VAR versus Expected Shortfall in Risk Measurement

A project presented to The Faculty of Mathematical Science and Informatics
University of Khartoum
In Partial fulfillment of the requirements for the degree of Bachelor of Statistics

Aasim Khalid Mohamed Khair Mohammed (Statistics-CS)
Reem Siddig Diab Idris (Statistics)

Project Supervisor: Mr.Alamin Hassan Salih

January 2025

إهداء

بسم الله الرحمن الرحيم

إلى أرواح من فقدناهم من شهداء حروب السودان، الذين حملوا بين جوانحهم آمال العيش في سلام، وجعلوا خسارتنا الإنسانية لا تعوّض إلا بعهدنا على تسديد دينهم ببناء وطننا

إلى شهداء حرب 15 أبريل، الذين ارتوت هذه الأرض بدمائهم

إلى الشهيد خطاب إسماعيل محمد خير، تقبّلك الله. كنت مقدامًا، لم تستكن رغم أن المعارك أثخنت جسدك، بل زدت عزيمةً ووهبت ذاك البدن اليافع درعًا منيعًا حتى فاضت روحك إلى بارئها ونلت ما تصبو له. جعل الله دماءك طُهرًا للوطن، وأدام ذكرك رمزًا للفراسة والجسارة

إلى كل الشعب السوداني الذي مزقته نيران الحروب و مشقات النزوح و اللجوء، أولئك الذين يحبون هذه الأرض ولكنهم أجبروا على الرحيل عنها

إلى خطاب و صحبه

إلى وطنٍ ومستقبلٍ يعمّه السلام

Contents

1.Introduction

The commencement of the Sudanese war in April of 2023 resulted in an immense loss of capital. Perhaps some organizations predicted the eruption of the conflict, but were not informed on the scale of destruction and devastation that would ensue from the battles fought on the streets of khartoum. Losses such as this are extremes and rarely occur as they deviate quite much from expected scenarios.

The definition of Probability distributions describes the probability of occurrence for every event, The normal distribution is the most commonly used distribution. The normal distribution is characterized by its bell shaped graph indicating its key trait which is that the most likely events to occur are closer to the mean and the least likely occurrences are the events further away from the center of the distribution, This defining trait is symmetrical around the mean of the distribution.

1.1.History

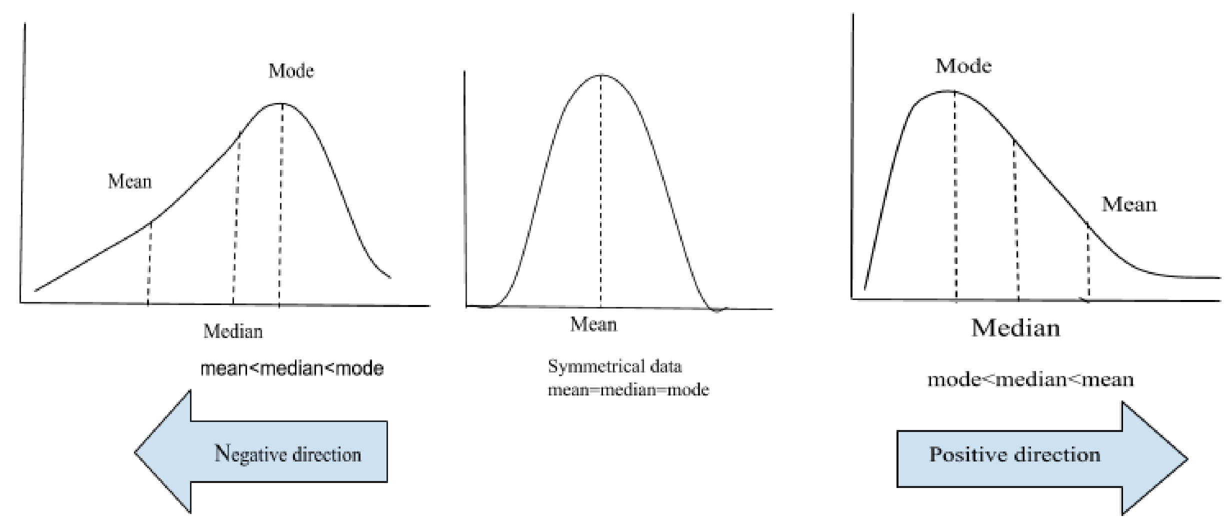
In the 17th century Galileo Galilei (1564-1642) discovered that the variation in the measurements of celestial objects weren’t entirely random, For instance the smaller the error the more likely it can occur, Smaller variations outnumbered larger ones, And he often found they were symmetrical around a central value. The early 19th century, Adrien-Marie Legendre (1752-1833) and Carl Friedrich Gauss (1777-1855) created a formula that closely matched the distribution of these mistakes, And Gauss proved its accuracy. The normal distribution appeared in other mathematical situations. Despite contributions from other mathematicians, Gauss's name became synonymous with the discovery, resulting in the term "Gaussian distribution" being used interchangeably with "normal distribution".

1.2.Properties and Features

The normal distribution is mainly described by its mean (μ) and standard deviation (σ). Assume a random variable x, then the probability density function for the normal distribution is given by

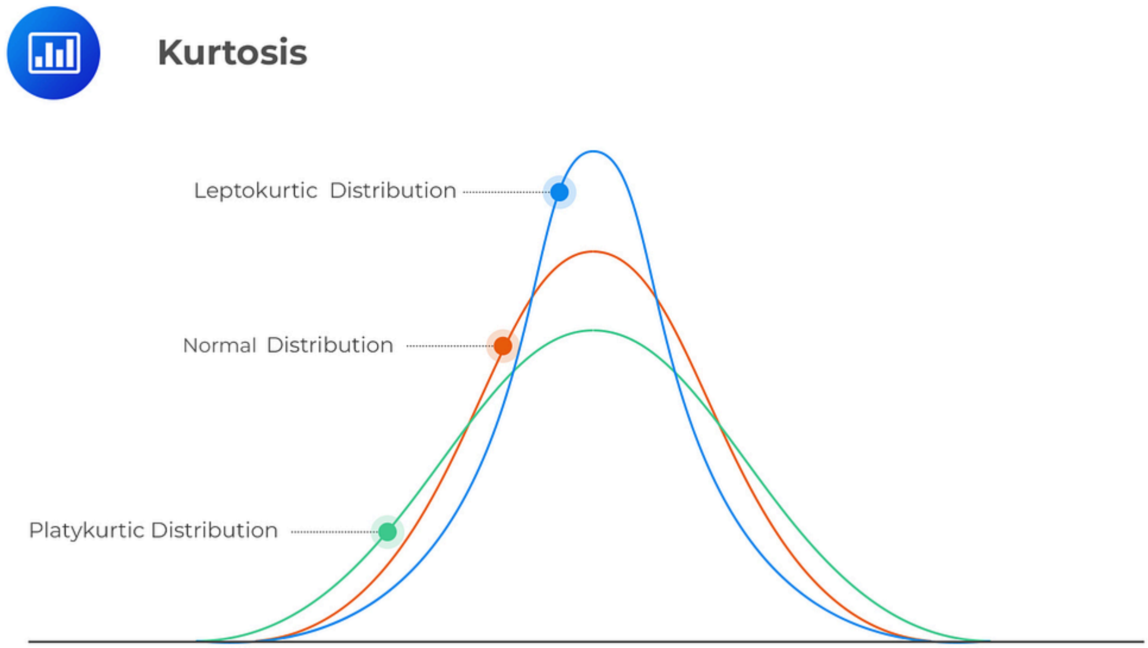
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

When a stochastic variable X follows the normal distribution we typically express that as X~ N(μ, σ^2) and it is understood that the random variable X follows the normal distribution with a mean equal to μ and standard deviation equal to σ. These parameters can be estimated by the sample mean and the sample standard deviation. The normal distribution is graphically presented as a bell shape that is symmetric around the mean, with half the observations falling on either side from the mean. Events that follow the normal distribution usually follow a trend of nearly 68.3% of the observations falling within one standard deviation from the mean, And 95.4% of the observations within two standard deviations from the mean, And 99.7% within three standard deviations from the mean. The tails of the normal distribution extend in both ways, Approaching but never reaching zero or the x axis. the more extreme the departure from the mean the lower the score's probability of occurrence. An essential characteristic is that the normal distribution has zero skew and a kurtosis of 3.



