Classification of Financial Risk (Basel Accord)

1. Credit risk: default risk
2. Market risk: risk due to changes in stock prices, interest rates, FX and commodity prices
3. Operational risk: includes legal and political risks

We start with the **market risk**, because

- high-quality data are available
- easier to understand
- the idea applicable to other types of risk.

**Risk measure and coherence:**

Let $L_t(\ell)$ be the loss variable of a financial position from time $t$ to time $t + \ell$. Here $\ell$ denotes a holding period. Financial loss is concerned with the distribution of $L_t(\ell)$. All risk measures available in the literature are **summary statistics** of $L_t(\ell)$.

**Coherence:**

A risk measure $\eta$ is **coherent** if for any two loss random variables $X$ and $Y$: 
1. **Subadditivity**: \( \eta(X + Y) \leq \eta(X) + \eta(Y) \).

2. **Monotonicity**: If \( X \leq Y \) almost surely, then \( \eta(X) \leq \eta(Y) \).

3. **Positive homogeneity**: For \( c > 0 \), \( \eta(cX) = c\eta(X) \).

4. **Translation invariance**: For \( c > 0 \), \( \eta(X + c) = \eta(X) + c \).

Subadditivity is related to diversification in finance.

**What is Value at Risk (VaR)(INPUT)?**

Simply put: An upper quantile of \( L_\ell(\ell) \) for a small tail probability.

- a measure of minimum loss of a financial position within a certain period of time for a given (small) probability

- the amount a position could decline in a given period, associated with a given probability (or confidence level)

Mathematically speaking: Let \( F_\ell(x) \) be the CDF of \( L_\ell(\ell) \). [The subscript of \( t \) is omitted from \( F \).]

- \( \text{VaR}_{1-p} = \inf \{ x | F_\ell(x) \geq 1 - p \} \)

- \( \Pr(L_\ell(\ell) \leq \text{VaR}_{1-p}) \geq 1 - p \) or \( \Pr(L_\ell(\ell) > \text{VaR}_{1-p}) \leq p \).

**Quantile**: \( x_q \) is the 100\( q \)th quantile of the continuous distribution \( F_\ell(x) \) if

\[
q = F_\ell(x_q), \quad \text{i.e.,} \quad q = P(L \leq x_q).
\]

In general, we have \( x_q = \min \{ x | P(L \leq x) \geq q \} \).
Figure 1: Definition of Value at Risk (VaR) for a continuous loss random variable based on the cumulative distribution function.

Figure 2: Definition of Value at Risk (VaR) for a continuous loss random variable based on the probability density function.
Discussion

First, VaR is commonly used. It is simple and some closed-form solutions are available:

1. Normal distribution: \( L_t \sim N(\mu_t, \sigma_t^2) \), then

   \[
   \text{VaR}_{1-p} = \mu_t + z_{1-p}\sigma_t,
   \]

   where \( z_{1-p} \) denotes the \((1 - p)\)th quantile of \( N(0, 1) \). In R, use \texttt{qnorm(0.95)} and \texttt{qnorm(0.99)}.

2. Student-\( t \) distribution: \( Y = (L_t - \mu_t)/\sigma_t \) is a Student-\( t \) distribution with \( v \) degrees of freedom, then

   \[
   \text{VaR}_{1-p} = \mu_t + t_{1-p, v}\sigma_t,
   \]

   where \( t_{1-p, v} \) is the \((1 - p)\)th quantile of \( t_v \). In R, use \texttt{qt(0.95, 5)} and \texttt{qt(0.99, 5)}.

3. Standardized \( t_v \) distribution: if \( v > 2 \), consider standardized \( Y \).

   \[
   Y^* = \frac{Y}{\sqrt{v/(v-2)}} = \frac{L_t - \mu_t}{\sigma_t\sqrt{v/(v-2)}}.
   \]

   Then, we have

   \[
   \text{VaR}_{1-p} = \mu_t + t^*_{1-p, v}\sigma_t\sqrt{v/(v-2)},
   \]

   where \( t^*_{1-p, v} \) is the \((1 - p)\)th quantile of standardized \( t_v \). In R with package \texttt{fGarch}, use \texttt{qstd(0.95, nu=5)} and \texttt{qstd(0.99, nu=5)}. 
Second, VaR is coherent for a normally distributed loss. For instance, consider the subadditivity:

\[
\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y) \\
= \sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y \\
\leq \sigma_x^2 + \sigma_y^2 + 2\sigma_x\sigma_y \\
= (\sigma_x + \sigma_y)^2, 
\]

where \( \rho = \text{cor}(X, Y) \). Therefore, \( \sigma_{x+y} \leq \sigma_x + \sigma_y \). This implies that \( z_{1-p}\sigma_{x+y} \leq z_{1-p}\sigma_x + z_{1-p}\sigma_y \), or equivalently, VaR of \( X + Y \) is less than or equal to the sum of VaR of \( X \) and VaR of \( Y \).

VaR, however, is not coherent in general. A simple counterexample; see also Example 3.13 of Klugman, Panjer and Willmot (2008).

**Example 1.** Suppose the CDF \( F_\ell(x) \) of a continuous loss random variable \( X \) satisfies the following probabilities:

\[
F_\ell(80) = 0.9215, \quad F_\ell(90) = 0.95, \quad F_\ell(100) = 0.97.
\]

For \( p = 0.05 \), the VaR of \( X \) is 90, because 90 is the 0.95th quantile of \( X \). We denote this by \( \text{VaR}_{0.95}^X = 90 \). Now, define two loss random variables \( X_1 \) and \( X_2 \) by

\[
X_1 = \begin{cases} 
X, & \text{if } X \leq 100 \\
0, & \text{if } X > 100
\end{cases}
\]

and

\[
X_2 = \begin{cases} 
0, & \text{if } X \leq 100 \\
X, & \text{if } X > 100.
\end{cases}
\]
These two loss variables are simply truncated versions of $X$ and we have $X = X_1 + X_2$. Since the total probability must be 1, the CDF $F^1_\ell(X)$ of $X_1$ satisfies

$$F^1_\ell(80) = 0.9215/0.97 = 0.95$$
$$F^1_\ell(90) = 0.95/0.97 = 0.9794$$
$$F^1_\ell(100) = 0.97/0.97 = 1.$$ 

