Classification of Financial Risk

1. Credit risk
2. Market risk
3. Operational risk

We start with the market risk, because

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

What is Value at Risk (VaR)?

- 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)

A formal definition:

- time period given: $\Delta t = \ell$
- loss in value: $L$
• CDF of the loss $F_\ell(x)$

• given (upper tail) probability: $p$

• VaR is defined as

$$p = Pr[L > \text{VaR}] = 1 - F_\ell(\text{VaR}).$$

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

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

and $F_\ell(.)$ is continuous. For discrete distribution, we have

$$x_q = \min\{x | P(L \leq x) \geq q\}.$$

Factors affect VaR:

1. the probability $p$.

2. the time horizon $\ell$.

3. the CDF $F_\ell(x)$. (or CDF of loss)

4. the mark-to-market value of the position.

Remark: Let $R(L)$ be the risk associated with loss $L$. From a theoretical point, $R(L)$ must possess the following basic properties:

1. Monotonicity: If $L_1 \leq L_2$ for all possible outcomes, then $R(L_1) \leq R(L_2)$. 
2. Sub-additivity: $R(L_1 + L_2) \leq R(L_1) + R(L_2)$ for any two portfolios.

3. Positive homogeneity: $R(hL) = hR(L)$, where $h > 0$.

4. Translation invariance: $R(L + a) = R(L) + a$, where $a$ is a positive real number.

The sub-additivity is associated with risk diversification. The equality holds when the two portfolios are perfectly positively correlated. A risk measure is called **coherent** if it satisfies the above four properties.

**Note.** If the loss involved is normally distributed, then VaR is a coherent risk measure. The sub-additivity can be seen because

\[
(\sigma_1 + \sigma_2)^2 = \sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2 \\
\geq \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2 \\
= \text{Var}(L_1 + L_2),
\]

where $\sigma_1$ and $\sigma_2$ are the standard errors of $L_1$ and $L_2$, respectively, and $\rho$ is the correlation between $L_1$ and $L_2$.

However, VaR fails to meet the sub-additive property under certain conditions. This is the reason that we shall also discuss expected shortfall or conditional VaR (CVaR), which is a coherent risk measure. Expected shortfall is the expected loss when the VaR is exceeded. Some people call expected shortfall as **Tail VaR** (TVaR) or expected tail loss (ETL). In the insurance literature, expected
shortfall is called *Conditional Tail Expectation* or *Tail Conditional Expectation* (TCE).

**Example** of incoherent VaR. See the book of Klugman, Panjer and Willmot (2008, Wiley). Let $Z$ denote a continuous loss random variable with the following CDF values

$$F_Z(1) = 0.91, \quad F_Z(90) = 0.95, \quad F_Z(100) = 0.96.$$  

It is clear that $\text{VaR}_{.95}(Z) = 90$. Now, define loss variables $X$ and $Y$ such that $Z = X + Y$, where

$$X = \begin{cases} 
Z, & \text{if } Z \leq 100 \\
0, & \text{if } Z > 100, 
\end{cases}$$

$$Y = \begin{cases} 
0, & \text{if } Z \leq 100 \\
Z, & \text{if } Z > 100. 
\end{cases}$$

The CDF of $X$ satisfies

$$F_X(1) = .91/.96 \approx 0.95, \quad F_X(90) = 0.95/.96 \approx 0.99, \quad F_X(100) = 1.$$  

Therefore, $\text{VaR}_{.95}(X) \approx 1$. Turn to $Y$. The CDF of $Y$ satisfies $F_Y(0) = 0.96$ so that $\text{VaR}_{.95}(Y) = 0$. Consequently,

$$\text{VaR}_{.95}(X) + \text{VaR}_{.95}(Y) = 1 < \text{VaR}_{.95}(Z).$$

In what follows, we shall use log returns in the analysis (simple returns can also be used).

Why use log returns?
log returns $\approx$ percentage changes.

VaR = Value \times (VaR of log return).

**Methods available for market risk**

1. RiskMetrics
2. Econometric modeling
3. Empirical quantile
4. Traditional extreme value theory (EVT)
5. EVT based on exceedance over a high threshold

Data used in illustrations:
Daily log returns of IBM stock
- span: July 3, 62 to Dec. 31, 98.
- size: 9190 points

Position: long on $10 million.

**Note:** For a long position, loss occurs at the left (or lower) tail of the returns. This is equivalent to using the right (or upper) tail if **negative** returns are used.

**RiskMetrics**
- Developed by J.P. Morgan
- $r_t$ given $F_{t-1}$: $N(0, \sigma_t^2)$
• $\sigma_t^2$ follows the special IGARCH(1,1) model

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

• VaR = $1.65\sigma_t$ if $p = 0.05$.

• $k$-horizon: VaR[$k$] = $\sqrt{k}$VaR

The square root of time rule

• Pros: simplicity and transparency

• Cons: model is not adequate

Example: IBM data

Model:

\begin{align*}
  r_t &= a_t, \quad a_t = \sigma_t \epsilon_t, \\
  \sigma_t^2 &= 0.9396 \sigma_{t-1}^2 + (1 - 0.9396) a_{t-1}^2
\end{align*}

Because $r_{9190} = -0.0128$ and $\hat{\sigma}_{9190}^2 = 0.0003472$,

$\hat{\sigma}_{9190}^2(1) = 0.000336$.