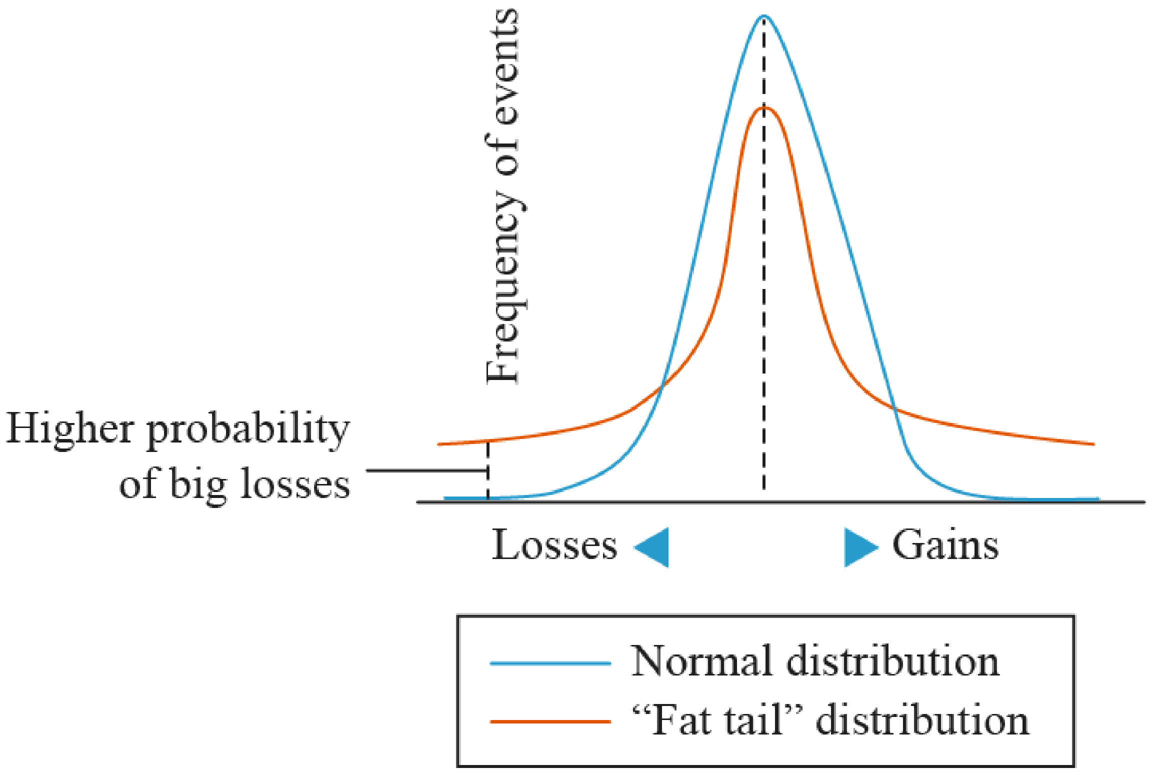
[image source]

Negative skew entails mode>median>mean it means values on the left have more probability of occurring relative to no skewness, Positive skewness is the opposite where mode<median<mean and the right tail is more likely to occur than the left relative to no skewness.



[image source]

Mesokurtic distributions have tails that resemble the normal distribution. The kurtosis for normal distribution is 3. If the kurtosis is more than 3 it is leptokurtic, this distribution may be identified as a narrow bell-shaped distribution with a higher peak than the typical distribution. In platykurtic distribution Kurtosis will be less than 3. The bell-shaped distribution in platykurtic will be wider, with a lower peak than in mesokurtic. A fat tailed distribution is one in which it is more likely for outliers to occur and more likely for them to be more extreme than the normal distribution.



[image source]

The normal distribution is one that is symmetric at the mean (i.e. zero skewness, mean=median=mode) and has kurtosis of 3 .

An important variation is the standard normal distribution which is a normal distribution with a mean of 0 and a standard deviation of 1 expressed as X~ N(0,1) which reads that the random variable X follows the standard normal distribution. A variable can be standardized by subtracting the mean from the variable and dividing by the standard deviation:

$$Z = \frac{X - \mu}{\sigma}$$

The probability density function of the standard normal distribution is:

$$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

1.3.Limitations and Solutions

The normal distribution is the most common distribution in nature, But not all phenomena follow the normal distribution. The normal distribution is limited when applied to risk modeling, it has shortcomings of accurately representing the tails of real-world distributions, The distribution of portfolio values often has "fat tails" which means that extreme values are more likely than the normal distribution would estimate. Furthermore the normal distribution curve assigns a positive for any value of future price, This doesn't make sense as the value of a portfolio can not be negative. The normal distribution is a well enough model for portfolios in many cases and is easy to use, However as markets become more unstable with crashes and booms the normal distributions limitations become more evident.

In the globalized highly dependent economy of present times, Accurately and reliably predicting and estimating losses, fluctuations and extreme events is ever more vital, As such another method of modeling these rare events at the tail of the distribution is necessary. And to that end extreme value theory (EVT) offers a solution to the limitations of the normal distribution.

EVT does not focus on the whole dataset, it looks at tails of the distribution and provides a much better understanding of extreme events. This is especially important in the field of risk measurement, as it allows us to properly assess both how likely rare yet threatening events are and what their impact could be. BMM along with the POT method are one of the two principal methods used by EVT. BMM is an aggregate extreme value model that assumes limiting distributions for the maximum, or minimum values over a specified time frame. The POT method is based on modeling exceedances distribution over a fixed threshold. The first and second approaches have a key position in extreme value theory, used throughout various fields such as finance, engineering and medicine to analyze and predict the behavior of extreme values. While the normal distribution assumption offers simplicity and is often a reasonable approximation in many situations, it can be a source of significant error in risk models during periods of market stress. EVT provides a more robust framework for estimating downside risk by focusing specifically on the tails of the distribution.

“Extreme value theory restricts the behavior of the distribution function in the tail basically to resemble a limited class of functions that can be fitted to the tail of the distribution function” (Haan & Ferreira, 2006). Extreme value theory offers a method for modeling extreme events that deviate significantly from the center of the distribution. Thus EVT is paramount for assessing risks and possible losses in rare events. Hence, it has become a prominent tool in risk measurement for multiple volatile fields.

1.4.Importance of EVT

Risk measurement is the practice of identifying, evaluating, and quantifying potential risks and losses in a project. Recently various fields require precise assessment and measurement of risks, including finance, engineering, and meteorology. EVT offers a solid theoretical basis and framework for extrapolation. “It leads to natural estimators for the relevant quantities and allows us to assess the accuracy of these estimators” (Haan & Ferreira, 2006).

In finance for example EVT can help in predicting probable market behaviors during periods of significant market turbulence, such as financial crisis. The 2008 financial crisis is a good example of why risk measurement methods that can anticipate extreme market behavior are essential. EVT focuses on the tail of the distribution, thus providing insights into the worst-case scenarios

1.5.Value at Risk and Expected Shortfall

EVT has two principle models for reaching results, Block maxima method (BMM) and peak over threshold (POT). “The block maxima the distribution of the standardized maximum is shown to follow extreme value distributions of Gumbel, Fréchet or Weibull distributions” (Singh et al., 2011), These distributions are special cases of the Generalized Extreme Value distribution (GEV) and so the series converges to GEV. When measuring the risk of a portfolio we are not always only monitoring the maximum or minimum observations rather occasionally we are interested in events that exceed a certain threshold, The peak over threshold models this. “EVT shows that the limiting distribution of excess is a Generalized Pareto Distribution” (Singh et al., 2011).

Two main methods of modeling risks are Value at Risk (VaR) and Expected Shortfall (ES). VaR estimates the maximum loss in a time frame for a confidence level. For example, the maximum possible loss of a stock price in one day might be ten dollars at a 95% confidence level. Expected Shortfall, however, creates an estimation of the loss if it happens to exceed the VaR threshold. Given our previous example, if the stock price lost more than ten dollars in one day, ES estimates how much that loss would be.

VaR and ES are essential Methods of risk measurement . VaR is simple and widely used but has been criticized for its limitation in modeling tail observations among other things, In the 2008 financial crisis most institutions used Var and to their detriment, Turner's (2009) study for the UK Financial Services Authority criticized the VaR models for failing on many days in a row, resulting in severe effects.ES addresses some of these limitations by focusing on the tail end of the loss distribution, driving more organizations to lean on ES in extreme scenarios.

1.6.Purpose of the Study

This study compares the benefits and drawbacks of VaR and ES in a risk measurement setting to further the conversation about these two methodologies. The study will provide academics and professionals new perspectives. Although ES has several limitations (Barrailler & Dufour, 2015), it has already been shown to be a more coherent model, However deciding to use either method depends on the use-case and other factors we will discuss further.

Through this study we aim to answer the following questions:

1. What are the advantages and limitations of VaR and ES in risk measurement?

2. How VaR and ES perform on modeling extreme risks?

3. What insights can be reached from using VaR and ES in varying fields?

In answering these questions, the research seeks to provide an in-depth understanding of VaR and ES in risk measurement context.