The 0.95th quantile of $X_1$ is 80. Therefore, $\text{VaR}^{1}_{0.95} = 80$, where the superscript 1 is used to denote $X_1$. On the other hand, $\text{Pr}(X_2 \leq 0) = P(X \leq 100) = 0.97$. Therefore, the 0.95th quantile of $X_2$ is less than or equal to 0. We denote this by $\text{VaR}^{2}_{0.95} \leq 0$. Taking the sum, we have $\text{VaR}^{1}_{0.95} + \text{VaR}^{2}_{0.95} \leq 80$.

In this particular instance, $X = X_1 + X_2$, yet $\text{VaR}^{x}_{0.95} = 90 > \text{VaR}^{1}_{0.95} + \text{VaR}^{2}_{0.95}$. Therefore, the subadditivity of VaR fails.

Finally, VaR does not describe the actual tail behavior of the loss random variable. It is not a perfect risk measure.

**Expected Shortfall (ES):** also known as tail value at risk (TVaR) and conditional VaR (CVaR).

Simply put, ES is the expected loss of a financial position after a catastrophic event. ES of a loss variable $X$ is defined as

$$\text{ES}_{1-p} = E(X|X > \text{VaR}) = \frac{\int_{\text{VaR}}^{\infty} xf(x)dx}{\text{Pr}(X > \text{VaR})}. \quad (4)$$

From the definition, ES is the expected loss of $X$ given that $X$ exceeds its VaR.
Assume that $X$ is continuous. Let $u = F(x)$ for $\text{VaR} \leq x \leq \infty$. Then, we have $du = f(x)dx$, $F(\text{VaR}) = 1 - p$, $F(\infty) = 1$, and $x = F^{-1}(u) = \text{VaR}_u$. Equation (4) becomes

$$\text{ES}_{1-p} = \frac{\int_{1-p}^{1} \text{VaR}_u du}{p}.$$ 

Thus, ES can be seen to average all $\text{VaR}_u$ for $1 - p \leq u \leq 1$. This averaging leads to coherence of ES. For the two loss densities in Figure 3, their ES are different with the dash line corresponding to a higher value.

Closed-form solutions for ES are also available for some loss distributions.

1. Normal distribution:

$$\text{ES}_{1-p} = \mu_t + \frac{f(z_{1-p})}{p} \sigma_t,$$  

(5)
where \( f(z) \) is the pdf of \( N(0, 1) \) and \( z_{1-p} \) is the \((1-p)\)th quantile of \( f(z) \).

2. Student-\( t_v \) loss distribution:

\[
ES_{1-p} = \mu_t + \sigma_t \frac{f_v(x_{1-p})}{p} \left( \frac{v + x^2_{1-p}}{v - 1} \right),
\]

(6)

where \( f_v(x) \) denotes the pdf of \( t_v \) and \( x_{1-p} \) is the \((1-p)\)th quantile of \( f_v(x) \).

3. Standardized Student-\( t_v \) loss.

\[
ES_{1-p} = \mu_t + \sigma_t \sqrt{\frac{v}{v-2}} \frac{f^*_v(x^*_{1-p})}{p} \left( \frac{(v - 2) + [x^*_{1-p}]^2}{v - 1} \right),
\]

(7)

where \( f^*_v(x) \) is the pdf a Standardized Student-\( t_v \) and \( x^*_{1-p} \) is the \((1-p)\)th quantile of \( f_v(x) \).

Calculation of VaR involves several factors:

1. Tail probability \( p \): \( p = 0.01 \) for risk management and \( p = 0.001 \) in stress testing.

2. The time horizon \( \ell \): 1 day or 10 days for market risk and 1 year or 5 years for credit risk.

3. The CDF \( F_\ell(x) \) or its quantiles of the loss random variable.

4. The amount of the financial position or the mark-to-market value of the portfolio.
The CDF $F_\ell(x)$ is the focus of econometric modeling. We define the loss random variable as

$$x_t = \begin{cases} r_t, & \text{if the position is short,} \\ -r_t, & \text{if the position is long.} \end{cases}$$

(8)

The dollar amount of VaR is then the cash value of the financial position times the VaR of the loss variable. That is, $\text{VaR} = \text{Value} \times \text{VaR}(x_t)$.

**Methods for calculating financial risk**

1. RiskMetrics
2. Econometric modeling
3. Quantile and quantile regression
4. Extreme value theory: traditional & Peaks over Thresholds
5. Serially correlated data: extremal index

**Demonstration:** To illustrate the various methods for assessing financial risk, we consider the daily log returns of IBM stock from January 2, 2001 to December 31, 2010 for 2515 observations. See Figure 4 and we assume a long position of 1 million on the stock. The loss $x_t = -r_t$.

**RiskMetrics:** Let $x_t$ denote the daily loss. RiskMetrics assumes
that $x_t|F_{t-1} \sim N(0, \sigma_t^2)$, where $\sigma_t^2$ follows the simple model:

$$
\sigma_t^2 = \alpha \sigma_{t-1}^2 + (1 - \alpha)x_{t-1}^2, \quad 1 > \alpha > 0.
$$

Therefore, the log price $p_t = \ln(P_t)$ of a portfolio satisfies the difference equation $p_t - p_{t-1} = a_t$, where $a_t = \sigma_t \epsilon_t$ is an IGARCH(1,1) process without drift. The value of $\alpha$ is often in the interval $(0.9, 1)$ with a typical value of 0.94.