For $p = 0.05$, VaR of $r_t = -1.65 \times \sqrt{0.000336} = -0.03025$

$$\text{VaR} = 10,000,000 \times 0.03025 = 302,500.$$ 

For $p = 0.01$, VaR of $r_t = -2.3262 \times \sqrt{0.000336} = -0.04265$, and

$\text{VaR} = 426,500$.

**Expected shortfall.** From the prior discussion, VaR is simply the

$100(1 - p)$th quantile of the loss function, where $p$ is the upper tail
probability. When an extreme loss occurs, i.e. VaR is exceeded, the actual loss can be much higher than VaR. To better quantify the loss and to employ a coherent risk measure, we consider the expected loss once the VaR is exceeded. Let $q = 1 - p$. The expected shortfall (ES) is then

$$ES_q = E(L|L > \text{VaR}_q).$$

For standard Normal distribution, we have $ES_q = f(\text{VaR}_q)/p$, where $f(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}x^2\right]$, which is the probability density function of the standard normal distribution. Consequently, the ES for a normal return $N(0, \sigma_t^2)$ is

$$ES_q = \frac{f(\text{VaR}_q)}{p} \times \sigma_t,$$

where $\sigma_t$ is the volatility and $f(x)$ denotes the density function of $N(0,1)$. Therefore, for a $N(0, \sigma_t^2)$ loss function, we have $ES_{0.99} = 2.6652\sigma_t$.

In general, for a normal distribution $N(\mu, \sigma_t^2)$, the ES is

$$ES_q = \mu + \frac{f(\text{VaR}_q)}{p} \times \sigma_t.$$

Expected shortfall can also be defined as the average VaR for small tail probabilities, i.e.

$$ES_{1-p} = \frac{1}{p} \int_0^p \text{VaR}_{1-u} du.$$

**Econometric models**

- $r_t = \mu_t + a_t$ given $F_{t-1}$
• \( \mu_t \): a mean equation (Ch. 2)

• \( \sigma_t^2 \): a volatility model (Ch. 3 or 4)

• Pros: sound theory

• Cons: a bit complicated.

IBM data:

**Case 1:** Gaussian

\[
 r_t = 0.00066 - 0.0247r_{t-2} + a_t, \quad a_t = \sigma_t \epsilon_t
\]

\[
 \sigma_t^2 = 0.00000389 + 0.0799a_{t-1}^2 + 0.9073\sigma_t^2.
\]

From \( r_{9189} = -0.00201 \), \( r_{9190} = -0.0128 \) and \( \sigma_{9190}^2 = 0.00033455 \), we have

\[
 \hat{r}_{9190}(1) = 0.00071 \quad \text{and} \quad \hat{\sigma}_{9190}^2(1) = 0.0003211.
\]

If \( p = 0.05 \), then

\[
 0.00071 - 1.6449 \times \sqrt{0.0003211} = -0.02877.
\]

VA\(R = 10,000,000 \times 0.02877 = 287,700. \)

If \( p = 0.01 \), then the quantile is

\[
 0.00071 - 2.3262 \times \sqrt{0.0003211} = -0.0409738.
\]

VA\(R = 409,738. \)
**Expected shortfall** for normal distribution $N(\mu_t, \sigma_t^2)$. Following the prior discussion, we have

$$ES_q = \mu_t + \frac{f(VaR_q)}{p} \sigma_t.$$  

**Case 2**: Student-$t_5$

$$r_t = 0.0003 - 0.0335r_{t-2} + a_t, \quad a_t = \sigma_t \epsilon_t$$

$$\sigma_t^2 = 0.000003 + 0.0559a_{t-1}^2 + 0.9350\sigma_{t-1}^2.$$  

From the data, $r_{9189} = -0.00201$, $r_{9190} = -0.0128$ and $\sigma_{9190}^2 = 0.000349$, we have

$$\hat{r}_{9190}(1) = 0.000367 \quad \text{and} \quad \hat{\sigma}_{9190}^2(1) = 0.0003386.$$  

If $p = 0.05$, then the quantile is

$$0.000367 - (2.015/\sqrt{5/3})\sqrt{0.0003386} = 0.000367 - 1.5608\sqrt{0.0003386} = -0.028354.$$  

VaR = $10,000,000 \times 0.028352 = $283,520.

If $p = 0.01$, the quantile is

$$0.000367 - (3.3649/\sqrt{5/3})\sqrt{0.0003386} = -0.0475943.$$  

VaR = $475,943.

**Discussion:**

- Effects of heavy-tails seen with $p = 0.01$.
- Multiple step-ahead forecasts are needed.
**Example 7.3 (continued).** 15-day horizon.

\( \hat{r}_{9190}[15] = 0.00998 \) and \( \sigma_{t}[15] = 0.0047948 \).

If \( p = 0.05 \), the quantile is \( 0.00998 - 1.6449 \sqrt{0.0047948} = -0.1039191 \).

15-day VaR = $10,000,000 \times 0.1039191 = $1,039,191.

RiskMetrics: VaR = $287,700 \times \sqrt{15} = $1,114,257.