2.Literature Review

Value at Risk (VaR) and Expected shortfall have been two of the most important techniques in past years for risk measurements, when VaR was introduced in the late 1980s it was very useful in estimating the maximum possible loss in a time frame given a confidence level. However, its limitations, especially in capturing tail risks, has led to developing and adoption of ES which focuses on the average loss at worst-case scenarios. In this section we will overview the historical and current state of research regarding both techniques, the gaps in knowledge, and the future direction in research, providing a deeper analysis in advantages and limitations, modeling, application, and implications of VaR and ES in risk measurement.

2.1.Current State of Research

When searching for Value at Risk (VaR) a quite number of results appear, it's not strange as the approch was used by big firms, it kept developing to ensure more accuracy, historically, VaR is not a unified method for measuring risk, as the different calculations methodologies each produce different VaR values, In order to properly account for extreme events and tail risks, recent research has concentrated on improving VaR models. Using EVT, that can significantly increase the accuracy of VaR predictions in the presence of fat tails and extreme swings, as demonstrated by (Embrechts et al., 2018). This technique overcomes a significant shortcoming of conventional VaR models, which frequently understate the risk of uncommon but serious occurrences.

Value at Risk has a wide range of applicability-one of the major advantages-with all the applications in VaR and advancement of technology, which encourages progress in the application of artificial intelligence and machine learning methods to VAR research. These technologies provide high-end tools for better risk prediction and modeling intricate market behaviors. In fact, other studies using machine learning algorithms, such as support vector machines and neural networks to improve the forecast of VaR, have shown that machine learning models capture nonlinear interactions more effectively than traditional statistical methods, thus offering more accurate risk assessments.

Furthermore the applicability of VaR in developing countries, where financial institutions and market dynamics can diverge greatly from developed markets, has also been studied. “Research has shown that risk management procedures in emerging markets can be improved by modifying VaR models to take into consideration specific features of these countries”. Wang et al. (2020)

Because it can capture tail risk and produce a meaningful risk metric, Expected Shortfall (ES), has gained a lot of attention as a better risk measure than Value at Risk (VaR). The current state of ES research reflects the field's expanding importance in risk measurement applications, both practically and theoretically. “ES focuses on the extreme events in the tail and gives information on the range of possible extreme losses with associated probability for each outcome. ES accumulates this information into a single number by computing the average outcomes in the tail, weighted by their probabilities” (P. F. Christoffersen, 2012).

Naturally, the use of EVT has been extended to ES, similar to its application in VaR, to improve tail risk estimation. But addressing the limitations of VaR has urged the advancement of ES, and this foundational work has prompted a major effort and research in refining ES models. “There has been significant progress in developing sophisticated modeling techniques for ES. One approach involves model dependencies between risk factors, which enhances effectiveness in providing more accurate risk assessments in ES”. (McNeil, Frey, and Embrechts 2015).

The growing studies and research on ES encourage its application in the fields of climate risk, finance, environment, and governance (ESG). Researchers are exploring how ES can be integrated into risk measurement of different fields.

2.2.Gaps in Knowledge

In spite of the widespread use and enormous amount of research on Value at Risk (VaR), several knowledge gaps remain, which leads to limiting the effectiveness of VaR as an accurate risk measurement approach. Logically, identifying these gaps is crucial for improving future research in risk measurements methods.

As previously indicated in this paper, VaR models often fail to capture extreme tail risks, particularly during periods of market stress. As a result EVT has been proposed as a solution, nevertheless, its practical implementation can be difficult and computationally demanding. “And more research is required to develop practical and robust methods for accurately estimating tail risk within the VaR framework”. (Embrechts et al., 2018).

Although machine learning techniques have shown promise in developing improved VaR models, the integration of machine learning into established methods of risk measurement remains very new. More empirical studies are, therefore, needed to confirm how effective the techniques really are under changing market conditions and to begin to establish guidelines for using them.

Expected Shortfall (ES), as previously mentioned in the paper, has been acknowledged for its advantages over Value at Risk (VaR) in terms of capturing tail risks and offering accurate risk measure. Nonetheless, a number of knowledge gaps continue to exist, which restricts its broad use and efficiency. Finding these gaps is critical to directing future studies and improving risk measurements procedures.

Calculating ES can be computationally complex, especially for large and sophisticated portfolios. Using ES in live risk measurement can be challenging. “In order to make ES more usable and accessible for daily use, improvements in computational methods and the creation of more effective algorithms are needed”. (McNeil, Frey, & Embrechts, 2015).

To evaluate the effects of extreme events, scenario analysis and stress testing frequently use ES, where it is still in development. “More research is needed to establish robust frameworks for using ES in stress testing and to evaluate its effectiveness in capturing the impact of rare but severe market events” (Yamai & Yoshiba, 2005).

Similar to VaR, ES is affected by the choice of the model and the assumptions made about the distribution. Additionally, the impact of model risk on ES estimates needs further investigation, as a result developing methods to evaluate and reduce model risk in ES is important in increasing its accuracy.

2.3.Future Direction for research

After discussing the current state of research and gaps in knowledge for both methods, this can be considered as a review for historical work, achievements, and challenges, serving as a ground to investigate more and answer the questions we brought up in the purpose of the study.

3.Methodology

3.1.Framework for Risk Measurement

Risk measurement is the quantification of potential loss, particularly in volatile market settings where such losses can be significant. A modern framework for risk measurement emphasizes the need for quantifying tail risks, which previous risk measurements cannot capture. The result of this quantification is a risk measure that helps decision makers make rational choices under uncertainty which is even more challenging in turbulent market conditions where assuming normal distributions for market returns (a must-have of several traditional risk models) frequently proves inaccurate. Conventional models relying heavily on historical behavior of the market do not usually predict or take into account extreme market movements such as these; calling for more advanced tools in this respect.

3.1.1.Coherent Risk Measures

An essential part of the modern risk measurement framework is coherent risk measures. Coherent risk measures were created to overcome the limitations of conventional risk measures especially in cases where tail risks or extreme losses tend to be more likely. A risk measure is said to be coherent if it satisfies four properties outlined in (Artzner et al. 1999). These properties serve as a method to judge the intuitiveness of risk measures. The feature of a coherent risk measure are:

Monotonicity: Consider two assets X and Y, if X is consistently riskier than Y be it that it loses money more frequently or at a higher rate, The risk measure for Y cannot be higher than that of X.

$$\text{For all } X \text{ and } Y \in G \text{ with } X \leq Y, \text{ we have } \rho(Y) \leq \rho(X)$$

Sub-additivity: Regarded as the most crucial property, Two portfolios that are combined together cannot produce a higher risk than the sum of their respective risks. For example, consider two independent stocks. If we create a portfolio by combining them then this portfolio cannot lose more than sum what they can both lose individually.

$$\text{For all } X_1 \text{ and } X_2 \in G, \rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$$

Homogeneity: A portfolio that is scaled up by factor should directly translate to its risk being scaled up by the same factor. Which is an essential feature when we want to compare two portfolios of different sizes.

$$\text{For all } \lambda \geq 0 \text{ and all } X \in G, \rho(\lambda X) = \lambda\rho(X)$$

Translation invariance: Adding a risk free asset to a portfolio should decrease its risk by that same amount.

$$X \in V, a \in R \Rightarrow \rho(X + a) = \rho(X) - a$$

These properties form the basis of construction of risk measures that are capable of modeling both regular risks and risks of extreme events. Coherent risk measures ensure a more structured approach to risk, with the measures adopted being appropriate and consistent with the underlying risk that exists in the market.

In the Basel III framework coherent risk measures are vital. Basel III regulates that organizations have enough capital to cover potential extreme losses, and this is where sub-additivity guarantees that the benefits of diversification are recognized when determining capital requirements, ensuring that institutions are neither over- nor under-capitalized. By adopting coherent risk measures, institutions can make better decisions about their capital allocations, aligning with regulatory standards and improving their overall risk management practices.

3.2.Calculating Value at Risk and Expected Shortfall

3.2.1.Value at Risk

Value at risk is a risk measure that calculates the maximum value that a portfolio can lose at a confidence level within a time window, VaR for example conveys that the likelihood of stock A to lose more than a hundred dollars in the next day is 5%. VaR provides one a single value to summarize and measure the possible risk. There are several methods to calculate VaR as discussed in (Linsmeier and Pearson 1996) and they are as follow:

Historical simulation

Historical simulation is a relatively straightforward method that doesn’t require many assumptions about the distribution of the underlying data. It uses past data e.g. historical prices or market rates. This method is based on relying on historical market behavior to predict Value at Risk.

This method can be described in these five steps:

1. Identify the main market factors that impact the portfolio for which VaR is being calculated.
2. Collect past data for these market factors through a predetermined time window (e.g., the past 1000 days).
3. For each day in the historical dataset, calculate the hypothetical value of your portfolio based on the historical values of the market factors.
4. Calculate the daily returns (percentage changes) of the portfolio values from step 3 and sort these returns from worst to best.

5. The VaR is determined by the return that corresponds to your chosen confidence level. For instance, for a 95% one-day VaR, you would look for the 5th percentile of the sorted returns as 5% of the returns would be worse than this value.

