An Application of Extreme Value Theory for Measuring Financial Risk in the Uruguayan Pension Fund

Guillermo Magnou

January 2016

Abstract

Traditional methods for financial risk measures adopts normal distributions as a pattern of the financial return behavior. Assessing the probability of rare and extreme events is an important issue in the risk management of financial portfolios. In this paper, we use Peaks Over Threshold (POT) model of Extreme Value Theory (EVT), and General Pareto Distribution (GPD) which can give a more accurate description on tail distribution of financial losses. The EVT and POT techniques provides well established statistical models for the computation of extreme risk measures like the Return Level, Value at Risk and Expected Shortfall. In this paper we apply this technique to a series of daily losses of AFAP SURA over an 18-year period (1997-2015), AFAP SURA is the second largest pension fund in Uruguay with more than 310,000 clients and assets under management over USD 2 billion.

Our major conclusion is that the POT model can be useful for assessing the size of extreme events. VaR approaches based on the assumption of normal distribution are definitely overestimating low percentiles (due to the high variance estimation), and underestimate high percentiles (due to heavy tails). The absence of extreme values in the assumption of normal distribution underestimate the Expected Shortfall estimation for high percentiles. Instead, the extreme value approach on POT model seems coherent with respect to the actual losses observed and is easy to implement.

Keywords. Extreme Value Theory, General Pareto Distribution, Peaks Over Threshold, Risk Measures, Value at Risk, Pension Fund.

1 Introduction

Traditional statistical methods for financial risk measures fits models to all data even if primary focus is on extremes. It is for this reason that it is common to see in literature the normal distribution assumption for financial returns. This assumption provides a good approximation for the average of financial returns (due the central limit theorem) but does not provide a good fit for the tail of the distribution.

The last 20 years have been characterized by significant instabilities in financial markets. This instability generates volatility in investment portfolios in sensitive areas such as social security. The Uruguayan pension funds are exposed to these volatility, especially because they are characterized for long term investment. This has led to numerous critics about the existing risk management systems and motivated the search for more appropriate methodologies for extreme risk measures. The Extreme Value Theory (EVT) is a powerful and fairly robust framework to study the tail behavior of a distribution. EVT became important in the 20s with problems primarily related to hydrology and led to the first fundamental theorem of Fisher-Tippet (1928), then Gnedenko (1948) characterized the asymptotic distribution of the maximum observed. Another point of view came in the 70s with the second fundamental theorem of Extreme Value Theory when Pickands (1975) and Balkema-de Haan (1974), characterized the asymptotic tail distribution as a Generalized Pareto Distribution (GPD) family.

1 The opinions expressed in this document are the sole responsibility of the author and do not compromise nor represent the position of the AFAP SURA. We thank Gonzalo Falcone and Joaquin Idoyaga for allowing the use of the data.
2 Master Mathematical Engineering, University of the Republic Uruguay.
3 Department of Legal, Compliance and Risk AFAP SURA. Mail: gmagnou@afapsura.com.uy
Extreme Value Theory provides well established statistical models for the computation of extreme risk measures like the Value at Risk, Expected Shortfall and Return Level. In this paper we utilize Peaks Over Threshold (POT) model of EVT, and GPD distribution which gives a more accurate description on tail distribution of financial losses.

2 Risk Measures

Financial risk is the prospect for financial loss due to unforeseen changes in underlying risk factors. Financial risks can be classified in different ways, such as market risk, credit risk (or the risk of loss arising from the failure of a counterparty to make a promised payment), liquidity risk, operational risk (or the risk of loss arising from the failures of internal systems or the people who operate in them) and others. Market risks, in turn, can be classified as interest rate risks, equity risks, exchange rate risks, commodity price risks, and so on.

In this paper we focus on the risk measures that study the behavior of extreme returns on investment portfolios. This corresponds to the determination of the value a given variable exceeds with a given (low) probability. A typical example of such measures is the Value-at-Risk (VaR). Other less frequently used measures are the expected shortfall (ES) and the return level (R).

Value at Risk

Value-at-Risk is generally defined as the sufficient capital to cover, in most instances, losses from a portfolio over a holding period of a fixed number of days. Suppose a random variable $X$ with continuous distribution function $F$ models losses on a certain financial portfolio over a certain time horizon. $VaR_\alpha$ can then be defined as the $\alpha$-th quantile of the distribution $F$

$$VaR_\alpha = F^{-1}(1 - \alpha), \quad (1)$$

where $F^{-1}$ is defined as the inverse of the distribution function $F$. For this paper we compute a 5%, 2.5%, 1% and 0.5% VaR over a one-day holding period. For example, under the assumption of normal distribution, $F \approx N(\mu, \sigma)$.