For standardized Student-\( t \) distribution with \( v \) degrees of freedom, the expected shortfall is given by

\[
ES_{1-p} = \frac{1}{p} f^*(x_q|v) \left( \frac{(v - 2) + x_q^2}{v - 1} \right),
\]

where \( x_q \) is the 100\( q \)th quantile of the distribution. If \( (Y - \mu)/\sigma \) follows a standardized Student-\( t \) distribution with \( v \) degrees of freedom, then

\[
ES_{1-p} = \mu + \frac{1}{p} f^*(x_q|v) \left( \frac{(v - 2) + x_q^2}{v - 1} \right) \sigma.
\]

**Empirical quantile**

Sample of log returns: \( \{ r_t | t = 1, \cdots, n \} \).

Order statistics:

\[
r_{(1)} \leq r_{(2)} \leq \cdots \leq r_{(n)}
\]

\( r_{(i)} \) as the \( i \)th order statistic of the sample.

\( r_{(1)} \) is the sample minimum

\( r_{(n)} \) the sample maximum.

Idea: Use the empirical quantile to estimate the theoretical quantile of \( r_t \).
For a given probability $p$, what is the empirical quantile?
If $np = \ell$ is an integer, then it is $r(\ell)$.
If $np$ is not an integer, find the two neighboring integers $\ell_1 < np < \ell_2$ and use interpolation.
The quantile is
\[
\hat{x}_p = \frac{p_2 - p}{p_2 - p_1} r(\ell_1) + \frac{p - p_1}{p_2 - p_1} r(\ell_2).
\]
IBM data:
n = 9190. If $p = 0.05$, then $np = 459.5$.
5% quantile is $(r(459) + r(460))/2 = -0.021603$.
VaR = $216,030$.
If $p = 0.01$, then $np = 91.9$ and the 1% quantile is
\[
\hat{x}_{0.01} = \frac{p_2 - 0.01}{p_2 - p_1} r(91) + \frac{0.01 - p_1}{p_2 - p_1} r(92)
= \frac{0.0001}{0.00011} (-3.658) + \frac{0.0001}{0.00011} (-3.657)
\approx -3.65709.
\]
VaR is $365,709$.

**Expected shortfall:**
\[
\text{ES}_q = \frac{1}{N_q} \sum_{i=1}^{n} x_{(i)} I[x_{(i)} > \hat{x}_q],
\]
where $N_q$ is the number of data that exceed the empirical quantile $\hat{x}_q$. In other words, $\text{ES}_q$ is the average of all the data that exceed the empirical 100$q$th quantile $\hat{x}_1$ of the data.
> nibm=-ibm
> quantile(nibm,c(.95,.99))
95%   99%
2.15855 3.63030 <= R uses a slightly different method, but the results
are close to those in the lecture note, which are 2.163 and 3.657,
respectively.

## Compute expected shortfall.
> idx=c(1:9190)[nibm > 2.15855]
> mean(nibm[idx])
[1] 3.172613
> idx=c(1:9190)[nibm > 3.6303]
> mean(nibm[idx])
[1] 5.097207

**Extreme value theory**: Focus on the tail behavior of $r_t$.

**Review of extreme value theory**

A properly normalized $r_{(n)}$ assumes a special distribution:

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

for $x < -1/\xi$ if $\xi < 0$ and for $x > -1/\xi$ if $\xi > 0$.

$\xi$: the shape parameter

$\alpha = 1/\xi$: tail index of the distribution.

Classification of distributions:

- Type I: $\xi = 0$, the Gumbel family. The CDF is
  
  $$F_*(x) = \exp[-\exp(x)], \quad -\infty < x < \infty. \quad (1)$$
• Type II: $\xi > 0$, the Fréchet family. The CDF is

$$F_*(x) = \begin{cases} 
\exp[-(1 + \xi x)^{-1/\xi}] & \text{if } x < -1/\xi \\
1 & \text{otherwise}.
\end{cases} \tag{2}$$

• Type III: $\xi < 0$, the Weibull family. The CDF here is

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

The probability density function (pdf) of the normalized minimum is

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

where $-\infty < x < \infty$ for $\xi = 0$, $x < -1/\xi$ for $\xi < 0$ and $x > -1/\xi$ for $\xi > 0$.

How to use the EVT distribution?

If we know the three parameters, we can compute the quantile!

**Empirical estimation**

Divide the sample into non-overlapping subsamples.

Suppose there are $T$ data points, we devide the data as

$$\{r_1, \ldots, r_n | r_{n+1}, \ldots, r_{2n} | r_{2n+1}, \ldots, r_{3n} | \ldots | r_{(g-1)n+1}, \ldots, r_{ng}\},$$

$n$: size of subgroup

Idea: find the minimum of each subgroup. These minima are the data used to estimate the three parameters.
Several estimation methods available. We use maximum likelihood estimates.

IBM data:

<table>
<thead>
<tr>
<th>n</th>
<th>g</th>
<th>Scale $\alpha_n$</th>
<th>Location $\beta_n$</th>
<th>Shape Par. $\xi_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Negative maximum returns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>437</td>
<td>0.823(0.035)</td>
<td>1.902(0.044)</td>
<td>0.197(0.036)</td>
</tr>
<tr>
<td>63</td>
<td>145</td>
<td>0.945(0.077)</td>
<td>2.583(0.090)</td>
<td>0.335(0.076)</td>
</tr>
<tr>
<td>126</td>
<td>72</td>
<td>1.147(0.131)</td>
<td>3.141(0.153)</td>
<td>0.330(0.101)</td>
</tr>
<tr>
<td>252</td>
<td>36</td>
<td>1.542(0.242)</td>
<td>3.761(0.285)</td>
<td>0.322(0.127)</td>
</tr>
<tr>
<td>(b) Maximal returns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>437</td>
<td>0.931(0.039)</td>
<td>2.184(0.050)</td>
<td>0.168(0.036)</td>
</tr>
<tr>
<td>63</td>
<td>145</td>
<td>1.157(0.087)</td>
<td>3.012(0.108)</td>
<td>0.217(0.066)</td>
</tr>
<tr>
<td>126</td>
<td>72</td>
<td>1.292(0.158)</td>
<td>3.471(0.181)</td>
<td>0.349(0.130)</td>
</tr>
<tr>
<td>252</td>
<td>36</td>
<td>1.624(0.271)</td>
<td>4.475(0.325)</td>
<td>0.264(0.186)</td>
</tr>
</tbody>
</table>