Variance-Covariance

In the variance-covariance method market factors are assumed to fit a multivariate normal distribution. Usingthis assumption, the Normal distribution of the profit and loss is computed. Once the distribution of profits and losses has been determined, the Normal distribution is used in order to compute the value at risk.

1. Identify the main market factors that impact the portfolio for which VaR is being calculated.
2. Assume a multivariate Normal distribution with zero means for the percentage changes in these market factors.
3. Estimate Parameters of the Normal Distribution: Estimate the means, standard deviations, and correlations of the chosen market factors using historical data.
4. Using the estimated parameters and the risk mapping, calculate the standard deviation of the portfolio's value.
5. Multiply the standard deviation of the portfolio, the appropriate percentile of the standard normal distribution for the confidence level selected (e.g., 1.65 for 95% confidence). Here, the properties of a normal distribution are utilized to calculate the VaR.

Monte Carlo Simulation

The Monte Carlo simulation method has some parallels to historical simulation. This method differs in using historical data for the market factors. Instead a statistical distribution is chosen to approximate potential changes in market factors, Then A pseudo-random number generator generates hundreds to tens of thousands of hypothetical variations of the market variables. These are used to generate thousands of hypothetical portfolio earnings and losses for the existing portfolio, as well as their distribution. The value at risk is computed based on this distribution. The monte carlo simulation follows these steps:

1. Fit a probability distribution on each of the market factors that influence the portfolio's value.
2. Generate a large number of random paths or scenarios for the market factors, based on the probability distributions with a pseudo random number generator.
3. For each simulated scenario, calculate the hypothetical value of your portfolio.
4. Determine VaR: Analyze the distribution of the simulated portfolio values. The VaR is then determined by the specific percentile of this distribution corresponding to your chosen confidence level.

3.2.2.Expected Shortfall

Expected shortfall (ES) or conditional VaR, is the expected loss when the loss is greater than the VaR level.
How ES is calculated:

1. Calculate the VaR at the required confidence level to use as a threshold. For example, if we want to calculate the 99% ES, we first need to determine the 99% VaR.
2. Calculate ES as the mean loss exceeding the VaR. This means we take the average of all losses that are greater than the calculated VaR.

So the value of the Expected Shortfall with a confidence level α is:

$$ES_\alpha = E(X \mid X \geq VaR_\alpha(X)\ )$$

where:
X represents the random variable for portfolio returns (or losses).
$VaR_\alpha(X)$ is the VaR of the portfolio at the confidence level 'α'.
E[...] denotes the expectation operator.

Methods for Calculating ES:
ES is calculated based on the VaR so it can be computed using all the methods of VaR be it historic, variance-covariance or monte-carlo simulation, then after calculating the VaR evaluate the average of the values that fall beyond the VaR.

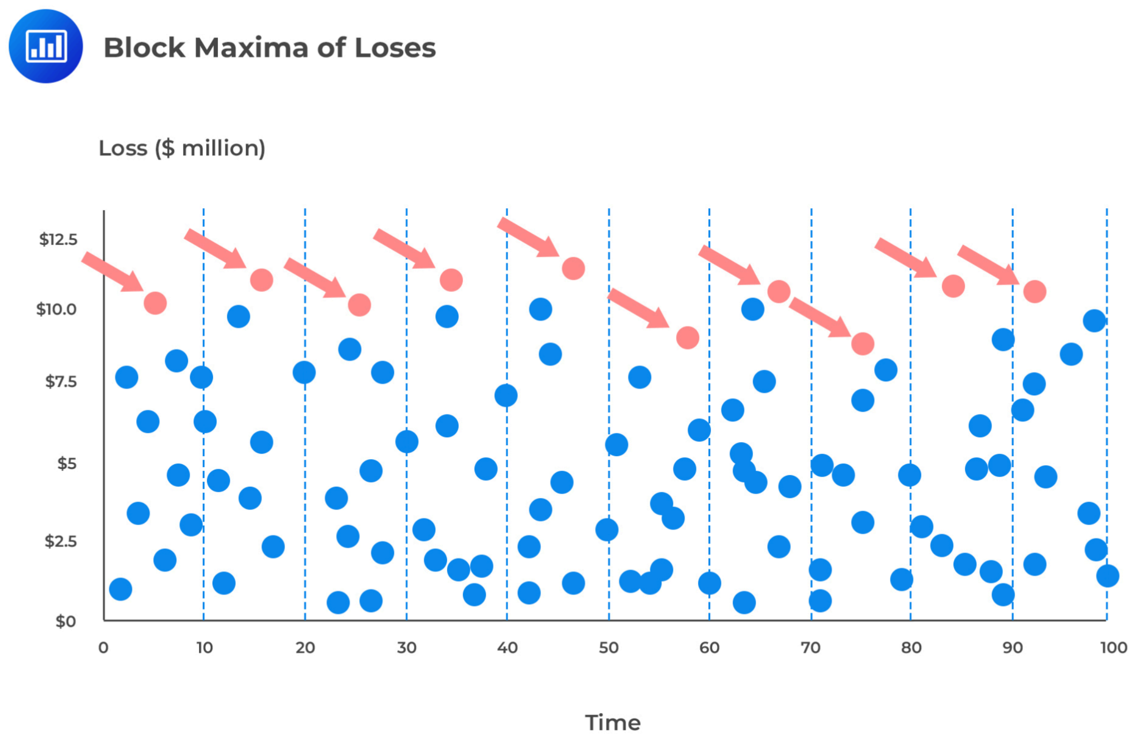
3.2.3.Using EVT to Improve VaR and ES Calculations

The extreme value theory improves the estimate of Value at Risk and Expected Shortfall, for giving a more accurate model better suited to fit into return distributions' tails. Traditional VaR and ES calculations are always underestimating the risk if the underlying data is exhibiting fat tails, most particularly under normal distribution. By specifically focusing on the tails of the distribution, EVT provides a more accurate risk measurement.
The process for using EVT to calculate VaR and ES is basically choosing the extreme value that would belong to the portfolio or asset under study then fitting an appropriate distribution to those extreme values.

There are mainly two methods for selecting extreme values:

Block Maxima Model

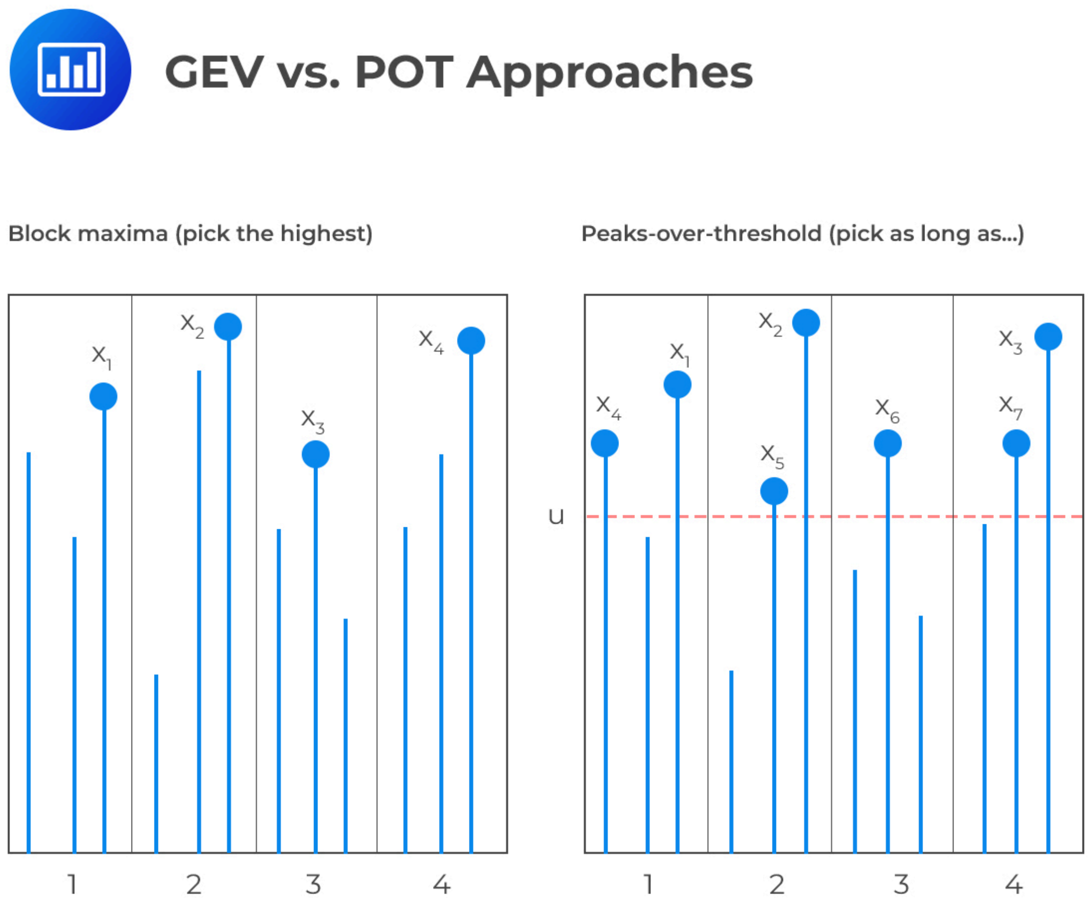
The block maxima method (BMM) is used for determining extreme data points from the overall data sample to later fit a distribution to these values, The block maxima method is mostly used in scenarios where the risk has some seasonality. In the block maxima method the data is divided into groups or blocks of equal size, From these blocks the maximum (or minimum depending on your purpose) is selected to be part of the extreme values sample used to fit the extreme distribution.