**Example 2.** Demonstration with IBM stock returns. We fit the special IGARCH(1,1) model of Eq. (9) to obtain an estimate of the parameter $\alpha$ and obtain $\hat{\alpha} = 0.943(0.007)$. In addition, using $x_{2515} = -0.061$ and $\hat{\sigma}_{2515} = 0.734$, we have $\hat{\sigma}_{2516} = 0.7133$. Consequently, using RiskMetrics, we have

$$
\text{VaR}_{0.95} = 1.173, \quad \text{VaR}_{0.99} = 1.659, \quad \text{ES}_{0.95} = 1.471, \quad \text{ES}_{0.99} = 1.901.
$$
Therefore,
\[ \text{VaR}_{0.95} = \frac{1.173}{100} \times 1,000,000 = 117,300, \quad \text{VaR}_{0.99} = 16,590. \]

Finally, for 15 days holding period, we have
\[
\begin{align*}
\text{VaR}_{0.95}(15) &= \sqrt{15} \times 11730 = 45,430 \\
\text{ES}_{0.95}(15) &= \sqrt{15} \times 14710 = 56,972.
\end{align*}
\]

**Discussion**: RiskMetrics has several advantages

1. Simplicity: Normal distribution, square-root of time rule, and multiple assets (portfolio)

2. Transparency

It also has some serious weaknesses:

1. Assumed model is rejected by empirical data

2. The square-root of time rule fails if either of the model assumptions is rejected.

Consider multiple positions: For 2 assets, we have
\[
\text{VaR} = \sqrt{\text{VaR}_1^2 + \text{VaR}_2^2 + 2\rho_{12}\text{VaR}_1\text{VaR}_2}.
\]

The generalization of VaR to a position consisting of \( m \) instruments is straightforward as
\[
\text{VaR} = \sqrt{\sum_{i=1}^{m} \text{VaR}_i^2 + 2 \sum_{i<j}^{m} \rho_{ij}\text{VaR}_i\text{VaR}_j},
\]
where $\rho_{ij}$ is the cross-correlation coefficient between returns of the $i$th and $j$th instruments and $\text{VaR}_i$ is the VaR of the $i$th instrument.

**Example 3.** Consider a simple portfolio consisting of 40% in AAA bonds and 60% on IBM stock. The market value of the portfolio is U.S. $1 million. To measure the bond returns, we employ the daily log return of the Bank of America Merrill Lynch U.S. Corp AAA total return index from January 2, 2001 to December 31, 2010. The data of bond index are obtained from the Federal Reserve Bank at St. Louis.

Figure 5 shows the log returns of the bond index. Like stock returns, bond returns also exhibit the pattern of volatility clustering and weak stationarity. For bond returns,

$$\sigma_t^2 = 0.9577\sigma_{t-1}^2 + (1 - 0.9577)r_{t-1}^2,$$

from which we have

$$\text{VaR}_{0.95} = 0.00705, \quad \text{and} \quad \text{VaR}_{0.99} = 0.00997.$$  

Recall, from Example 7.2, that for the daily log returns of IBM stock, we have

$$\text{VaR}_{0.95} = 0.01173, \quad \text{and} \quad \text{VaR}_{0.99} = 0.01659.$$  

The sample correlation coefficient of the log returns between IBM stock and AAA bond index is $-0.2215$. Consequently, for the portfolio we have

$$\text{VaR}^e_{0.95} = 0.01173 \times 0.6 = 0.00704, \quad \text{VaR}^b_{0.95} = 0.00705 \times 0.4 = 0.00282.$$
where the superscripts $e$ and $b$ denote equity and bond returns, respectively. The VaR$_{0.95}$ for the portfolio is then

$$\text{VaR}_{0.95} = \sqrt{(\text{VaR}_{0.95}^e)^2 + (\text{VaR}_{0.95}^b)^2 + 2(-0.2215)\text{VaR}_{0.95}^e\text{VaR}_{0.95}^b}$$

$$= 0.006978.$$ 

For this particular instance, we see that with tail probability $p = 0.05$, the VaR of the portfolio is less than the VaR of each component. More specifically, with $1$ million investment, we have

1. Equity market only: $\text{VaR}_{0.95} = \$11,730$.
2. Bond market only: $\text{VaR}_{0.95} = \$7,050$.
3. Portfolio (60-40): $\text{VaR}_{0.95} = \$6,978$.

This result is expected because VaR is a coherent risk measure under the normality assumption. The example, thus, demonstrates the value of diversification.

**Econometric Modeling:**

1. Use a time-series model to predict the mean return, e.g. AR or cross-sectional
2. Use a volatility model to predict the volatility, e.g. GARCH with $t_v$ innovations.

The approach considers modeling seriously, but it requires human intervention. Also, multiple-period risk measures become tedious.
Example 4. Consider again the daily log returns of IBM stock. Model 1: a Gaussian GARCH(1,1) model. The fitted model is
\begin{align*}
x_t &= -6.01 \times 10^{-4} + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0, 1) \\
\sigma_t^2 &= 4.378 \times 10^{-6} + 0.101a_{t-1}^2 + 0.884\sigma_{t-1}^2.
\end{align*}

The 1-step ahead predictions at $T = 2515$ are $-6.01 \times 10^{-4}$ and $7.82 \times 10^{-3}$, respectively, for mean and volatility. Consequently, we have
\begin{align*}
\text{VaR}_{0.95} &= 0.01227, \quad \text{ES}_{0.95} = 0.01554, \\
\text{VaR}_{0.99} &= 0.01760, \quad \text{ES}_{0.99} = 0.02025.
\end{align*}

These results imply \( \text{VaR}_{0.95} = $12,270 \) and \( \text{ES}_{0.95} = $15,540 \) for the next trading day.

Model 2: a GARCH(1,1) model with standardized Student-\( t \) innova-
tions. The fitted model is
\[ x_t = -4.113 \times 10^{-4} + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t^*_{5.751} \]
\[ \sigma_t^2 = 1.922 \times 10^{-6} + 0.0645a_{t-1}^2 + 0.9286\sigma_{t-1}^2. \]

The 1-step ahead predictions of the model for the mean and volatility at \( T = 2515 \) are \(-4.113 \times 10^{-4}\) and 0.00801, respectively. Therefore, the risk measures for the financial position are \( \text{VaR}_{0.95} = 15,450 \) and \( \text{ES}_{0.95} = 21,850 \).