EVT to VaR: Use a two-step procedure, because of the division into subgroup.

VaR for $r_t$:

\[
\text{VaR} = \begin{cases} 
\beta_n - \frac{\alpha_n}{\xi_n} \left\{ 1 - \left[ -n \ln(1 - p) \right]^{-\xi_n} \right\} & \text{if } -\xi_n \neq 0 \\
\beta_n + \alpha_n \ln \left[ -n \ln(1 - p) \right] & \text{if } \xi_n = 0.
\end{cases}
\]

Let $r_{n,i}$ be the maximum of a subperiod of length $n$. Under the traditional EVT, $r_{n,i}$ follows a generalized extreme value distribution with parameter $(\xi, \sigma, \mu)$. 
What is the relationship between quantile of $r_{n,i}$ and the return $r_t$? Let $Q$ be a real number.

\[
P(r_{n,i} > Q) = 1 - P(r_{n,i} \leq Q)
\]
\[
= 1 - P(\text{all } r_t \text{ in the subperiod } \leq Q)
\]
\[
= 1 - \prod_{t=1}^{n} P(r_t \leq Q) \quad \text{(use independence)}
\]
\[
= 1 - \left[ P(r_t \leq Q) \right]^n \quad \text{(because of same distribution)}
\]

Consequently, let $p$ be a small upper tail probability of $r_t$ and $Q$ be the corresponding quantile. That is,

\[
P(r_t \leq Q) = 1 - p
\]

From the above equation, we have

\[
P(r_{n,i} > Q) = 1 - (1 - p)^n.
\]

Therefore,

\[
P(r_{n,i} \leq Q) = 1 - P(r_{n,i} > Q) = (1 - p)^n.
\]

This means that $Q$ is the $(1 - p)^n$-th quantile of the generalized extreme value distribution.

**Takeaway:** For a small probability $p$, compute $(1 - p)^n$, where $n$ is the length of subperiod, then VaR can be obtained by finding the $(1 - p)^n$-th quantile of the extreme value distribution.

For IBM data, if $n = 63$ (quarterly minima), then $\hat{\alpha}_n = 0.945$, $\hat{\beta}_n = -2.583$, and $\hat{\xi}_n = 0.335$. If $p = 0.01$, the VaR is

\[
\text{VaR} = 2.583 - \frac{0.945}{-0.335} \left[ 1 - [-63 \ln(1 - 0.01)]^{-0.335} \right]
\]
VaR is $304,969.
If $p = 0.05$, then VaR is $166,641.
For $n = 21$, the results are:

VaR = $340,013$ for $p = 0.01$;
VaR = $184,127$ for $p = 0.05$.

**Discussion:**

- Results depend on the choice of $n$
- VaR seems low, but it might be due to the choice of $p$.
  
  If $p = 0.001$, then
  
  VaR = $546,641$ for the Gaussian AR(2)-GARCH(1,1) model
  VaR = $666,590$ for the extreme value theory with $n = 21$.

Additional information on applying extreme value theory to value at risk calculation.

To traditional approach of EVT

**Return Level:** It is a risk measure based on the idea of subperiods.
The $g$ $n$-subperiod return level, $L_{n,g}$, is the level that is exceeded in one out of every $g$ subperiods of length $n$.

$$P(r_{n,i} < L_{n,g}) = \frac{1}{g},$$

where $n$ is the length of subperiod used in estimating the GEV model and $r_{n,i}$ denotes subperiod minimum. For sufficiently large $n$,

$$L_{n,g} = \beta_n + \frac{\alpha_n}{\xi_n} \left\{ -\ln(1 - 1/g) \right\}^{-\xi_n} - 1,$$
where the shape parameter $\xi_n \neq 0$.

For a short position, the return level is

$$L_{n,g} = \beta_n + \frac{\alpha_n}{\xi_n} \{ 1 - [\ln(1 - 1/g)]^{-1/\xi_n} \}.$$

**Summary** of IBM data:

Position = $10$ million.

If $p = 0.05$, then

1. $302,500$ for the RiskMetrics,
2. $287,200$ for an AR(2)-GARCH(1,1) model,
3. $283,520$ for an AR(2)-GARCH(1,1) with $t_5$
4. $216,030$ using the empirical quantile, and
5. $184,127$ for EVT with $n = 21$.

If $p = 0.01$, then

1. $426,500$ for the RiskMetrics,
2. $409,738$ for an AR(2)-GARCH(1,1) model,
3. $475,943$ for an AR(2)-GARCH(1,1) model with $t_5$
4. $365,800$ for empirical quantile, and
5. $340,013$ for EVT with $n = 21$.