[Image Source]

Peak Over Threshold

Peak over threshold (PoT) involves picking a threshold in the data and selecting every data point that falls beyond it (higher or lower depending on the portfolio). Selecting a threshold that is too low would include more values that might not be extreme but would entail having more data to fit the distribution to, conversely setting a very high threshold would lead to ensuring only extreme values are represented but might result in too few data points to fit the distribution to.



[Image Source]

EVT primarily relies on two families of distributions to model extreme values:

Generalized Extreme Value Distribution

The generalized extreme value distribution (GEV) provides a way to model the distribution of extreme events, often it is the block maxima, which is obtained from BMM, and It contains special cases of the three specific extreme value distributions, each case captures a different tail behavior.

1. Parameters and Characteristics:

This equation has three parameters which are

Location (μ): This parameter moves the distribution along the horizontal axis, so that the increase in μ shifts the distribution to the right, and a decrease in μ shifts it to the left.

Scale (σ): This is the parameter of the spread or dispersion of the distribution, larger values of σ imply a wider distribution, and with smaller values of σ implying a narrower distribution.

Shape (ξ): Because it controls the tail behavior of the distribution and produces the three GEV variants, this is considered to be a very important parameter.

2. GEV Equation:

$$H_{\xi,\mu,\sigma}(x) = \begin{cases} \exp\left(-(1+\xi\frac{x-\mu}{\sigma})^{-1/\xi}\right) & \text{if } \xi \neq 0 \\ \exp\left(-e^{-(x-\mu)/\sigma}\right) & \text{if } \xi = 0 \end{cases}$$

Where:

$$(1+\xi\frac{x-\mu}{\sigma}) > 0$$

3. Parameter Estimation:

Estimation of parameters of the GEV distribution is one of the crucial steps in modeling an extreme event accurately; this proper estimation allows fitting the GEV distribution to the observed block maxima effectively, hence enabling more reliable risk assessments.

The most common approach in the estimation of these parameters is called the Maximum Likelihood Estimation; this process gives the estimates of the parameters that maximize the likelihood of the data at hand.

The method involves using the likelihood function for a sample of block maxima and it is subject to the condition that $1+\xi\frac{x-\mu}{\sigma} > 0$, ensuring the function remains valid.

Then to estimate the parameters using MLE, we solve a following system of equations derived from the partial derivatives of the log-likelihood function with respect to μ, σ, and ξ. These equations are typically solved numerically, as closed-form solutions are generally not available, the MLE method provides parameter estimates that best fit the observed data to the GEV distribution.

While MLE is the most applied method, the Method of L-Moments and Bayesian Estimation are used as well. After the estimation of parameters, a good fit of the GEV model must be evaluated by assessing how well it fits the data.

There are 3 cases for the tail behavior of the distribution, resulting in the three GEV variants.

1- Fréchet Distribution ($\xi > 0$)

also referred to as the Type II extreme value distribution or Cauchy-Fréchet type, Is a distribution characterized by its heavy right tail that is polynomially decreasing. This heavy tail implies that it assigns a greater probability to extreme values.

$$H(x) = \begin{cases} \exp(-y^{-\alpha}) & y > 0 \\ 0 & y \leq 0 \end{cases}$$

Where $\alpha = \xi^{-1}$, $y = 1+\xi\frac{x-\mu}{\sigma}$

2- Gumbel Distribution ($\xi = 0$)

Also known as the Type I extreme value distribution, It is often employed to model the distribution of the maximum of a large set of independent events when the underlying distribution exhibits characteristics of an exponentially decreasing tail.

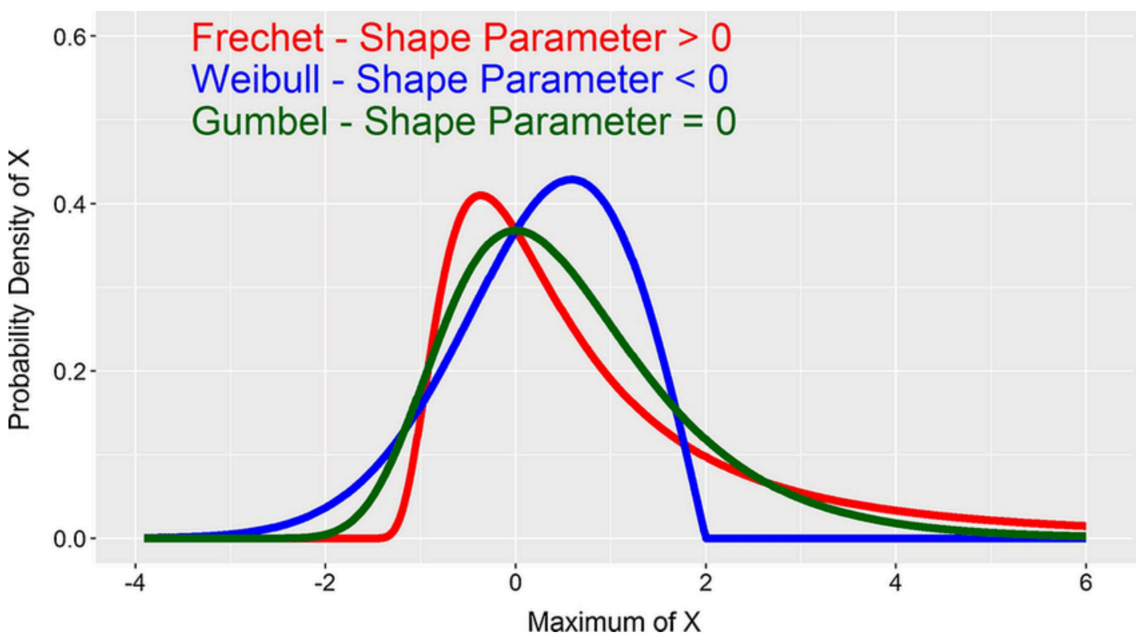
$H(x) = \exp(-e^{-(x-\mu)/\sigma})$, $x \in \mathbb{R}$

3- Weibull Distribution ($\xi < 0$):

While commonly associated with modeling minima, such as lifetimes of products, the Weibull distribution can also describe the distribution of maxima under certain conditions. When applied to maxima, it is known as the Type III extreme value distribution. A key characteristic of the Weibull distribution (for maxima) is its finite upper bound, meaning there is a limit to how large the maximum value can be.

$$H(x) = \begin{cases} \exp(-(-y)^{-\alpha}) & y < 0 \\ 1 & y \geq 0 \end{cases}$$

where $\alpha = \xi^{-1}$, $y = -(1 + \xi \frac{x - \mu}{\sigma})$



Generalized Pareto Distribution

The generalized Pareto distribution (GPD) is a three-parameter distribution often employed to model the distribution of values that exceed a specific high threshold, referred to as threshold exceedances. It is closely related to the GEV distribution: given a sufficiently high threshold, the distribution of values surpassing that threshold can be approximated by the GPD.

GPD is considered to be more efficient in modelling limited data since it fits the exceedance over a specific threshold which makes it less dependent on a large data set as BMM requires.

1. Parameters and Characteristics: The GPD is characterized by location (μ), scale (σ), and shape (ξ) parameters.

- **Location (μ)**: This is the parameter responsible for the distribution along the x-axis; it thus defines essentially the threshold value *u* above which excesses are considered.
- **Scale (σ)**: It is a parameter of dispersion that characterizes the scale or variability of the distribution around the threshold.
- **Shape (ξ)**: It is a crucial parameter that describes the tail behavior of the distribution. The shape parameter can take positive, zero, and negative values, which implies the following three different tail behaviors:
    - $\xi>0$: Indicates a heavy-tailed distribution. The probability of very large values is higher, this is quite common in financial markets, as extreme losses can happen.
    - $\xi=0$: The exponential distribution: Lighter tail, where the extreme values are less likely to occur.
    - $\xi<0$: Bounded distribution, where there is a finite upper bound to the values.

In the context of threshold exceedances, the GPD is particularly useful in addressing this issue, since it can represent the extreme market events in a very efficient way, modeling precisely the tail of the distribution beyond a selected high threshold *u*.