**Empirical quantile and quantile regression**

Most statistical software provides empirical quantiles for a given data set.

**Example 5.** Consider the daily log returns of IBM stock. Since \( 2515 \times 0.95 = 2389.25 \), we let \( \ell_1 = 2389, \ell_2 = 2390, p_1 = 2389/2515 \) and \( p_2 = 2390/2515 \). The empirical 95% quantile of the negative log returns can be obtained as
\[ \hat{x}_{0.95} = 0.75x_{(2389)} + 0.25x_{(2390)} = 0.02654, \]

\( x_{(i)} \) is the \( i \)th order statistic of the loss variable \( x_t \). In this particular instance, \( x_{(2389)} = 0.02652 \) and \( x_{(2390)} = 0.02657 \). Finally, with \( p = 0.05 \), the sample expected shortfall is \( \text{ES}_{0.95} = 39,949 \) for the next trading day.

**Quantile Regression**: see Koenker and Bassett (1978).

Estimating the conditional quantile \( x_q|F_{t-1} \) of \( x_t \) given \( F_{t-1} \) as
\[ \hat{x}_q|F_{t-1} \equiv \inf\{\beta'z|R_q(\beta) = \min\}, \quad (10) \]
where “\( R_q(\beta_o) = \min \) means that \( \beta_o \) is obtained by
\[
\beta_o = \text{argmin} \beta \sum_{t=1}^{n} w_q(x_t - \beta' z_t),
\]
where \( w_q(.) \) is defined as before.

**Example 6.** Again, consider the daily log returns of IBM stock. We employ a quantile regression with two predictors. The first predictor is the lag-1 daily volatility of the IBM stock and the second predictor is the lag-1 VIX index of Chicago Board Options Exchange (CBOE). More specifically, we consider the quantile regression
\[
Q(q|z_t) = \sum_{t=2}^{2515} w_q(x_t - \beta_0 - \beta_1 s_{t-1} - \beta_2 v_{t-1}),
\]
where \( x_t = -r_t \) with \( r_t \) being the daily log return of IBM stock, \( s_{t-1} \) is the lag-1 daily IBM stock volatility obtained from fitting a Gaussian GARCH(1,1) model to \( x_t \), and \( v_{t-1} \) is the lag-1 VIX index obtained from CBOE. Here we use the VIX index, not the percentage VIX.

Applying the quantile regression in (11) with \( q = 0.95 \), we obtain
\[
\hat{\beta}_0 = -0.001(0.003), \quad \hat{\beta}_1 = 1.17724(0.22268), \quad \hat{\beta}_2 = 0.02809(0.01615),
\]
where the number in parentheses denotes standard error. As expected the 95th quantile of the IBM negative daily log returns depends critically on the lag-1 IBM daily volatility and marginally on the lag-1 VIX index. Based on the model, we have \( \hat{Q}(0.95|z_{2515}) = 0.013385 \). This implies that \( \text{VaR}_{0.95} = \$13,385 \) for the financial position. Figure 6 shows the negative IBM log returns \( x_t = -r_t \) and
Figure 6: Time plot of the negative daily log returns of IBM stock from January 3, 2001 to December 31, 2010. The upper line shows the 95th quantiles obtained by the quantile regression of Equation (11)

the fitted values of the quantile regression with probability $q = 0.95$. The plot also shows that VaR is time-varying and highlights the fact that the actual loss may vary when the loss exceeds VaR.

**Methods based on extreme value theory**

**Review of Extreme Value Theory (EVT):** Let $x_{(n)}$ be the sample maximum of a loss variable. EVT is concerned with the limiting distribution of $x_{(n)}$, after some proper scaling and centering, as $n \to \infty$.

For independent samples, the limiting distribution of the normalized
maximum \( r^*_n = [x_n - \mu_n]/\sigma_n \) is given by

\[
F_*(x) = \begin{cases} 
\exp[-(1 + \xi x)^{-1/\xi}] & \text{if } \xi \neq 0 \\
\exp[-\exp(-x)] & \text{if } \xi = 0
\end{cases}
\]

(12)

for \( x < -1/\xi \) if \( \xi < 0 \) and for \( x > -1/\xi \) if \( \xi > 0 \), where the subscript * signifies the normalized maximum.

The parameter \( \xi \) is referred to as the shape parameter that governs the tail behavior of the limiting distribution. The parameter \( \alpha = 1/\xi \) is called the tail index of the distribution.

The result encompasses the three types of limiting distribution of Gnedenko (1943):

- **Type I:** \( \xi = 0 \), the Gumbel family. The CDF is

  \[
  F_*(x) = \exp[-\exp(-x)], \quad -\infty < x < \infty. \quad (13)
  \]

- **Type II:** \( \xi > 0 \), the Fréchet family. The CDF is

  \[
  F_*(x) = \begin{cases} 
\exp[-(1 + \xi x)^{-1/\xi}] & \text{if } x > -1/\xi, \\
0 & \text{otherwise.}
\end{cases}
\]

  (14)

- **Type III:** \( \xi < 0 \), the Weibull family. The CDF here is

  \[
  F_*(x) = \begin{cases} 
\exp[-(1 + \xi x)^{-1/\xi}] & \text{if } x < -1/\xi, \\
1 & \text{otherwise.}
\end{cases}
\]

Two important implications of EVT.
Figure 7: Probability density functions of extreme value distributions for normalized maximum. The solid line is for a Gumbel distribution, the dotted line is for the Weibull distribution with $\xi = -0.5$, and the dashed line is for the Fréchet distribution with $\xi = 0.9$.

