FINANCIAL RISK MANAGEMENT 2

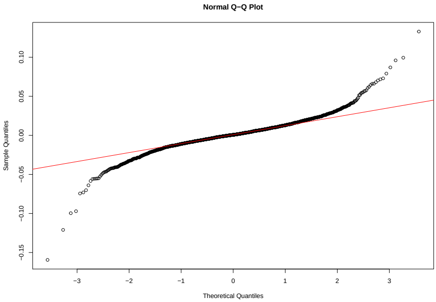
Ta Quoc Bao

Department of Mathematics, International University-VNUHCM

1. Type of Tails

- In FRM we focus principally on events that occur with a 1% or 5% probability. This is fine for most day-to-day applications in financial institutions, and methods such as GARCH and historical simulation are well suited to provide VaR for these purposes.
- In most risk analysis we are concerned with the negative observations in the lower tails, hence to follow the convention, we can pre-multiply returns by -1
- In most risk applications, we do not need to focus on the entire distribution. Since all we care about are large losses, which usually belong in the tails. For example, GARCH modeling is done with the entire distribution of returns.
- EVT, on the other hand, focuses explicitly on analyzing the tail regions of distributions (i.e., the probability of uncommon events)

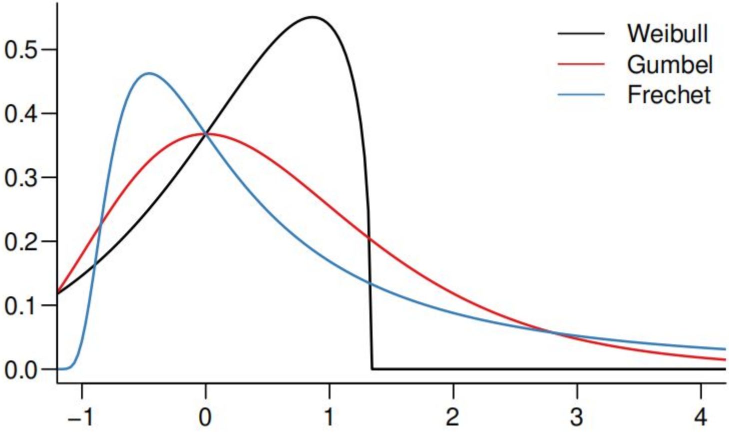
- Furthermore, there is no reason to believe the distribution of returns is symmetric; the upper and lower tails do not have the same thickness or shape. In most cases, the upper tail of returns is thinner than lower tail. For example, the upper tail of return of Microsoft is thinner than the lower tail.



- EVT can be useful in such situations as it enables us to explicitly identify the type of asymmetry in the extreme tails.

- In most risk applications we do not need to focus on the entire distribution of returns since all we care about are large losses, which usually belong in the tails.
- The main result of EVT states that, regardless of the overall shape of the distribution,the tails of all distributions fall into one of three categories as long as the distribution of an asset return series does not change over time. This means that for risk applications we only need to focus on one of these three categories:
  - Weibull Thin tails where the distribution has a finite endpoint (e.g., the distribution of mortality and insurance/re-insurance claims).
  - Gumbel Tails decline exponentially (e.g., the normal and log-normal distributions).
  - Frechet Tails decline by a power law; such tails are known as “fat tails” (e.g., the Student-t and Pareto distributions).

The following figure shows the distributions of Weibull, Gumbel and Frechet distributions



- From the figure above we see that: (i) the Weibull clearly has a finite endpoint. (ii) the Frechet tail is thicker than the Gumbel's.
- In most applications in finance, we know that returns are fat tailed

2. Generalized extreme value (GEV) distribution

Quantitative financial risk management is concerned with maximal losses (worst-case losses). Let X_i be i.i.d with distribution F continuous. Then the block maximum is given by

$$M_n := \max(X_1, X_2, ..., X_n)$$

Definition (Maximum domain of attraction)

Suppose we find normalizing sequences of real numbers $c_n > 0$ and d_n, such that $(M_n - d_n)/c_n$ converge in distribution, i.e.,

$$\begin{aligned} P((M_n - d_n)/c_n \leq x) &= P(M_n \leq c_n x + d_n) \\ &= P(M_i \leq c_n x + d_n, i = 1, 2, .., n) \\ &= F^n(c_n x + d_n) \to H(x), n \to \infty \end{aligned}$$

for some non-degenerate df H (not a unit jump). Then F is in the maximum domain of attraction of $H (F \in MDA(H))$.