2. GPD Equation:

$$F_{\xi,\mu,\sigma}(x) = \begin{cases} 1-(1+\xi\frac{x-\mu}{\sigma})^{-1/\xi} & \xi \neq 0 \\ 1-e^{-(\frac{x-\mu}{\sigma})} & \xi = 0 \end{cases}$$

3. Parameter Estimation:

The parameter estimation for the GPD is a major step in applying EVT to the data. and the most widely used method for parameter estimation is the Maximum Likelihood Estimation. Given the observed data, this method basically determines the values of the parameters that maximize the likelihood function; it is especially suited to applications in EVT because it gives efficient and unbiased estimates even in the cases of extreme values.

The steps below are followed in applying MLE to the GPD:

1. Choose a high threshold *u* and determine the number of exceedances.
2. Derive the likelihood function using the GPD, after using the exceedances.
3. Optimize the likelihood function for parameters estimation μ**,** σ and ξ.

The choice of the threshold *u* is important because it has a direct influence on the number of exceedances, which determines the estimates of parameters stability. A threshold that is too low may include non-extreme values, while a threshold that is too high may result in too few exceedances for reliable estimation. Given the difficulties in estimating this diverging threshold, researchers' advice to plot estimators of various thresholds and try to find a stable region, in this case, the estimator has a significant bias, which complicates interval estimation.

Integrating EVT with VaR and ES Calculations:

After demonstrating the standard methods of calculating VaR and ES, as well as a detailed demonstration of EVT, the BMM and POT, and the EVT-based distribution, in this section, we will clarify the integration of EVT with VaR and ES calculations. As stated previously, standard methods frequently fall short, especially in the presence of fat-tailed return distributions, emphasizing that EVT serves as an important complement to these traditional methods, providing greater accuracy in estimating extreme events while also addressing the limitations of methods such as variance-covariance and historical simulation.
EVT provides a comprehensive theoretical foundation to the models describing extreme events, EVT's unique feature is it quantifies a process's stochastic behavior at varying scales, and it usually requires estimation of the probability of events that are more extreme than any other that has been previously observed.

There are two primary approaches to integrating EVT with VaR and ES calculations:

Static Approach: This approach utilizes the GPD for modeling the distribution of exceedances beyond a selected threshold *u*. The fitting of the GPD is performed based on maximum likelihood estimation, which obtains parameters for deriving VaR and ES using analytical expressions. This approach provides a straightforward method for estimating extreme risks in a static context, but does not account for time-varying market dynamics.

The **(Pickands (1975), Balkema and de Haan (1974))** theorem provides the asymptotic tail distribution of a random variable, when its true distribution is unknown. Unlike the Fisher-Tippett-Gnedenko theorem, which deals with a sample's maximum, the **Pickands-Balkema-De Haan theorem** describes values over a threshold.

For unknown distribution function F of a random variable X the **Pickands-Balkema-De Haan theorem** describes the conditional distribution function F_u of the variable X over a certain threshold u, This is the so-called conditional excess distribution function, defined as:

$$F_u(y) = P(X - u \leq y | X > u) = \frac{F(u+y) - F(u)}{1 - F(u)}$$

for $0 \leq x < x_F - u$ where x_F is either the finite or infinite right endpoint of the underlying distribution F. The function F_u describes the distribution of the excess value over a threshold u given that the threshold is exceeded.

The theorem states that:
Let F_u be the conditional excess distribution function, for a large class of underlying distribution function, and large u, F_u is well approximated by the generalized pareto distribution in the following sense. Suppose that there exist functions $a(u)$, $b(u)$, with $a(u) > 0$, such that $F_u(a(u)y + b(u)) \to G_{k,\sigma}(y), \text{ as } u \to \infty,$ converge to a non-degenerate distribution, then such limit is equal to the generalized Pareto distribution:

$$F_u(a(u)y + b(u)) \to G_{k,\sigma}(y), \text{ as } u \to \infty,$$

where

- $G_{k,\sigma}(y) = 1 - (1 + ky/\sigma)^{-1/k}$, if $k \neq 0$
- $G_{k,\sigma}(y) = 1 - e^{-y/\sigma}$, if $k = 0$.

For $\sigma > 0$, $y \geq 0$, when $k \geq 0$, and $0 \leq y < -\sigma/k$ when $k < 0$.

Given the GPD equation and If there is an extreme distribution F with right endpoint x_F , we can assume that for some threshold u, and $F_u(x) = G_{\xi,\sigma}(x)$ for $0 \leq x < x_F - u$ and $\xi \in R$ and $\sigma > 0$. For $x \geq u$,

$$\begin{aligned}
\bar{F}(x) &= P(X > u)P(X > x | X > u) \\
&= \bar{F}(u)P(X - u > x - u | X > u) \\
&= \bar{F}(u)\bar{F}_u(x - u) \\
&= \bar{F}(u)\left(1 + \xi\frac{x - u}{\sigma}\right)^{-1/\xi}
\end{aligned}$$

given $F(u)$, this gives a formula for tail probabilities. The inverse of the above formula gives the high quantile of the distribution or VaR. For $\alpha \geq F(u)$, VaR is given by:

$$VaR_\alpha = q_\alpha(F) = u + \frac{\sigma}{\xi}\left(\left(\frac{1-\alpha}{\bar{F}(u)}\right)^{-\xi} - 1\right)$$

For $\xi < 1$ the ES is given by:

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 q_x(F)dx = \frac{VaR_\alpha}{1-\xi} + \frac{\sigma - \xi u}{1-\xi}$$

It has been known that the relationship between VaR and ES is that while VaR measures the threshold of extreme losses, ES gives a deeper look into the risk by considering the expected losses beyond the threshold. That is to say, ES is always greater than or equal to VaR for the same confidence level, since it averages potential losses that are worse than VaR. ES is directly related to VaR because it includes an adjustment factor that accounts for the shape parameter ξ and the variability beyond the threshold u.

Now let's consider VaR at a confidence level α represents the maximum loss not exceeded with probability α, For a loss random variable L and ES, is the expected loss given that the loss exceeds the VaR at level α. The relationship between ES and VaR can be expressed as:

$$\mathrm{ES}_\alpha(L) = \tfrac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(L)\, du$$

This integral formulation indicates that ES is the average of VaR over all confidence levels exceeding α.

On another hand the shape and scale parameters of the GPD (ξ and σ) play a crucial role in defining both VaR and ES:

- As $\xi \to 0$, the GPD resembles an exponential distribution, and the difference between VaR and ES diminishes.
- When $\xi > 0$ (heavy tails), the difference between VaR and ES widens, hence capturing extreme tail risks becomes necessary through ES.
- VaR and ES are directly proportional to σ. Higher σ leads to a higher VaR and ES, reflecting greater variability in extreme losses.

Thus, VaR and ES can also be defined as a function of estimated GPD parameters as follows:

$$F(x) = (1 - F(u))F_u(y) + F(u)$$

If n is the total observation and N_u is the number of observations above u and we replace F_u by the GPD and $F(u)$ by $(n - N_u)/n$ we get an estimator for tail probabilities:

$$\hat{F}(x) = 1 - \frac{N_u}{n}\left(1 + \frac{\hat{\xi}}{\hat{\sigma}}(x-u)\right)^{-1/\hat{\xi}}$$

The inverse of above equation with a probability p gives the VaR:

$$\widehat{VaR}_p = u + \frac{\hat{\sigma}}{\hat{\xi}}\left(\left(\frac{n}{N_u}p\right)^{-\hat{\xi}} - 1\right)$$

And using the ES equation mentioned above, the ES is given by:

$$\widehat{ES}_p = \frac{\widehat{VaR}_p}{1-\hat{\xi}} + \frac{\hat{\sigma} - \hat{\xi}u}{1-\hat{\xi}}$$

Based on the above, in POT method GPD is fitted to the excess distribution (value above threshold a u) by MLE and the confidence interval estimates are calculated by profile likelihood and then the unconditional or static estimates for VaR and ES are calculated.

Dynamic Approach: The dynamic approach for VaR forecasting, using an EVT approach as suggested by a number of research studies, among them McNeil and Frey (2000), embeds a time-varying volatility component. This is done through first fitting the historical data into a GARCH(1,1) model in order to generate residuals; second, through the application of EVT, by using the POT method on such residuals, estimating the VaR and ES becomes possible. This will enable the model to make dynamic adjustments in view of the observed market fluctuations and hence be more suitable for the forecast of extreme risks in dynamic markets.