1. The tail behavior of the CDF $F(x)$ of $x_t$ determines the limiting distribution $F_*(x)$ of the normalized maximum.

The sequences $\{\mu_n\}$ and $\{\sigma_n\}$, however, depend on the CDF $F(x)$. See McNeil, Frey and Embtrechts (2005, Chapter 7).

2. The tail index $\xi$ does not depend on the time interval of $x_t$. That is, the tail index is invariant under time aggregation; handy in the VaR calculation.

**Empirical Estimation & Demonstration:**

1. Block Maxima Method

2. Maximum Likelihood Method
Figure 8: Block maximum of daily log returns of IBM stock, in percentages, when the sub-period is 21 trading days. The data span is from January 2, 2001 to December 31, 2010 so that there are 120 blocks. The upper plot is for positive returns and the lower plot for negative returns.

Table 1: Results of the Hill Estimator for Daily Log Returns of IBM Stock from July 3, 1962 to December 31, 1998. Standard errors are in parentheses.

<table>
<thead>
<tr>
<th>q</th>
<th>110</th>
<th>130</th>
<th>150</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_t$</td>
<td>0.380(0.036)</td>
<td>0.399(0.035)</td>
<td>0.398(0.032)</td>
</tr>
<tr>
<td>$-r_t$</td>
<td>0.356(0.034)</td>
<td>0.383(0.034)</td>
<td>0.405(0.033)</td>
</tr>
</tbody>
</table>

3. The Nonparametric Approach: shape parameter $\xi$ Hill estimator or Pickands estimator.

**Application to risk measures:** It requires a 2-step procedure because of block maxima.

For a given small upper tail probability $p$, the VaR of a financial
Figure 9: Scatterplots of the Hill estimator for the daily log returns of IBM stock. The sample period is from January 2, 2001 to December 31, 2010: the upper plot is for positive returns and the lower plot for negative returns.

Table 2: Maximum Likelihood Estimates of the Extreme Value Distribution For Daily Log Returns of IBM Stock, in percentages, from January 2, 2001 to December 31, 2010. Standard errors are in parentheses.
Figure 10: Residual analysis of fitting a GEV distribution to the negative IBM daily log returns, in percentages, from January 2, 2001 to December 31, 2010. The sub-period length used is 21 days.

The position with loss variable $x_t$ is

$$\text{VaR} = \begin{cases} 
\mu_n - \frac{\sigma_n}{\xi_n} \{1 - [-n \ln(1 - p)]^{-\xi_n}\} & \text{if } \xi_n \neq 0 \\
\mu_n - \sigma_n \ln[-n \ln(1 - p)] & \text{if } \xi_n = 0,
\end{cases} \quad (15)$$

where $n$ is the length of sub-periods.

**Summary**

We summarize the approach of applying the traditional extreme value theory to VaR calculation as follows:

1. Select the length of the sub-period $n$ and obtain sub-period maxima $\{x_{n,i}\}, i = 1, \ldots, g$, where $g = [T/n]$.

2. Obtain the maximum likelihood estimates of $\mu_n, \sigma_n$, and $\xi_n$. 


3. Check the adequacy of the fitted extreme value model; see the next section for some methods of model checking.

4. If the extreme value model is adequate, apply Eq. (15) to calculate VaR.

Multi-period VaR under EVT:

\[ \text{VaR}(\ell) = \ell^{1/\alpha} \text{VaR} = \ell^\xi \text{VaR}, \]

**Example 7.** Consider the daily log returns, in percentage, of IBM stock.

From Table 2, we have \( \hat{\alpha}_n = 1.029, \hat{\beta}_n = 1.966, \) and \( \hat{\xi}_n = 0.251 \) for \( n = 21 \). Therefore, for the left-tail probability \( p = 0.05 \), the corresponding VaR is

\[
\text{VaR} = 1.966 - \frac{1.029}{0.251} \left\{ 1 - \left[ -21 \ln(1 - 0.05) \right]^{-0.251} \right\} = 1.8902.
\]

Thus, for negative daily log returns of the stock, the upper 1% quantile is 1.8902%. Consequently, we have \( \text{VaR}_{0.95} = 1,000,000 \times 0.018902 = 18,902 \). If the probability is 0.01, then the corresponding VaR is $39,242.

If we chose \( n = 42 \) (i.e., approximately 2 months), then \( \hat{\alpha}_n = 1.1, \hat{\beta}_n = 2.489, \) and \( \hat{\xi}_n = 0.287 \). The upper 1% quantile of the loss variable based on the extreme value distribution is

\[
\text{VaR} = 2.489 - \frac{1.1}{0.287} \left\{ 1 - \left[ -42 \ln(1 - 0.01) \right]^{-0.287} \right\} = 3.5655.
\]
Therefore, for a long position of $1,000,000, the corresponding 1-day horizon VaR is $35,655 at the 1% risk level. If the probability is 0.05, then the corresponding VaR is $17,313. In this particular case, the choice of $n = 21$ gives higher VaR values.

**Discussion:** Applications using daily log returns of IBM stock.

1. $11,730 for the RiskMetrics,

2. $12,270 for a Gaussian GARCH(1,1) model,

3. $15,450 for a GARCH(1,1) model with a standardized Student-$t$ distribution with 5.75 degrees of freedom,

4. $26,540 for using the empirical quantile,

5. $13,385 for using quantile regression, and

6. $18,901 for applying the traditional extreme value theory using $n = 21$ for the length of sub-periods.

If the tail probability is 1%, then the VaR is

1. $16,590 for the RiskMetrics, a

2. $15,540 for a Gaussian GARCH(1,1) model,

3. $25,420 for a GARCH(1,1) model with a standardized Student-$t$ distribution with 5.75 degrees of freedom,

4. $50,132 for using the empirical quantile, and
5. $39,242 for applying the traditional extreme value theory using \( n = 21 \).

**Peaks Over Thresholds (POT):** A two-dimensional framework (exceedance and exceeding times)

The basic theory of the POT approach is to consider the conditional distribution of \( x = y + \eta \) given \( x > \eta \) for the limiting distribution of the maximum given in Eq. (12).