If $p = 0.001$, then
1. $566,443 for the RiskMetrics,
2. $546,641 for an AR(2)-GARCH(1,1) model,
3. $836,341 for an AR(2)-GARCH(1,1) model with $t_5$
4. $780,712 for quantile, and
5. $666,590 for EVT with $n = 21$.

**Multi-period VaR with EVT**

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

where $\alpha$ is the tail index and $k$ is the shape parameter.

For IBM data with $p = 0.05$ and $n = 21$,

\[ \text{VaR}(30) = (30)^{0.197} \text{VaR} = 1.954 \times 184,127 = 359,841. \]
New approach to VaR
Based on Exceedances over a high threshold
Idea: frequency of big returns and their magnitudes are important.
Statistical theory:
Two-dimensional Poisson process
Two possible cases:
Homogeneous: parameters are fixed over time
Non-homogeneous case: parameters are time-varying, according to some explanatory variables.
IBM data: homogeneous model

<table>
<thead>
<tr>
<th>Thr.</th>
<th>Exc.</th>
<th>Shape Par. $\xi$</th>
<th>Log(Scale) $\ln(\alpha)$</th>
<th>Location $\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Original log returns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.0%</td>
<td>175</td>
<td>0.30697(0.09015)</td>
<td>0.30699(0.12380)</td>
<td>4.69204(0.19058)</td>
</tr>
<tr>
<td>2.5%</td>
<td>310</td>
<td>0.26418(0.06501)</td>
<td>0.31529(0.11277)</td>
<td>4.74062(0.18041)</td>
</tr>
<tr>
<td>2.0%</td>
<td>554</td>
<td>0.18751(0.04394)</td>
<td>0.27655(0.09867)</td>
<td>4.81003(0.17209)</td>
</tr>
<tr>
<td>(b) Removing the sample mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.0%</td>
<td>184</td>
<td>0.30516(0.08824)</td>
<td>0.30807(0.12395)</td>
<td>4.73804(0.19151)</td>
</tr>
<tr>
<td>2.5%</td>
<td>334</td>
<td>0.28179(0.06737)</td>
<td>0.31968(0.12065)</td>
<td>4.76808(0.18533)</td>
</tr>
<tr>
<td>2.0%</td>
<td>590</td>
<td>0.19260(0.04357)</td>
<td>0.27917(0.09913)</td>
<td>4.84859(0.17255)</td>
</tr>
</tbody>
</table>

VaR calculation:
$$
\text{VaR} = \begin{cases} 
\beta + \frac{\alpha}{-\xi} \left(1 - \left[-T \ln(1 - p)\right]^{-\xi}\right) & \text{if } -\xi \neq 0 \\
\beta + \alpha \ln\left[-T \ln(1 - p)\right] & \text{if } \xi = 0 
\end{cases}
$$
where $T = 252$, the number trading days in a year.

IBM data: VaR of 5% & 1%

• Case I: original returns
  1. $\eta = 3.0\%$: $228,239$ & $359.303$.
  2. $\eta = 2.5\%$: $219,106$ & $361,119$.
  3. $\eta = 2.0\%$: $212,981$ & $368.552$.

• Case II: remove sample mean
  1. $\eta = 3.0\%$: $232,094$ & $363,697$.
  2. $\eta = 2.5\%$: $225,782$ & $364,254$.

Peaks over the Threshold (POT)

**Generalized Pareto Distribution:** For simplicity, assume that the shape parameter $k \neq 0$. Consider the extreme value distribution of maximum (Eq. (7.16) of the textbook)

$$F_*(r) = \exp \left[ - \left( 1 - \frac{-\xi (r - \beta)}{\alpha} \right)^{-1/\xi} \right].$$

The distribution of $r \leq x + \eta$ given $\eta$, where $x \geq 0$, is

$$\Pr(r \leq x + \eta | r > \eta) \approx 1 - \left( 1 + \frac{\xi x}{\psi(\eta)} \right)^{-1/\xi},$$

where $\psi(\eta) = \alpha + \xi (\eta - \beta)$, which depends on $\eta$. 
The distribution with cumulative distribution function
\[ G(x) = 1 - \left[ 1 + \frac{\xi x}{\psi(\eta)} \right]^{-1/\xi}, \]
is called a generalized Pareto distribution (GPD).

**Selection of the high threshold**

**Mean Excess:** Given a high threshold \( \eta_o \), suppose the excess \( r - \eta_o \) follows a GPD with parameter \( \xi \) and \( \psi(\eta_o) \), where \( 0 > -\xi > -1 \).
Then the mean excess over the threshold is
\[ E(r - \eta_o|r > \eta_o) = \frac{\psi(\eta_o)}{1 - \xi}. \]
For any \( \eta > \eta_o \), the mean excess function is defined as
\[ e(\eta) = E(r - \eta|r > \eta) = \frac{\psi(\eta_o) + \xi(\eta - \eta_o)}{1 - \xi}. \]
The fact that, for a given \( \xi \), \( e(\eta) \) is a linear function of \( \eta \), where \( \eta > \eta_o \), provides a simple method to infer the threshold \( \eta_o \) for GPD.
Define the empirical mean excess as
\[ e_T(\eta) = \frac{1}{N_\eta} \sum_{i=1}^{N_\eta} (r_{t_i} - \eta), \]
where \( N_\eta \) is the number of returns that exceed \( \eta \) and \( r_{t_i} \) are the values of the corresponding returns.
The scatterplot \( e_T(\eta) \) versus \( \eta \) is called the mean excess plot, which should be linear for \( \eta > \eta_o \).
In R, the command is `meplot`.