GARCH(1,1) process is given by:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2,$$

where $\varepsilon = R_{t-1} - \mu_{t-1}$, $\mu_t = \lambda R_{t-1}$, $\alpha_0, \alpha_1, \beta > 0$, $\beta + \alpha_1 < 1$ and $|\lambda| < 1$

In contrast to static risk modeling using EVT, where we model the unconditional distribution FX(x) and are interested in loss for k days in general, the dynamic approach models the conditional return distribution conditioned on the historical data to forecast the loss over the next $k \geq 1$ days. If we follow the GARCH(1,1) model the one day ahead forecast of VaR and ES are calculated as:

$$\begin{aligned} VaR_q &= \mu_{t+1} + \sigma_{t+1} VaR(Z_q) \\ ES_q &= \mu_{t+1} + \sigma_{t+1} ES(Z_q) \end{aligned}$$

The method is implemented in following two steps:

1. A GARCH(1,1) model is fitted to the historical data by pseudo maximum likelihood estimation (PML), the model in this step gives the residuals for step-2 and also 1 day ahead predictions of μ_{t+1} and σ_{t+1}.

2. EVT (POT method) is applied to the residuals extracted from step-1 for a constant choice of threshold u to estimate $VaR(Z_q)$and $ES(Z_q)$to calculate the risk measures using the above equation.

In summary, we can say that the steps for **Integrating EVT with VaR and ES Calculations** are:

1. We first find the extreme values in the data of the returns of the asset or portfolio by using either BMM or POT.
2. We then fit the appropriate EVT distribution (usually GEV for BMM, GPD for POT) on the identified extreme values. The fitted distribution now gives a model on the behavior that the extreme values are expected to behave.
3. Estimate the parameters of the selected EVT distribution (location, scale, and shape) using your data. Sources mention the maximum likelihood estimation (MLE) as a common method for parameter estimation.
4. Calculate VaR and ES from the distribution.

4. Results and Discussion

4.1. Introduction

The code and the data used in this analysis can be found in this GitHub Repository

An extreme value theory analysis was done on conflict-related data from Sudan. A model was fitted to the data using the peak over threshold method, And the Value at risk (VaR) with a 95% confidence level for nearly five months was calculated, the expected shortfall (ES) was calculated as well, The threshold was selected based on sensitivity analysis on the parameters as well as VaR and ES, and the model adequacy was checked.

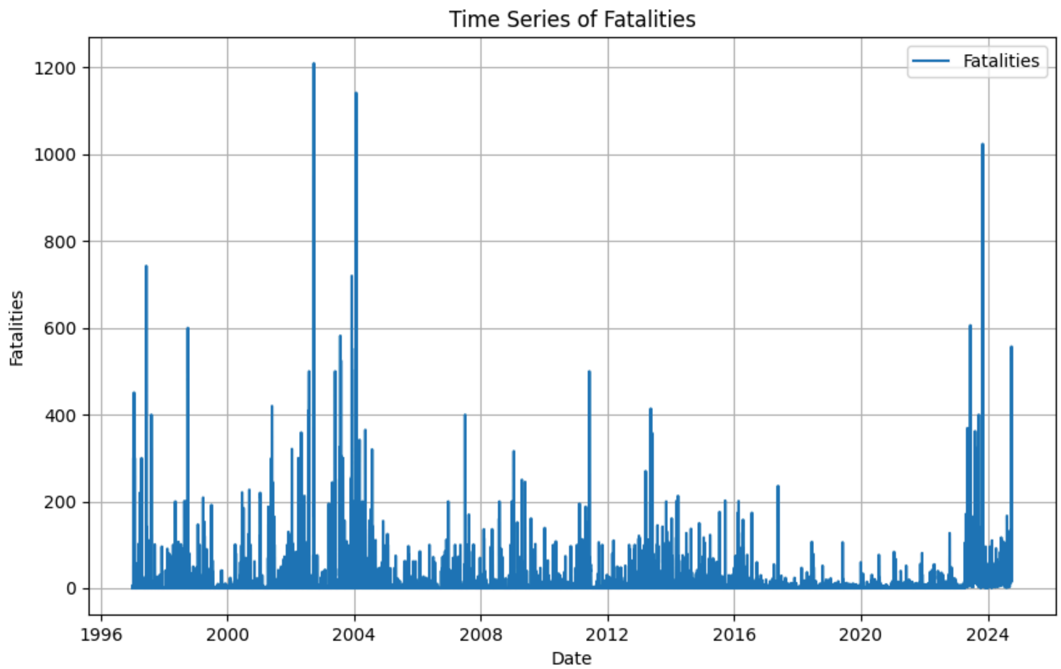
4.2. Data Source

4.2.1. Data Description

The dataset consists of 6,618 observations of conflict-related fatalities in Sudan, spanning from January 1997 to October 2024. The data shows the number of conflict fatalities per day. The dataset also contains several days with zero fatalities, reflecting periods of relative calm. The data was collected by Armed Conflict Location & Event Data (ACLED).

4.2.2. Data Descriptives

The data contains 6,618 values corresponding to days of the full date range from 1/1997 to 4/10/2024. The mean fatalities per day is 11.64 deaths, with a standard deviation of 45.3, and a median of 0 fatalities per day, only the 75% percentile is higher than 0 at 5 fatalities per day. However, the maximum number of fatalities recorded in a single event was 1,210, This shows that the data has very extreme values.



4.3. Code Stack

Microsoft Excel was used to preprocess the data and prepare it for analysis. Python programming language was used to process and analyze the data, This language was picked for its generality and wide use in EVA. The library of pandas was used to handle the data, And matplotlib to visualize the results of the analysis clearly, The main library used is pyextremes which is a library aimed at performing univariate Extreme Value Analysis.

4.4. Preprocessing

The original data source contained conflict data from all of Africa between 1997 and Oct of 2024 And captured several variables other than fatalities. First all records from countries other than Sudan were dropped, Second only the Date and Fatalities variables were kept, Third the fatalities were aggregated on the date by summing the deaths for each day, the data contained duplicated entries which were removed to ensure it captured only the total fatalities per day. The data was loaded into the Python script and the Date column was transformed into a datetime variable which is a requirement of pyextremes, this datetime column was used as an index for the data, And the fatalities variable was selected as the target of the analysis. Missing dates were filled with 0 as these days weren't recorded in data collection and hence had no conflict events and thus no fatalities. Any null values were dropped.

4.5. Extreme Value Theory (EVT) Application

4.5.1. Model Fit

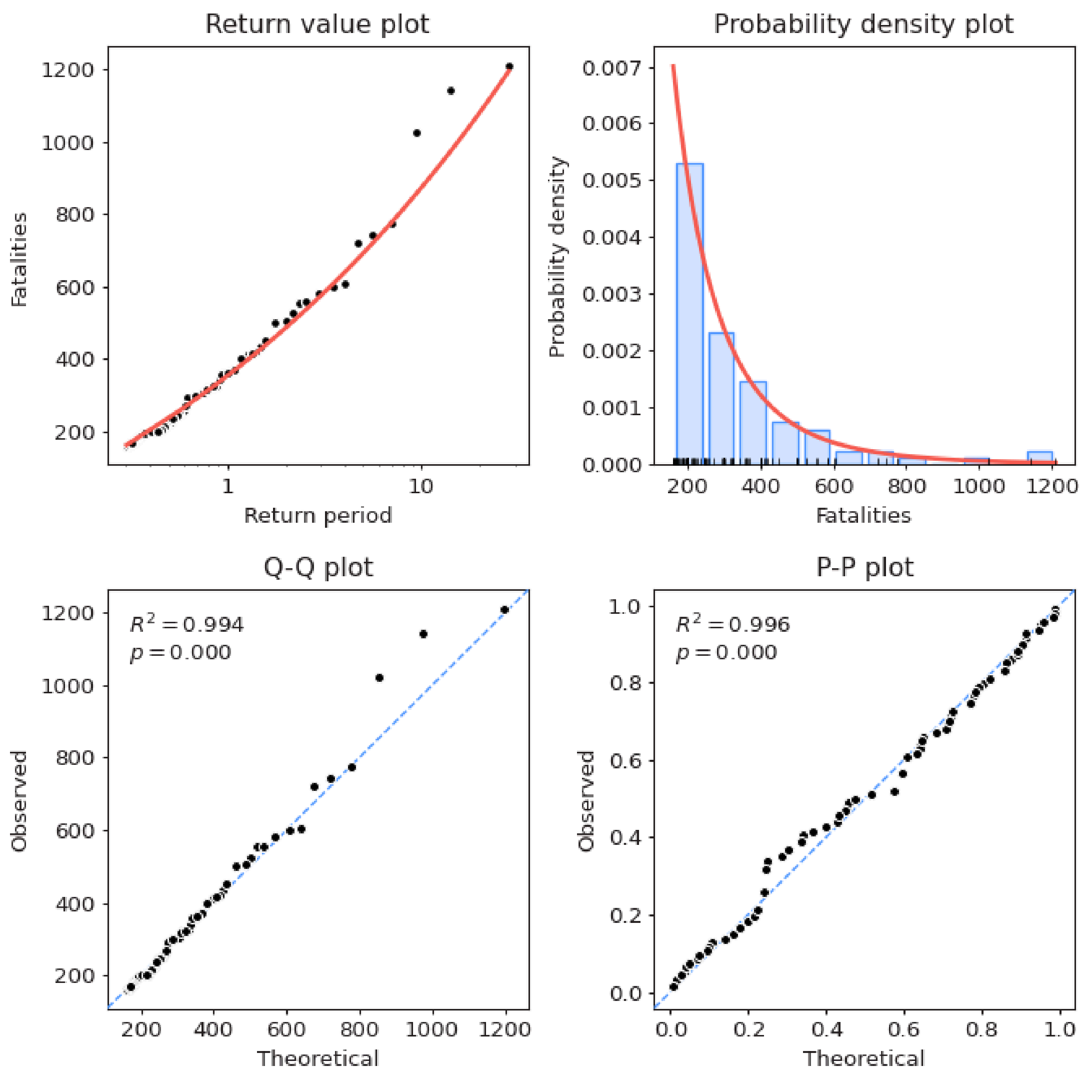
To fit the EVT model. Extreme values were selected using the peak-over-threshold method. A declustering window of 1 day was used to ensure that the data is independent which is a requirement of GPD. A threshold of 160 fatalities was selected based on analysis of the stability of the scale and shape parameters as well as the return value, mean residual life was also used to acquire a range of valid thresholds, This range was further reduced with sensitivity analysis on the VaR and ES values as well as diagnostic plots of return value, probability density, Q-Q and P-P plots, further analysis on the model fit using Kolmogorov-Smirnov test to check if which thresholds produce a model that fits the data, the smallest of these thresholds was selected ensuring enough extreme values to fit the model on, Which was 160.