However, by definition $VaR_\alpha$ gives no information about the size of the losses which occur with probability smaller than $1 - \alpha$, i.e. the measure does not tell how bad it gets if things go wrong. Given these problems with $VaR_\alpha$, we seek an alternative measure which satisfies this.

Expected Shortfall

Another measure of risk is the expected shortfall (ES) or the tail conditional expectation which estimates the potential size of the loss exceeding VaR. The expected shortfall is defined as the expected size of a loss that exceeds $VaR_\alpha$

$$ES_\alpha = E(X \mid X > VaR_\alpha) \quad (2)$$

Expected Shortfall, as opposed to Value at Risk, is a coherent risk measure in the sense that satisfies properties of monotonicity, sub-additivity, homogeneity, and translational invariance.

Return Level

If $H$ is the distribution of the maximum observed over successive non overlapping periods of equal length, the return level $R_m^k = H^{-1}(1 - 1/m)$ is the level expected to be exceeded in one out of $m$ periods of length $k$. For example, assuming a model for the annual maximum, the 15-years return level $R_{15}^{365}$ is on average only exceeded in one year out of every 15 years.

The return level can be used as a measure of the maximum loss of a portfolio, a rather more conservative measure than the Value-at-Risk.

3 Extreme Value Theory

When modeling the maximum of a random variable, Extreme Value Theory (EVT) plays the same fundamental role as the central limit theorem when modeling sums of random variables. In both cases, the theory tells us what the limiting distributions are. EVT provides for simple parametric
models to capture the extreme tails of a distribution. Mainly there are two broad methods of applying EVT: the first of which is based on the generalized extreme value distributions known as the Block Maximum (BMM) approach, whilst the second one is based on the Generalized Pareto Distribution (GPD) and is known as the Peak Over Threshold (POT) approach. However, the POT method uses data more efficiently, it is for that reason, we will use this approach. In the following pages, the fundamental theoretical results underlying the threshold method are presented.

**Peak Over Threshold**

The Peak Over Threshold (POT) method, considers the distribution of exceedances over a certain threshold. Our problem is illustrated in Figure 1, we consider an (unknown) distribution function $F$ of a random variable $X$. We are interested in estimating the distribution function $F_u$, for values of $x$ above a certain threshold $u$.

![Figure 1: Distribution function $F$ and excess distribution $F_u$](image)

The distribution function $F_u$ is called the excess distribution function and is defined as

$$F_u(y) = P(X - u \leq y \mid X > u) = \frac{F(u + y) - F(u)}{1 - F(u)}, \quad 0 \leq y \leq x_F - u$$  \hspace{1cm} (3)

where $X$ is a random variable, $u$ is a given threshold, $y = x - u$ are the excesses and $x_F \leq \infty$ is the right endpoint of $F$.

**Theorem Pickands (1975), Balkema and de Haan (1974).** For a large class of underlying distribution $F$, the excess distribution function $F_u$ can be approximated by GPD for increasing threshold $u$.

$$F_u(y) \approx G_{\xi, \beta}(y), \quad u \to \infty$$

where $G_{\xi, \beta}$ is the Generalized Pareto Distribution (GPD) which is given by

$$G_{\xi, \beta}(y) = \begin{cases} (1 + \frac{\xi y}{\beta})^{-1/\xi} & \text{if } \xi \neq 0, \\ 1 - e^{-y/\beta} & \text{if } \xi = 0. \end{cases}$$  \hspace{1cm} (4)

for $y \in [0, (x_F - u)]$ if $\xi \geq 0$ and $y \in [0, -\beta/\xi]$ if $\xi < 0$. Here $\xi$ is the shape parameter and $\beta$ is the scale parameter for GPD.

We define the mean excess function for the GPD with parameter $\xi < 1$ as

$$e(z) = E(X - z \mid X > z) = \frac{\beta + \xi z}{1 - \xi}, \quad \beta + \xi z > 0. \hspace{1cm} (5)$$

This function gives the average of the excesses of $X$ over a varying values of the threshold $z$.