**Use of GPD in VaR**
For a given threshold, estimate GPD to obtain parameters $k$ and $\psi(\eta)$. Check the adequacy of the fit; see demonstration. Provided that the model is adequate, the VaR can be computed by

$$\text{VaR}_q = \eta + \frac{\psi(\eta)}{-\xi} \left\{ 1 - \left[ \frac{T}{N_\eta} (1 - q) \right]^{-\xi} \right\},$$

where $q = 1 - p$ with $0 < p < 0.05$, $T$ is the sample size and $N_\eta$ is the number of exceedances.

Alternatively, one can use the formula in Eq. (7.36) of the textbook when one treats the exceedances and exceeding times as a two-dimensional Poisson process. The VaR results obtained are close.

**Expected Shortfall (ES):** the expected loss given that the VaR is exceeded. Specifically,

$$\text{ES}_q = E(r|r > \text{VaR}_q) = \text{VaR}_q + E(r - \text{VaR}_q|r > \text{VaR}_q).$$

For GPD, it turns out that

$$\text{ES}_q = \frac{\text{VaR}_q}{1 - \xi} + \frac{\psi(\eta) - \xi \eta}{1 - \xi}.$$

In **evir**, the command is **riskmeasures**.

Non-homogeneous case:

$$k_t = \gamma_0 + \gamma_1 x_{1t} + \cdots + \gamma_v x_{vt} \equiv \gamma_0 + \gamma'x_t$$

$$\ln(\alpha_t) = \delta_0 + \delta_1 x_{1t} + \cdots + \delta_v x_{vt} \equiv \delta_0 + \delta'x_t$$

$$\beta_t = \theta_0 + \theta_1 x_{1t} + \cdots + \theta_v x_{vt} \equiv \theta_0 + \theta'x_t.$$
For IBM data, explanatory variables include past volatilities, etc. See Chapter 7 for more details and estimation results.

Illustration:
For December 31, 1998, we have $x_{3,9190} = 0$, $x_{4,9190} = 0.9737$ and $x_{5,9190} = 1.9766$. The parameters become

$$
\xi_{9190} = 0.01195, \quad \ln(\alpha_{9190}) = 0.19331, \quad \beta_{9190} = 6.105.
$$

If $p = 0.05$, then quantile = 3.03756% and

$$
\text{VaR} = 10,000,000 \times 0.0303756 = 303,756.
$$

If $p = 0.01$, then VaR is $497,425$.

For December 30, 1998, we have $x_{3,9189} = 1$, $x_{4,9189} = 0.9737$ and $x_{5,9189} = 1.8757$ and

$$
\xi_{9189} = 0.2500, \quad \ln(\alpha_{9189}) = 0.52385, \quad \beta_{9189} = 5.8834.
$$

The 5% VaR becomes

$$
\text{VaR} = 10,000,000 \times 0.0269139 = 269,139.
$$

If $p = 0.01$, then VaR becomes $448,323$.

**R Demonstration:**
Use the library: `evir`. Download the library before using the commands.

The program`'evir'` uses maxima (right tail) so that one should use`'minus returns'` for the left tail.
(* pgev, dgev, qgev and rgev are commands for CDF, pdf, quantile and random draw of the generalized extreme value distribution *)
(* For example, to obtain the 95th quantile, use below *)
> qgev(0.95,xi=0.5,mu=0,sigma=1) % obtain quantile
[1] 6.830793


> library(evir)
> da=read.table("d-ibmln98.txt")
> ibm=da[,1]
> plot(ibm,type='l')
> qqnorm(ibm) % normal probability plot

> nibm=-ibm % Focus on the left tail.

> m1=gev(nibm,block=21) % fit gen. extreme value dist.
> m1
$n.all
[1] 9190
$n
[1] 438
$data
.....
$block
[1] 21

$par.est
   xi    sigma     mu
0.1956199 0.8239793 1.9031998
$par.ses
   xi    sigma     mu
$\varcov$

\[
\begin{bmatrix}
[1,] & [2,] & [3,] \\
[1,] & 1.263428e-03 & -2.782725e-05 & -0.0004338483 \\
[2,] & -2.782725e-05 & 1.208770e-03 & 0.0008475859 \\
[3,] & -4.338483e-04 & 8.475859e-04 & 0.0019480124 \\
\end{bmatrix}
\]

$\text{converged}$

[1] 0

$nllh.\text{final}$

[1] 654.3337

attr("class")

[1] "gev"

> names(m1)

[1] "n.all" "n" "data" "block" "par.ests" [6] "par.ses" "varcov" "converged" "nllh.\text{final}"

> m1$n  \ % numbers of monthly maximum

[1] 438

> ymax=m1$data

> hist(ymax)

> ysort=sort(ymax)

> plot(ysort,-log(-log(ppoints(ysort))),xlab='Monthly maximum')

  % Gumbel qq-plot

> plot(m1)  \ % Model checking plots

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

Selection: 1

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

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

> 1-pgev(max(ymax),xi=.196, mu=1.90, sigma=.824)
[1] 5.857486e-05 % Prob. that the drop will exceed the maximum.