The parameter estimation method used was maximum likelihood estimation. Using the clustering window, the threshold, MLE and POT method the model was fit to the tail of data, The parameters of the distribution are:

1. **Location (μ)**: This is the threshold itself at 160 fatalities per day, Only values over this were considered as extreme.
2. **Scale (σ)**: Around 141.58 and describes the spread of extreme values.
3. **Shape (ξ):** Equal to 0.195 which is greater than zero indicating the distribution is a Frechet type and that it exhibits heavy tails so extreme fatalities are more likely to occur.

4.5.2. Model Evaluation

A diagnostic plot was used to assess the model, which included 4 plots:



- **Return Value Plot:** The plot displays a close fit between the theoretical return values and the observed, indicating a good fit.
- **Probability density Plot:** The theoretical and observed probabilities are closely matched, supporting the fit of the distribution.

- **Q-Q Plot:** The plot also shows a good match between theoretical and observed values, with a small deviation for higher values which is expected of EVA data. The R^2 =0.994 highlights a strong similarity between the distribution of the data and our model, With a significant p-value =0.000<0.05.

- **P-P Plot:** This plot further supports that the model fits the data well with most points falling on or close to the line, The R^2 = 0.996 with a p-value of 0.000 confirming that the model fits the data very well.

To further test the goodness of fit of the model, A Kolmogorov-Smirnov (KS) Test was performed, with a significance level alpha=0.05, Null hypothesis: the data follows the general Pareto distribution and the alternative hypothesis that the data doesn't follow the general Pareto distribution.

The result of the test was that the test statistic was 0.0935, and the p-value was 0.3668.
Since the p-value is larger than the significance level, we can't reject the null hypothesis, thus we assume the data follows the general Pareto distribution at alpha=0.05, this is further supported by the test statistic being less than the critical value (0.1389).

4.5.3. VaR & ES calculation

After checking the adequacy of the mode, VaR was calculated with a confidence level of 95% and a return period of 1/(1 - confidence level) = 20 periods with period size = 7 days as 196.3, Thus the probability of fatalities exceeding 196.3 is 5% in each week for the next 20 weeks (140 days).

Using the VaR calculation and the ES equation provided earlier:

$$ES_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 q_x(F)dx = \frac{VaR_\alpha}{1-\xi} + \frac{\sigma - \xi u}{1-\xi}$$

The Expected shortfall was calculated with a 95% confidence level as 380.928, Meaning that given that a week's maximum fatality exceeds the VaR value, then the average or expectation for that fatality count is 380.928. There is a significant jump from the VaR to the ES showing that given the number of fatalities per day exceeded the var threshold, this number is more likely to to reach severely extreme levels.

4.5.4. VaR & ES Comparison

The difference between the VaR value and ES value highlights the importance of expected shortfall as it captures more information and not only gives a threshold for the extreme event but also what to expect in severe cases where the fatalities exceed that value.

5.Conclusion and Recommendations

5.1.Conclusion

The project focused on a comprehensive understanding of both VaR and ES with regards to risk measurement. The study questions guided the project on both the advantages and limitations that come inherently with both the methods. While VaR is a good starting point, because of its ease of computation and widespread adoption, VaR has limitations in the area of accurate modeling of the tail risks-especially during very extreme events. On the other hand, ES focuses on the average loss in the worst cases (i.e. the cases worse than the VaR), It gives a more general view of potential risk, though at a high computational cost.

This project evaluated the performance of VaR and ES in modeling high risks. The inherent limits of standard calculations of VaR that rely on the normal distribution assumption become very clear in these situations. EVT evolved as a powerful technique to enhance both VaR and ES calculations and provided a complex framework for modeling the tails of distributions. Using approaches such as BMM and POT, EVT enables a detailed understanding of extreme value behavior in order to pick and fit proper distributions such as GEV and GPD. The study then considered the calculations using both static and dynamic approaches, it showed equations of VaR and ES as well as their relationship.

The study focused on the insights that the application of VaR and ES in different kinds of domains brought. Apart from financial applications, these risk measures have been applied in climate risk, engineering, and environmental management. In all these fields, identification and management of risks are very important, and both VaR and ES, when their robustness is enhanced by EVT, are a valued basis for making decisions under uncertainty. The evolving context of the study of risk measurement also shows the need to go beyond simple methodologies and to employ advanced ones that could respond properly to modern market complexities.

The study uses a real-world application of the conflict in Sudan, using the POT method combined with GPD for modeling fatalities data; this shows how EVT improves risk assessment outside a financial framework.

5.2.Recommendations

Following what we have discussed during the study, various recommendations have emerged, as follows:

- The need to move beyond simple methods when modeling risks, especially when dealing with turbulent and more complex and dynamic markets, as the classic methods often fails to accurately estimate extreme values, and the use for dynamic approaches can incorporate time-varying volatility models (e.g., GARCH(1,1)). These dynamic models fit better for forecasting extreme risks in dynamic markets.

- Prioritize Expected Shortfall over Value at Risk for the measurement of risk since ES is more comprehensive as it accounts for the average of losses beyond the threshold given by VaR. Given that VaR only offers a threshold for potential losses, ES should be preferred for risk management, especially in highly volatile environments, where extreme losses are possible, it is especially crucial in conflict situations where extreme events are more likely and can result in severe consequences.

- Extending the application of VaR and ES enhanced by EVT in various fields, while the common use is in financial fields, the principles can be applied in climate risk, engineering, environmental events, and as done in this study in conflict data, as EVT can be used as a reliable basis for decision making under uncertainty in many domains.

- The need for extension and adoption, relying on risk modeling for decision making in Sudan especially during this conflict time and uncertainty, with the proper interpretation and presentation of results taking into account the sensitivity of the data and how the events and numbers are linked to life losses, injuries, and tragic incidents.

Bibliography

1. Abhay K. Singh, David E. Allen, and Robert J. Powell. Value at Risk Estimation Using Extreme Value Theory.

2. Turner, A. (2009). The Turner Review: A regulatory response to the global banking crisis. Financial Services Authority

3. Samuel Kotz, and Saralees Nadarajah. Extreme Value Distributions: Theory and Applications.

4. Martin Odening, and Jan Hinrichs. Using extreme value theory to estimate Value-at-Risk.

5. Laurens de Haan, Ana Ferreira. Extreme Value Theory *An Introduction.*

6. Matthieu Barrailler & Thibaut Dufour. Value at Risk and Expected Shortfall.

7. Paul Embrechts, Sidney I. Resnick, and Gennady Samorodnitsky. Extreme Value Theory as a Risk Management Tool.

8. Yasuhiro Yamai, and Toshinao Yoshiba. Value-at-risk versus expected shortfall: A practical perspective.

9. Alexander John McNeil, and Rüdiger Frey. Quantitative Risk Management: Concepts, Techniques, and Tools.

10. Philippe Artzner, Freddy Delbaen, Eber Jean-Marc, David D Heath. Coherent Measures of Risk.

11. Thomas J. Linsmeier and Neil D. Pearson, Risk Measurement: An Introduction to Value at Risk.

12. Torben G. Andersen, Tim Bollerslev, Peter F. Christoffersen, Francis X. Diebold. Financial Risk Measurement For Financial Risk Management.