c is higher, between -0.5 and -1. This result cannot be rejected by the current analysis due to the sensitivity to both the block duration (e.g., monthly, or yearly) and data (e.g., Close or Adjusted Close) used.

To view the results of our back test, we plot the returns from 2019 together with different measures of VaR corresponding to different distributions.

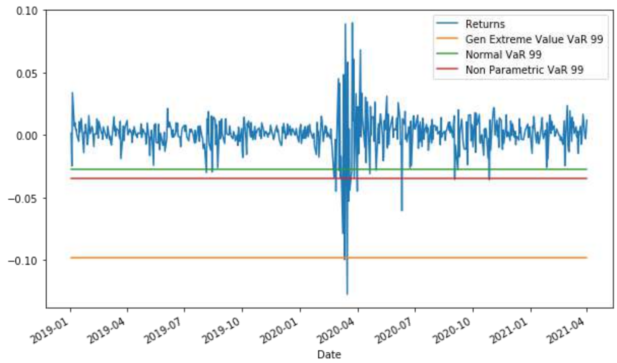

Figure 5 - VaRs estimations

The extreme value at risk only intersects the returns twice. When divided by number of returns in the period gives a number less than 1%. This means that this quantity is not rejected in the back-test as a 1% measure of risk. The same cannot be said for the VaR under the normal and the non-parametric distributions.

```
Percentage of retruns that exceed the Extreme Value VaR 99 equals 0.004
Percentage of retruns that exceed the non-parametric VaR 99 equals 0.019
Percentage of retruns that exceed the Gaussian VaR 99 equals 0.039
```

PEAK OVER THRESHOLD

In our second method, we use peaks over threshold to model the tail of the distribution. The values above a certain threshold are used to estimate a Generalized Pareto distribution that is the best fit to the tail. With the following PDF

$$f(x;c) = (1 + cx)^{-1-1/c}$$

defined for $x \geq 0$ if $c \geq 0$, and $0 \leq x < 1/-c$ if $c < 0$. Taking the integral we reach the CDF given by

$$F(x;c) = 1 - \frac{1}{(1 + cx)^{1/c}}$$

Under positive c this implies a polynomial decline for the CDF identical to the CDF of the Generalized Extreme Value distribution. More generally it can be proved that the class of distributions that have stable tail distributions also have stable distributions for the maxima. The following log-frequency plot shows the polynomial decline of the tail. The intuition behind this plot is that if the tail PDF were to be a polynomial, then the logarithm with respect to returns should be linear.

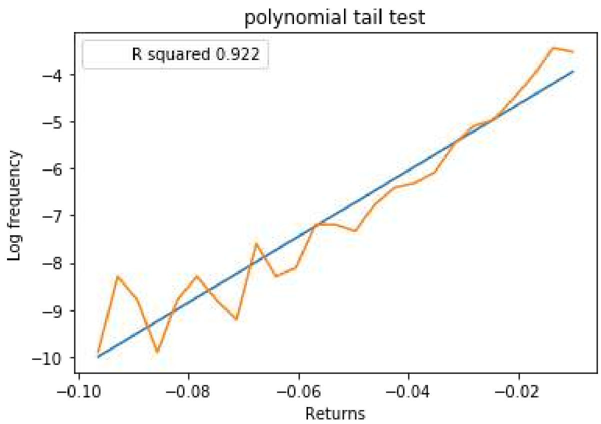

Figure 6 - Proof of polynomial tail

We compare the resulting fit with that of a normal distribution for losses over the 90% quantile, the left tail with an empirical mass of 10%. This is done by first filtering the data for loss values above the 90% quantile and fitting, using the *genpareto* of SciPY in python.

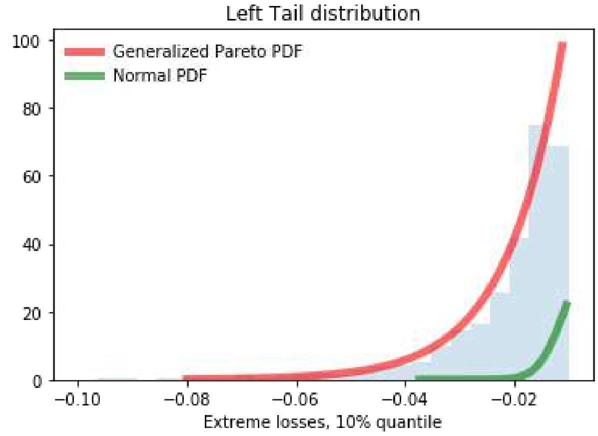

Figure 7 - Left tail distribution of the returns

As can be seen, the Generalized pareto is a perfect fit for the tail. More importantly it is above the histogram for most tail values, making it a more conservative estimate for tail probabilities, probabilities that dictate blow up in finance. The normal distribution is instead ill matched for the extreme events, not coming close even in the last 10% of probabilities.

CONCLUSION

In this article we have seen two basic ways of using extreme value theory with financial data. These models are robust to extreme events. Events which seem to be quite common in the financial data. The main drawback is the filtering required to compute the distributions; this means that we lose a significant portion of the dataset when fitting distributions. Note that increasing the time scale does result in more return data, as venturing into higher frequencies implies incurring more noise in the sample. Overall, these techniques can help us obtaining more realistic probabilities for extreme events and consequently computing the VaRs and other risk measures.

REFERENCES

Source code and dataset:

https://github.com/theonlyonetorise/Extreme-Value-Theory

Bhattacharyya, Malay and Ritolia, Gopal. 2008. "Conditional VaR using EVT - Towards a planned margin scheme." *International Review of Financial Analysis* 382-395.

Haan, Laurens and Ferreira, Ana. 2006. *Extreme Value Theory: An Introduction.* Springer.

Taleb, Nassim Nicholas. 2020. *Statistical Consequences of Fat Tails: Real World Preasymptotics, Epistemology, and Applications.*

DISCLAIMER