Definition. The (standard) generalized extreme value (GEV) distribution is given by

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

where $1 + \xi x > 0$. Depending on the value of ξ, H_ξ become one of the three distributions:

- if $\xi > 0$, H_ξ is the Frechet
- if $\xi < 0$, H_ξ is the Weibull
- if $\xi = 0$, H_ξ is the Gumbel

Fisher-Tippet and Gnedenko theorems

- The theorems state that the maximum of a sample of properly normalized IID random variables converges in distribution to one of the three possible distributions: the Weibull, Gumbel or the Frechet
- Let $X_1, X_2, ..., X_T$ denote IID random variables (RVs) and the term MT indicate maxima in sample of size T
- The standardized distribution of maxima, $M_T := \max(X_1, X_2, ..., X_T)$, is

$$\lim_{T\to\infty} \mathbb{P}\Big(\frac{M_T - a_T}{b_T} \leq x\Big) = H(x)$$

where the constants a_T and b_T > 0 exist and are defined as $a_T = T\mathbb{E}(X_1)$ and $b_T = \sqrt{Var(X_1)}$

Example 1(Exponential distribution)
Let $X_i \sim$ Exp(λ), choosing $c_n = 1/\lambda$ and $d_n = \log(n)/\lambda$, we obtain

$$F^n(c_n x + d_n) = (1 - e^{-x}/n)^n \to \exp(-e^{-x}) = H(x)(Gumbell)$$

Example 2. (Pareto distribution)
Let X_i be i.i.d and have Pareto distribution with

$$F(x) = 1 - (\frac{\kappa}{\kappa + x})^{\theta}, x \geq 0, \theta, \kappa > 0$$

Chosing $c_n = \kappa n^{1/\theta}/\theta$ and $d_n = \kappa(n^{1\theta} - 1)$. Then we get

$$F^n(c_n x + d_n) = \left(1 + \frac{-(1 + x/\theta)^{-\theta}}{n}\right)^n \to \exp(-(1 + x/\theta))^{-\theta} = H_{1/\theta}(x)$$

This is Frechet distribution.

3. Asset returns and Fat tails

- The term “fat tails” can have several meanings, the mostcommon being “extreme outcomes occur more frequently than predicted by normal distribution”
- The most frequent definition one may encounter is Kurtosis, but it is not always accurate at indicating the presence of fat tails $(Kur > 3)$
- This is because kurtosis is more concerned with the sides of the distribution rather than the heaviness of tails

A formal definition of fat tails

The formal definition of fat tails comes from regular variation

Regular variation. A random variable, X, with distribution F has fat tails if it varies regularly at infinity; that is there exists a positive constant $\tau > 0$ such that:

$$\lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\tau}, x > 0$$

Tail distributions

- In the fat-tailed case, the tail distribution is Frechet:

$$H(x) = \exp(-x^{-\tau})$$

Lemma. A random variable X has regular variation at infinity (i.e. has fat tails) if and only if its distribution function F satisfies the following condition:

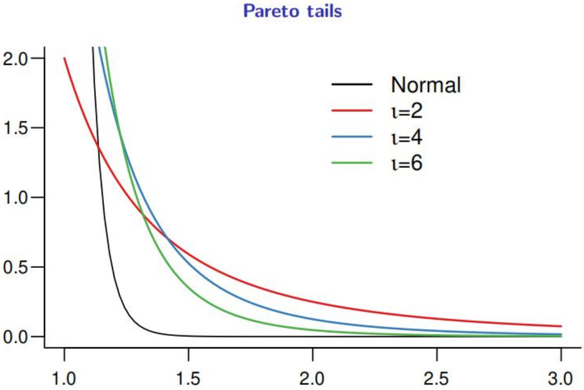
$$1 - F(x) = P(X > x) = Ax^{-\tau} + o(x^{-\tau})$$

for positive constant A, when $x \to \infty$.

- The expression $o(x^{-\tau})$ is the remainder term of the Taylor-expansion of $P(X > x)$, it consists of terms ofthe type $Cx^{-\tau}$ for constant C and $j > \tau$
- As $x \to \infty$ the tails are asymptotically Pareto distributed:

$$F(x) \approx 1 - Ax^{-\tau}$$

The following figure presents normal and fat tail distribution with respect to $\tau = 2, 4, 6$



Normal and fat distributions

- The definition demonstrates that fat tails are defined by how rapidly the tails of the distribution decline as we approach infinity
- As the tails become thicker, we detect increasingly large observations that impact the calculation of moments:

$$\mathbb{E}(X^m) = \int x^m f(x)dx$$

- If $\mathbb{E}(X^m)$ exists for all positive m, such as for the normal distribution, the definition of regular variation implies that moments $m \geq 1$ are not defined for fat-tailed data

4. Implementing EVT in practice

There are two main approaches

1. Block Maxima
2. Peaks over thresholds (POT)

Block maxima approach

- This approach follows directly from the regular variation definition where we estimate the GEV by dividing the sample into blocks and using the maxima in each block
- The procedure is rather wasteful of data and a relatively large sample is needed for accurate estimate for estimation (only the maxima of large blocks are used).