**Risk Measures under Extreme Value Theory**

Assuming a GPD function for the tail distribution, $VaR_\alpha$, $ES_\alpha$ and $R_m^\alpha$ can be defined as a function of GPD parameters. For equation (3), if we denote $x = u + y$ then

$$F(x) = (1 - F(u))F_u(y) + F(u)$$
and replacing \( F_u \) by the GPD and \( F(u) \) by the empiric estimate \( (n - N_u)/n \), where \( n \) is the total number of observations and \( N_u \) the number of observations above the threshold \( u \), we obtain

\[
\hat{F}(x) = \frac{N_u}{n} \left( 1 - (1 + \frac{\hat{\xi}}{\hat{\beta}(x-u)})^{-1/\hat{\xi}} \right) + \frac{1 - N_u}{n} = \frac{1 - N_u}{n} \left( 1 + \frac{\hat{\xi}}{\hat{\beta}(x-u)} \right)^{-1/\hat{\xi}}.
\]  

(6)

Inverting equation (6) for a given probability \( \alpha \) gives

\[
\bar{VaR}_\alpha = u + \frac{\hat{\beta}}{\hat{\xi}} \left( \left( \frac{n}{N_u} \right)^{-\xi} - 1 \right).
\]  

(7)

If we add and subtract \( VaR_\alpha \) in the equation (2) and we obtain

\[
ES_\alpha = VaR_\alpha + E(X - VaR_\alpha | X > VaR_\alpha)
\]

where the second term on the right is the expected value of the exceedances over the threshold \( VaR_\alpha \). So, for equation (5) where \( z = VaR_\alpha - u \) and \( \xi < 1 \) then

\[
ES_\alpha = VaR_\alpha + \frac{\hat{\beta} + \hat{\xi}(VaR_\alpha - u)}{1 - \xi} = \frac{VaR_\alpha}{1 - \xi} + \frac{\hat{\beta} - \hat{\xi}u}{1 - \xi}.
\]  

(8)

We know that

\[
P(X > x | X > u) = \frac{P(X > x)}{P(X > u)} = \left[ 1 + \frac{\hat{\xi}}{\hat{\beta}}(x-u) \right]^{-1/\hat{\xi}}.
\]

Hence, the level \( x_m \) that is exceeded on average once every \( m \) observations is the solution of

\[
P(X > x) = \frac{N_u}{n} \left[ 1 + \frac{\hat{\xi}}{\hat{\beta}}(x_m - u) \right]^{-1/\hat{\xi}} = \frac{1}{m}
\]

where \( P(X > u) = N_u/n \) is the empiric estimate. Rearranging,

\[
x_m = u + \frac{\hat{\beta}}{\xi} \left[ \left( \frac{mN_u}{n} \right)^{\hat{\xi}} - 1 \right].
\]

For presentation, it is often more convenient to give return levels on an annual scale, so that the \( M \)-year return level is the level expected to exceed once every \( M \) years. If there are \( k \) observations per year, this corresponds to the \( m \)-observation return level, where \( m = M \times k \). Hence, the \( M \)-year return level is defined by

\[
x_M = u + \frac{\hat{\beta}}{\xi} \left[ \left( \frac{kMN_u}{n} \right)^{\hat{\xi}} - 1 \right].
\]

4 Empirical results

We consider an extreme value approach, working with the daily losses series of AFAP SURA NAV over a period of eighteen years (1997-2015). AFAP SURA is an Uruguayan pension fund manager, with more than 310,000 clients (almost 10% of Uruguay’s total population) and assets under management over USD 2 billion, being the second largest pension fund manager in Uruguay. The empirical study uses the series of daily losses of AFAP SURA NAV, containing 4,802 trading days. The left panel of Figure 2 shows a graph of the daily evolution of AFAP SURA NAV values, and the right panel the daily return.
As shown in Figure 2, returns don’t appear to have a normal distribution and they exhibit dependence in the second moment. Something common in financial returns.

In practice, we have to consider two important aspects, the selection of the threshold $u$ and the independence of the exceedances. For example, the left panel of Figure 3 shows 182 exceedances for the threshold $u = 0.5$, clearly there is a concentration of exceedances in the years 2002 and 2009. In the right panel we use a cluster technique to reduce dependence of the exceedances and then we have 59 exceedances. The clusters are identified as follows. The first exceedance of the threshold initiates the first cluster. The first cluster then remains active until either ten consecutive values fall below (or are equal to) the threshold. The next exceedance of the threshold (if it exists) initiates the second cluster, and so on.