The conditional distribution of \( x \leq y + \eta \) given \( x > \eta \) is

\[
\Pr(x \leq y + \eta | x > \eta) = \frac{\Pr(\eta \leq x \leq y + \eta)}{\Pr(x > \eta)} = \frac{\Pr(x \leq y + \eta) - \Pr(x \leq \eta)}{1 - \Pr(x \leq \eta)}.
\]

Using the CDF \( F_*(\cdot) \) of Eq. (12) and the approximation \( e^{-z} \approx 1 - z \) and after some algebra, we obtain that

\[
\Pr(x \leq y + \eta | x > \eta) = \frac{F_*(y + \eta) - F_*(\eta)}{1 - F_*(\eta)}
\]

\[
= \exp \left[ - \left(1 + \frac{\xi(y+\eta-\mu)}{\sigma}\right)^{-1/\xi} \right] - \exp \left[ - \left(1 + \frac{\xi(\eta-\mu)}{\sigma}\right)^{-1/\xi} \right]
\]

\[
= \frac{1 - \exp \left[ - \left(1 + \frac{\xi(\eta-\mu)}{\sigma}\right)^{-1/\xi} \right]}{1 - \exp \left[ - \left(1 + \frac{\xi y}{\sigma + \xi(\eta - \mu)}\right)^{-1/\xi} \right]}
\]

\[
\approx 1 - \left(1 + \frac{\xi y}{\sigma + \xi(\eta - \mu)}\right)^{-1/\xi},
\]

where \( y > 0 \) and \( 1 + \xi(\eta - \mu)/\sigma > 0 \). The case of \( \xi = 0 \) is taken as the limit of \( \xi \to 0 \) so that

\[
Pr(x \leq y + \eta | x > \eta) \approx 1 - \exp(-y/\sigma).
\]
Generalized Pareto distribution:

The probability distribution with cumulative distribution function

\[
G_{\xi, \psi(\eta)}(y) = \begin{cases} 
1 - \left[1 + \frac{\xi y}{\psi(\eta)}\right]^{-1/\xi} & \text{for } \xi \neq 0, \\
1 - \exp[-y/\psi(\eta)] & \text{for } \xi = 0,
\end{cases}
\]  

(18)

where \(\psi(\eta) > 0\), \(y \geq 0\) when \(\xi \geq 0\), and \(0 \leq y \leq -\psi(\eta)/\xi\) when \(\xi < 0\), is called the generalised Pareto distribution (GPD). For GPD, suppose that the excess distribution of \(x\) given a threshold \(\eta_o\) is a GPD with shape parameter \(\xi\) and scale parameter \(\psi(\eta_o)\). Then, for an arbitrary threshold \(\eta > \eta_o\), the excess distribution over the threshold \(\eta\) is also a GPD with shape parameter \(\xi\) and scale parameter \(\psi(\eta) = \psi(\eta_o) + \xi(\eta - \eta_o)\).

Mean Excess Function

\[
e_T(\eta) = \frac{1}{N_\eta} \sum_{i=1}^{N_\eta} (x_{t_i} - \eta),
\]

(19)

where \(N_\eta\) is the number of returns that exceed \(\eta\) and \(x_{t_i}\) are the values of the corresponding returns.

Estimation:

For a given high threshold, use GPD to obtain parameter estimates. For the long position of $1 million on IBM stock, we have

\[
\text{VaR}_{0.95} = $25,855, \quad \text{ES}_{0.95} = $39,625,
\]

for the first trading day of 2011 when the threshold of 1% is used. If the threshold is 1.2%, we have

\[
\text{VaR}_{0.95} = $26,115, \quad \text{ES}_{0.95} = $39,603.
\]
Table 3: Maximum Likelihood Estimates of the Generalized Pareto Distribution For Negative Daily Log Returns of IBM Stock from January 2, 2001 to December 31, 2010. Standard errors are in parentheses and n.exceed denotes the number of exceedances.

<table>
<thead>
<tr>
<th>Thr.</th>
<th>η</th>
<th>n.exceed</th>
<th>Shape ξ</th>
<th>Scale σ</th>
<th>Location μ</th>
<th>ψ(η)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>504</td>
<td>0.107(0.042)</td>
<td>0.009(0.001)</td>
<td>-0.006(0.001)</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>410</td>
<td>0.075(0.044)</td>
<td>0.010(0.001)</td>
<td>-0.007(0.002)</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>0.8</td>
<td>610</td>
<td>0.106(0.039)</td>
<td>0.009(0.001)</td>
<td>-0.006(0.001)</td>
<td>0.010</td>
<td></td>
</tr>
</tbody>
</table>

Figure 11: Diagnostic plots for GPD fit to the daily negative log returns of IBM stock from January 2, 2001 to December 31, 2010.

Finally, for threshold of 0.8%, we have

$$\text{VaR}_{0.95} = $25,866, \quad \text{ES}_{0.95} = $39,620.$$  

**An Alternative Parameterization**: Use $\psi(\eta) = \alpha + \xi(\eta - \beta)$.

The parameter of GPD becomes $(\xi, \psi(\eta))$. 
R Demonstration for the lecture note

> da=read.table("d-ibm-0110.txt",header=T)
> head(da)
   date   return
1 20010102 -0.002206

> ibm=log(da[,2]+1)*100
> source("Igarch.R")
> mm=Igarch(ibm)

Coefficient(s):

|     | Estimate | Std. Error | t value | Pr(>|t|) |
|-----|----------|------------|---------|----------|
| alpha | 0.942857 | 0.007172   | 131.464 | < 2.22e-16 *** |

--- ### You may use the result to calculate volatility forecast and VaR ###

> fvariance=m1$par*m1$volatility[2515]^2+(1-m1$par)*ibm[2515]^2
> fvol=sqrt(fvariance)
> fvol
   beta
0.7133031