Peaks over thresholds approach

- This approach is based on models for all large observations that exceed a high threshold and hence makes better use of data on extreme values
- There are two common approaches to POT
  (i) Fully parametric models (e.g. the Generalized Pareto distribution or GPD)
  (ii) Semi-parametric models (e.g. the Hill estimator)

(i) Generalized Pareto distribution

- Consider a random variable X, fix a threshold u and focus on the positive part of $X - u$
- The distribution $F_u(x)$ (called excess distribution over u) is defined as follows

$$F_u(x) := P(X - u \leq x \mid X > u) = \frac{F(x+u) - F(u)}{1 - F(u)}$$

Mean excess function. If $\mathbb{E}(| X |) < \infty$ the mean excess function is defined by

$$e(u) := \mathbb{E}(X - u \mid X > u)$$

- Interpretation: F_u is the distribution of the excess loss $X - u$ over u, given that $X > u$. $e(u)$ is the mean of F_u as a function of u.
- For continuous $X \sim F$ with $\mathbb{E}(X) < \infty$ the following formula holds:

$$ES_\alpha = e(VaR_\alpha(X)) + VaR_\alpha(X)$$

- Note that if u is VaR of X then $F_u(x)$ is the probability that we exceed VaR by a particular amount (a shortfall) given that VaR is violated.
- The key result is that as $u \to \infty, F_u(x)$ converges to Generalized Pareto distribution (GPD), say $G_{\xi,\beta}(x)$
- $$G_{\xi,\beta}(x) = \begin{cases} 1 - (1 + \xi \frac{x}{\beta})^{\frac{1}{\xi}}, & \xi \neq 0 \\ 1 - \exp(\frac{x}{\beta}), & \xi = 0 \end{cases}$$

  where $\beta > 0$ is the scale parameter, ξ is known as shape. $x \geq 0$ when $\xi \geq 0$ and $0 \leq x \leq -\frac{\beta}{\xi}$ when $\xi < 0$.
- Therefore we need to estimate both shape ξ and scale β parameters when applying GDP
- Recall, for certain values of ξ the shape parameters, $G_{\xi,\beta}(x)$ becomes one of the three distributions (Frechet ($\xi > 0$); Weibull ($\xi < 0$), and Gumbel ($\xi = 0$)).

VaR and ES under GPD

- Assume that $F_u(x) = G_{\xi,\beta}(x), \xi \neq 0$, and some u.
- We obtain the following GPD-based formula for tail probabilities:

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

- Assuming we know $\bar{F}(u)$, inverting this formula for $\alpha \geq F(u)$ leads to

$$VaR_\alpha(X) = F^{\leftarrow}(\alpha) = u + \frac{\beta}{\xi}\left[\left(\frac{(1-\alpha)^{-\xi}}{\bar{F}(u)}\right) - 1\right]$$

and

$$ES_\alpha(X) = \frac{VaR_\alpha(X)}{1-\xi} + \frac{\beta - \xi u}{1-\xi}, \quad \xi < 1$$

(ii) Hill method

We have the approximation

$$F(x) = 1 - Ax^{-\tau}$$

The tail index τ can be approximated by the Hill method given as

$$\hat{\xi} = \frac{1}{\hat{\tau}} = \frac{1}{C_T}\sum_{i=1}^{C_T} \log \frac{x_{(i)}}{u}$$

order statistics

where C_T is the number of observations in the tail, $2 \leq C_T \leq T$, as $T \to \infty$ then $C_T \to \infty$, and $C_T/T \to 0$, the notation $x_{(i)}$ indicates sorted data, where the maxima is denoted by $x_{(1)}$, and the second-largest observation by $x_{(2)}$

Which method to choose?

- GPD, as the name suggests, is more general and can be applied to all three types of tails
- Hill method on the other hand is in the maximum domain of attraction (MDA) of the Frechet distribution
- Hence Hill method is only valid for fat-tailed data

Hill-based tail and risk measure estimates

- We have $\bar{F}(x) = 1 - F(x) \approx Ax^{-\tau}, x \geq u$. Estimate τ by $\hat{\tau}$, and u by $x_{(C_T)}$ (for C_T sufficiently small)
- Note that $A = u\bar{F}(u)$, hence $\hat{A} = (x_{(C_T)})^{\hat{\tau}}\hat{\bar{F}}(x_{(C_T)}) \approx (x_{(C_T)})^{\hat{\tau}}\frac{C_T}{T}$. So

$$\hat{\bar{F}}(x) = \frac{C_T}{T}\left(\frac{x}{x_{(C_T)}}\right)^{-\hat{\tau}}$$

then

$$\widehat{VaR_\alpha} = \left(\frac{T}{C_T}(1-\alpha)\right)^{-\frac{1}{\hat{\tau}}} x_{(C_T)}$$

and

$$\widehat{ES_\alpha} = \frac{\hat{\tau}}{\hat{\tau}-1}\widehat{VaR_\alpha}$$