> rlevel.gev(m1,k.blocks=36) %return level & its 95% conf. interval.

> source('evtVaR.R') % R script to calculate VaR using extreme value.
> evtVaR(.196,.824,1.903,21,.01)
[1] 3.401727

** Peak Over the Threshold approach (homogeneous case).

> meplot(nibm) % mean excess plot

> m2=gpd(nibm,threshold=2.5)
> names(m2)
[1] "n"      "data"      "threshold"   "p.less.thresh"
[5] "n.exceed" "method"    "par.est"     "par.se"   
[9] "varcov"  "information" "converged"  "nllh.final"
> m2$threshold
[1] 2.5
> m2$n.exceed
[1] 310

> m2 % Obtain all output
$n
[1] 9190
$data
......


$threshold
[1] 2.5
$p.less.thresh
[1] 0.9662677
$n.exceed
[1] 310
$method
[1] "ml"

$par.est
   xi beta
0.2641593 0.7786761

$par.ses
   xi beta
0.06659234 0.06714131

$varcov
      [,1]      [,2]
[1,] 0.004434540 -0.002614442
[2,] -0.002614442 0.004507955

$information
[1] "observed"
$converged
[1] 0
$nllh.final
[1] 314.375
attr("class")
[1] "gpd"

> plot(m2) # Model checking. Should see all plots.

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: 1
[1] "threshold = 2.5 xi = 0.264 scale = 0.779 location= 2.5"

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

> shape(nibm) % A plot showing the stability of the estimates.

> riskmeasures(m2,c(0.95,0.99)) % Compute VaR and expected shortfall.
    p quantile    sfall
[1,] 0.95 2.208932 3.162654
[2,] 0.99 3.616487 5.075507

Credit Risk:


Some techniques for credit risk measurement

1. Long-term credit rating (High to Low)
<table>
<thead>
<tr>
<th>S&amp;P</th>
<th>Moody</th>
<th>Fitch</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>Aaa</td>
<td>AAA</td>
</tr>
<tr>
<td>AA</td>
<td>Aa</td>
<td>AA</td>
</tr>
<tr>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>BBB</td>
<td>Baa</td>
<td>BBB</td>
</tr>
<tr>
<td>BB</td>
<td>Ba</td>
<td>BB</td>
</tr>
<tr>
<td>B</td>
<td>B</td>
<td>B</td>
</tr>
<tr>
<td>CCC</td>
<td>Caa</td>
<td>CCC</td>
</tr>
<tr>
<td>CC</td>
<td>Ca</td>
<td>CC</td>
</tr>
<tr>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
</tbody>
</table>

2. Credit quality over time (transition)

S&P One-year transition matrix

(Source: Standard & Poor’s, Feb. 1997)

<table>
<thead>
<tr>
<th>Ini. Rat.</th>
<th>Rating at year-end(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AAA</td>
</tr>
<tr>
<td>AAA</td>
<td>88.5</td>
</tr>
<tr>
<td>AA</td>
<td>0.76</td>
</tr>
<tr>
<td>A</td>
<td>0.08</td>
</tr>
<tr>
<td>BBB</td>
<td>0.03</td>
</tr>
<tr>
<td>BB</td>
<td>0.02</td>
</tr>
<tr>
<td>B</td>
<td>0.00</td>
</tr>
<tr>
<td>CCC</td>
<td>0.17</td>
</tr>
</tbody>
</table>
3. CreditMetrics (JP Morgan & other sponsors)

Now, there is a R package CreditMetrics available.

4. Altman Z score (Mainly is U.S.)

\[
Z = 3.3 \left( \frac{\text{Earnings before Interest and Taxes [EBIT]}}{\text{Total Assets}} \right) + 1.0 \left( \frac{\text{Sales}}{\text{Total Assets}} \right) + 0.6 \left( \frac{\text{Market Value of Equity}}{\text{Book Value of Debt}} \right) + 1.4 \left( \frac{\text{Retained Earnings}}{\text{Total Assets}} \right) + 1.2 \left( \frac{\text{Working Capital}}{\text{Total Assets}} \right)
\]

5. KMV Corporation’s credit risk model (now merged with Moody)

**CreditMetrics**: developed by J.P. Morgan and other sponsors in 1997.

Simply put, CreditMetrics addresses the question:

“How much will one lose on his loans and loan portfolios next year for a given confidence level?”

From the assessment of market risk, the current market value and its volatility of a financial position play an important role in VaR calculation. Application of VaR methodology to *nontradable loans* encounters some immediate problems:

1. The current market value of the loan is not directly observable, because most loans are not traded.

2. No time-series data available to estimate the volatility.
To overcome the difficulties, we make use of

1. Available data on a borrower’s credit rating
2. The probability that the rating will change over the next year (the rating transition matrix)
3. Recovery rates on defaulted loans
4. Credit spreads and yields in the bond (or loan) market.

**Example:** Consider a five-year fixed-rate loan of $100 million made at 6% annual interest, and the borrower is rated BBB.

**Note:** The numerical numbers used in this example are from Chapter 6 of the reference book cited above.

**Rating migration:** One-year transition probabilities for BBB-rated borrower

<table>
<thead>
<tr>
<th></th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>Defaulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>prob.</td>
<td>0.02</td>
<td>0.33</td>
<td>5.95</td>
<td>86.93</td>
<td>5.30</td>
<td>1.17</td>
<td>0.12</td>
<td>0.18</td>
</tr>
</tbody>
</table>

**Valuation** Rating change (upgrades and downgrades) will affect the required credit risk spreads or premiums on the loan’s remaining cash flows and, hence, the implied market value of the loan. Downgrade → credit spread premium rises → present value of the loan should fall.