> VaR_0.95=qnorm(.95)*fvol
> VaR_0.95
   beta
1.173279

### A summary of the calculation is given below

<table>
<thead>
<tr>
<th>prob</th>
<th>VaR</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1,]</td>
<td>0.950</td>
<td>1.173279</td>
</tr>
<tr>
<td>[2,]</td>
<td>0.990</td>
<td>1.659391</td>
</tr>
<tr>
<td>[3,]</td>
<td>0.999</td>
<td>2.204272</td>
</tr>
</tbody>
</table>

#### GARCH

> xt=-log(da$return+1) % calculate negative log returns.
> library(fGarch)
> m1=garchFit(~garch(1,1),data=xt,trace=F)
> m1

Title: GARCH Modelling
Call: garchFit(formula = ~garch(1, 1), data = xt, trace = F)

Mean and Variance Equation:
    data ~ garch(1, 1) [data = xt]

Conditional Distribution: norm
Error Analysis:

|       | Estimate  | Std. Error | t value | Pr(>|t|) |
|-------|-----------|------------|---------|----------|
| mu    | -6.010e-04 | 2.393e-04  | -2.511  | 0.012044 * |
| omega | 4.378e-06  | 1.160e-06  | 3.774   | 0.000161 *** |
| alpha1| 1.011e-01  | 1.851e-02  | 5.463   | 4.67e-08 *** |
| beta1 | 8.841e-01  | 1.991e-02  | 44.413  | < 2e-16 *** |

---

> predict(m1,3)

<table>
<thead>
<tr>
<th></th>
<th>meanForecast</th>
<th>meanError</th>
<th>standardDeviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.0006009667</td>
<td>0.007824302</td>
<td>0.007824302</td>
</tr>
<tr>
<td>2</td>
<td>-0.0006009667</td>
<td>0.008043298</td>
<td>0.008043298</td>
</tr>
<tr>
<td>3</td>
<td>-0.0006009667</td>
<td>0.008253382</td>
<td>0.008253382</td>
</tr>
</tbody>
</table>

> source("RMeasure.R")

> m11=RMeasure(-.000601,.0078243)

Risk Measures for selected probabilities:

<table>
<thead>
<tr>
<th>prob</th>
<th>VaR</th>
<th>ES</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1,]</td>
<td>0.950</td>
<td>0.01226883 0.01553828</td>
</tr>
<tr>
<td>[2,]</td>
<td>0.990</td>
<td>0.01760104 0.02025244</td>
</tr>
<tr>
<td>[3,]</td>
<td>0.999</td>
<td>0.02357790 0.02574412</td>
</tr>
</tbody>
</table>

> m2=garchFit(~garch(1,1),data=xt,trace=F,cond.dist="std")

> m2

Title: GARCH Modelling
Call: garchFit(formula = ~garch(1,1), data=xt, cond.dist = "std", trace=F)

Mean and Variance Equation:

data ~ garch(1, 1) [data = xt]

Conditional Distribution: std

Error Analysis:

|       | Estimate  | Std. Error | t value | Pr(>|t|) |
|-------|-----------|------------|---------|----------|
| mu    | -4.113e-04 | 2.254e-04  | -1.824  | 0.06811 . |
| omega | 1.922e-06  | 7.417e-07  | 2.592   | 0.00954 ** |
| alpha1| 6.448e-02  | 1.323e-02  | 4.874   | 1.09e-06 *** |
| beta1 | 9.286e-01  | 1.407e-02  | 65.993  | < 2e-16 *** |
| shape | 5.751e+00  | 6.080e-01  | 9.459   | < 2e-16 *** |

---

> predict(m2,3)

<table>
<thead>
<tr>
<th></th>
<th>meanForecast</th>
<th>meanError</th>
<th>standardDeviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.0004112738</td>
<td>0.008100872</td>
<td>0.008100872</td>
</tr>
<tr>
<td>2</td>
<td>-0.0004112738</td>
<td>0.008191119</td>
<td>0.008191119</td>
</tr>
<tr>
<td>3</td>
<td>-0.0004112738</td>
<td>0.008279772</td>
<td>0.008279772</td>
</tr>
</tbody>
</table>
> m22=RMeasure(-.0004113,.0081009,cond.dist="std",df=5.751)

Risk Measures for selected probabilities:
   prob   VaR   ES
[1,] 0.950 0.01545311 0.02184843
[2,] 0.990 0.02542061 0.03294803
[3,] 0.999 0.04289786 0.05332908

##### Empirical quantiles
> ibm=-log(da[,2]+1)
> prob1=c(0.9,0.95,0.99,0.999) % probabilities of interest
> quantile(ibm,prob1)
   90%   95%   99%   99.9%
0.01736836 0.02653783 0.05013151 0.07198369
> sibm=sort(ibm) % Sorting into increasing order
> 0.95*2515
[1] 2389.25
> es=sum(sibm[2390:2515])/(2515-2389)
> es
[1] 0.03994857

### Quantile regression

> dd=read.table("d-ibm-rq.txt",header=T) % Load data
> head(dd)
   nibm vol  vix
1  -0.10947840 0.01700121  29.99
2   0.01530858 0.01614694  26.60
   ....
6  -0.00940860 0.03211091  27.99
> dim(dd)
[1] 2514  3
> dd[,3]=dd[,3]/100
> library(quantreg)
> mm=rq(nibm~vol+vix,tau=0.95,data=dd) % Quantile regression
> summary(mm)
Call: rq(formula = nibm ~ vol + vix, tau = 0.95, data = dd)
tau: [1] 0.95 % probability
Coefficients:
      Value Std. Error t value  Pr(>|t|)
(Intercept) -0.00104 0.00257 -0.40317 0.68686
 vol  1.17724 0.22268  5.28660 0.00000
 vix  0.02809 0.01615  1.73977 0.08202
> names(mm)
coefficients   x           y           residuals
dual           fitted.values   formula   terms
xlevels        call           tau         rho