The choice of the threshold $u$ is the important issue to deal with, $u$ too high results in too few exceedances and consequently high variance estimators. On the other hand, $u$ too small provides biased estimators and the approximation to a GPD could not be feasible. So far, there is no algorithm with a satisfactory performance for the selection of the threshold $u$ available. It is for this reason that we use graphical approaches to select the threshold $u$.

For different thresholds $u$, the maximum likelihood estimates for the shape and the modified scale parameter (modified by subtracting the shape multiplied by the threshold) are plotted against the thresholds (see, Figure 4). If the threshold $u$ is a valid threshold to be used for peaks over threshold modeling, the parameter estimates depicted should be approximately constant above $u$. After seeing Figure 4, we choose the threshold $u = 0.5$. 

---

Figure 2: Daily evolution and return of AFAP SURA NAV since 1997 to 2015.

Figure 3: Daily losses over the threshold $u = 0.5$. 

---
The results of maximum likelihood estimation of the GPD parameters (with the chosen threshold \( u = 0.5 \)) are \( \xi = 0.5175 \) (s.e 0.1919) and \( \beta = 0.3568 \) (s.e 0.0792). Figure 5 shows how GPD fits to the 59 exceedances.

One of the purposes of this paper is to determinate the maximum loss of the portfolio, in Figure 6 we show the return level for different periods of time. The return levels are interpreted as follows, a maximum loss of 3.26% in the portfolio is expected once every five years, a maximum loss of 4.75% is expected every 10 years, 6.88% every 20 years and 11.14% every 50 years.
In Table 1 we report 95%, 97.5%, 99% and 99.5% Value at Risk and Expected Shortfall estimates for two different estimation methods. The performance of the different methods can be evaluated by comparing the estimates with the actual losses observed (empirical result). VaR approaches based on the assumption of normal distribution are definitely overestimating low percentiles (due to the high variance estimation), and underestimating high percentiles (due to heavy tails). The absence of extreme values in the assumption of normal distribution underestimates the Expected Shortfall estimation for high percentiles. On the other hand, the extreme value approach on GPD models seems coherent with the actual losses observed and is easy to implement.

<table>
<thead>
<tr>
<th>Value at Risk: one day horizon</th>
<th>( \alpha = 5% )</th>
<th>( \alpha = 2.5% )</th>
<th>( \alpha = 1% )</th>
<th>( \alpha = 0.5% )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal model</td>
<td>0.808</td>
<td>0.975</td>
<td>1.169</td>
<td>1.301</td>
</tr>
<tr>
<td>GPD model</td>
<td>0.408</td>
<td>0.666</td>
<td>1.185</td>
<td>1.777</td>
</tr>
<tr>
<td>Empirical Result</td>
<td>0.397</td>
<td>0.664</td>
<td>1.201</td>
<td>1.769</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expected Shortfall: one day horizon</th>
<th>( \alpha = 5% )</th>
<th>( \alpha = 2.5% )</th>
<th>( \alpha = 1% )</th>
<th>( \alpha = 0.5% )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal model</td>
<td>1.030</td>
<td>1.175</td>
<td>1.348</td>
<td>1.468</td>
</tr>
<tr>
<td>GPD model</td>
<td>1.049</td>
<td>1.583</td>
<td>2.658</td>
<td>3.887</td>
</tr>
<tr>
<td>Empirical Result</td>
<td>0.991</td>
<td>1.468</td>
<td>2.314</td>
<td>3.134</td>
</tr>
</tbody>
</table>

Table 1: VaR and ES estimates for two different estimation methods

5 Conclusion

We have illustrated how Extreme Value Theory can be used to model financial risk measures such as Value at Risk, expected shortfall and return level, applying it to daily returns of AFAP SURA. Our major conclusion is that the POT model can be useful for assessing the size of extreme events. From a practical point of view we discussed how to handle the selection of the threshold \( u \) and the independence of the exceedances. After that we estimate the model parameters through maximum likelihood and quantified the return level for 5, 10, 20 and 50 years.

Finally, we compared traditional methods for risk measures with the POT model, noting that the last one provides a superior adjustment. This is because traditional models don’t take into account the instability of financial markets that cause extreme values.
References