Upgrade has the opposite effect.

return to the example. (after one-year and a credit rating change)

\[
P = 6 + \frac{6}{1 + r_{1,1} + s_1} + \frac{6}{(1 + r_{1,2} + s_2)^2} + \frac{6}{(1 + r_{1,3} + s_3)^3} + \frac{106}{(1 + r_{1,4} + s_4)^4},
\]
where $r_{1,i}$ are the risk-free rates on zero-coupon U.S. Treasury bonds expected to exist one year into the future and $s_i$ is the annual credit spread on loans of a particular rating class of 1-year, 2-year, 3-year and 4-year maturities (derived from observed spreads in the corporate bond market over Treasuries).

One-year forward zero curves plus credit spreads by credit rating category:

<table>
<thead>
<tr>
<th>Category</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>3.60</td>
<td>4.17</td>
<td>4.73</td>
<td>5.12</td>
</tr>
<tr>
<td>AA</td>
<td>3.65</td>
<td>4.22</td>
<td>4.78</td>
<td>5.17</td>
</tr>
<tr>
<td>A</td>
<td>3.72</td>
<td>4.32</td>
<td>4.93</td>
<td>5.32</td>
</tr>
<tr>
<td>BBB</td>
<td>4.10</td>
<td>4.67</td>
<td>5.25</td>
<td>5.63</td>
</tr>
<tr>
<td>BB</td>
<td>5.55</td>
<td>6.02</td>
<td>6.78</td>
<td>7.27</td>
</tr>
<tr>
<td>B</td>
<td>6.05</td>
<td>7.02</td>
<td>8.03</td>
<td>8.52</td>
</tr>
<tr>
<td>CCC</td>
<td>15.05</td>
<td>15.02</td>
<td>14.03</td>
<td>13.52</td>
</tr>
</tbody>
</table>

Suppose that, during the first year, the borrower gets upgraded from BBB to A. The present value of the loan is

$$P = 6 + \frac{6}{1.0372} + \frac{6}{(1.0432)^2} + \frac{6}{(1.0493)^3} + \frac{106}{(1.0532)^4} = 108.66.$$ 

Value of the loan at the end of Year 1, under different rating changes (including first-year coupon):

<table>
<thead>
<tr>
<th>Rating</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>Defaulty value</th>
</tr>
</thead>
<tbody>
<tr>
<td>value</td>
<td>109.37</td>
<td>109.19</td>
<td>108.66</td>
<td>107.55</td>
<td>102.02</td>
<td>98.10</td>
<td>83.64</td>
<td>51.13</td>
</tr>
</tbody>
</table>
Calculation of VaR

<table>
<thead>
<tr>
<th>Year-end Rating</th>
<th>Probability of State(%)</th>
<th>New Loan Value Plus Probability</th>
<th>Probability Weighted from</th>
<th>Difference of Value Mean ($)</th>
<th>Probability Weighted Difference Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0.02</td>
<td>109.37</td>
<td>0.02</td>
<td>2.28</td>
<td>0.0010</td>
</tr>
<tr>
<td>AA</td>
<td>0.33</td>
<td>109.19</td>
<td>0.36</td>
<td>2.10</td>
<td>0.0146</td>
</tr>
<tr>
<td>A</td>
<td>5.95</td>
<td>108.66</td>
<td>6.47</td>
<td>1.57</td>
<td>0.1474</td>
</tr>
<tr>
<td>BBB</td>
<td>86.93</td>
<td>107.55</td>
<td>93.49</td>
<td>0.46</td>
<td>0.1853</td>
</tr>
<tr>
<td>BB</td>
<td>5.30</td>
<td>102.02</td>
<td>5.41</td>
<td>(5.06)</td>
<td>1.3592</td>
</tr>
<tr>
<td>B</td>
<td>1.17</td>
<td>98.10</td>
<td>1.15</td>
<td>(8.99)</td>
<td>0.9446</td>
</tr>
<tr>
<td>CCC</td>
<td>0.12</td>
<td>83.64</td>
<td>1.10</td>
<td>(23.45)</td>
<td>0.6598</td>
</tr>
<tr>
<td>Default</td>
<td>0.18</td>
<td>51.13</td>
<td>0.09</td>
<td>(55.96)</td>
<td>5.6358</td>
</tr>
</tbody>
</table>

Form the table, the mean value of the loan is $107.09 (sum of the 4-th column). The variance of the value is 8.9477 (sum of the last column).

Consequently, the standard deviation is \( \sigma = \sqrt{8.9477} = 2.99 \).

If normal distribution of the loan value is used,

- 5% VaR: \( 1.65 \times \sigma = 4.93 \)
- 1% VaR: \( 2.33 \times \sigma = 6.97 \)

If actual distribution is used,
• 6.77% VaR: $107.09-102.02 = $5.07
• 1.47% VaR: $107.09-98.10 = $8.99
• 1% VaR: $107.09-92.29 = $14.80.

The 1% number 92.29 is obtained by interpolation as (1.47%, 98.10) and (0.3%, 83.64).