> fit=mm$fitted.values
> tdx=c(2:2515)/252+2001
> plot(tdx,dd$nibm,type='l',xlab='year',ylab='neg-log-rtn')
> lines(tdx,fit,col='red')
> v1[2515]
[1] 0.008018202
> vix[2515]
[1] 17.75
> vfit=-.00104+1.17724*v1[2515]+0.02809*vix[2515]/100
> vfit
[1] 0.01338532
> mm=rq(xt~vol+vix,tau=0.99,data=dd) # 99th quantile
> summary(mm)
Call: rq(formula = xt ~ vol + vix, tau = 0.99, data = dd)
tau: [1] 0.99
Coefficients:
   Value     Std. Error  t value     Pr(>|t|)
(Intercept) 0.01182   0.00831  1.42190    0.15518
vol          1.03129   0.73125  1.41031    0.15857
vix           0.04409   0.05335  0.82641    0.40865

### Extreme value theory
> library(evir) # Load package
> par(mfcol=c(2,1))
> m1=gev(xt,block=21)
> m1
$n.all
[1] 2515
$n
[1] 120
$data
$block
[1] 21
$par.est
 sigma     mu
0.251353 1.028910 1.965850
$par.ses
 sigma     mu
0.08847742 0.09013351 0.10932034
$\text{varcov}$
\[
\begin{array}{ccc}
[1,] & [2,] & [3,] \\
[1,] & 0.007828254 & -0.001080741 & -0.003453668 \\
[2,] & -0.001080741 & 0.008124049 & 0.006145413 \\
[3,] & -0.003453668 & 0.006145413 & 0.011950936 \\
\end{array}
\]

$\text{converged}$

\[
[1] \ 0
\]

> plot(m1)

Make a plot selection (or 0 to exit):
1: plot: Scatterplot of Residuals
2: plot: QQplot of Residuals
Selection: 1

### POT method

> da=read.table("d-ibm-0110.txt",header=T)
> ibm=log(da[,2]+1)
> xt=-ibm
> m1=pot(xt,threshold=0.01)
> m1

$n$

\[
[1] \ 2515
\]

$\text{period}$

\[
[1] \ 1 \ 2515
\]

$\text{data}$

\[
[1] \ 0.01530858 \ 0.01074553 \ 0.01139063 \ \ldots
\]

attr("times")

\[
[1] \ 3 \ 6 \ 10 \ \ldots
\]

$\text{span}$

\[
[1] \ 2514
\]

$\text{threshold}$

\[
[1] \ 0.01
\]

$\text{p.less.thresh}$

\[
[1] \ 0.7996024
\]

$n.exceed$

\[
[1] \ 504
\]

$\text{par.est}$

\[
\begin{array}{cccc}
\text{xi} & \text{sigma} & \text{mu} & \text{beta} \\
0.107268254 & 0.008914461 & -0.005634968 & 0.010591597
\end{array}
\]

$\text{par.ses}$

\[
\begin{array}{cccc}
\text{xi} & \text{sigma} & \text{mu} \\
0.0415025597 & 0.0009052881 & 0.0012156539
\end{array}
\]

$\text{intensity}$

\[
[1] \ 0.2004773
\]

32
$converged
[1] 0
> plot(m1)
Make a plot selection (or 0 to exit):

1: plot: Point Process of Exceedances
2: plot: Scatterplot of Gaps
3: plot: Qplot of Gaps
4: plot: ACF of Gaps
5: plot: Scatterplot of Residuals
6: plot: Qplot of Residuals
7: plot: ACF of Residuals
8: plot: Go to GPD Plots
Selection: 0

> riskmeasures(m1,c(0.95,0.99))
   p       quantile      sfall
[1,] 0.950 0.02585540 0.03962479
[2,] 0.990 0.04744964 0.06381374
> riskmeasures(m2,c(0.95,0.99)) % Threshold=0.012
   p       quantile      sfall
[1,] 0.950 0.02611524 0.03960353
[2,] 0.990 0.04745886 0.06267327
> riskmeasures(m3,c(0.95,0.99)) % Threshold=0.008
   p       quantile      sfall
[1,] 0.950 0.02586561 0.03962012
[2,] 0.990 0.04744180 0.06376612

### Generalized Pareto distribution
> library(evir)
> da=read.table("d-ibm-0110.txt",header=T)
> ibm=log(da[,2]+1)
> xt=-ibm
> m1gpd=gpd(xt,threshold=0.01)
> m1gpd
$n
[1] 2515
$data
   [1] 0.01530858 0.01074553 0.01139063 ..... 
$threshold
[1] 0.01
$p.less.thresh
[1] 0.7996024
$n.exceed
$method
[1] "ml"

$par.est

<table>
<thead>
<tr>
<th>xi</th>
<th>beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10703752</td>
<td>0.01059601</td>
</tr>
</tbody>
</table>

$par.ses

<table>
<thead>
<tr>
<th>xi</th>
<th>beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0544269528</td>
<td>0.0007255951</td>
</tr>
</tbody>
</table>

$converged
[1] 0

$nllh.final
[1] -1733.994

> names(m1gpd)
[1] "n"  "data"  "threshold"  "p.less.thresh"
[5] "n.exceed"  "method"  "par.est"  "par.ses"
[9] "varcov"  "information"  "converged"  "nllh.final"

> par(mfcol=c(2,2))
> plot(m1gpd)
Make a plot selection (or 0 to exit):

1: plot: Excess Distribution
2: plot: Tail of Underlying Distribution
3: plot: Scatterplot of Residuals
4: plot: QQplot of Residuals

Selection: 0

> riskmeasures(m1gpd,c(0.95,0.99))

<table>
<thead>
<tr>
<th>p</th>
<th>quantile</th>
<th>sfall</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1,]</td>
<td>0.950</td>
<td>0.02585941</td>
</tr>
<tr>
<td>[2,]</td>
<td>0.990</td>
<td>0.04745161</td>
</tr>
</tbody>
</table>