Financial Econometrics
Lecture 9: Value at Risk, Expected Shortfall & Risk Management

Prof Hamed Ghoddusi
2019

Introduction

- Classification of Financial Risks (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 ℓ 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 η 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 $c > 0, \eta(X + c) = \eta(X) + c$.

What is Value at Risk (VaR)?

- Simply put: An upper quantile of $L_t(\ell)$ 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 $L_t(\ell)$. [The subscript of t is omitted from F.]

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

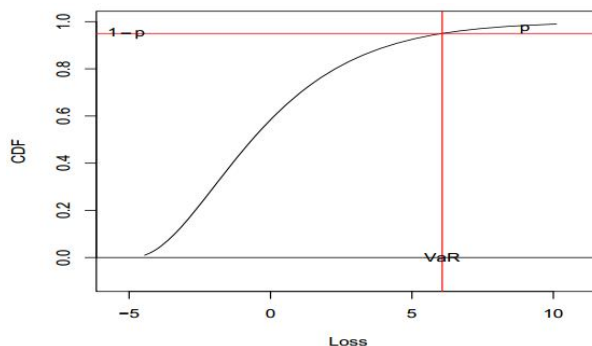 - $\Pr(L_t(\ell) \leq \text{VaR}_{1-p}) \geq 1 - p$ or $\Pr(L_t(\ell) > \text{VaR}_{1-p}) \leq p$.

Figure 1: Definition of Value at Risk (VaR) for a continuous loss random variable based on the cumulative distribution function.

What is Value at Risk (VaR)?

Quantile: x_q is the $100q$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 $x_q = \min\{x|P(L \leq x) \geq q\}$.

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 $Z_t = \frac{L_t - \mu_t}{\sigma_t}$ is standard normal and the VaR for L_t is

$$\text{VaR}_{1-p} = \mu_t + z_{1-p}\sigma_t, \tag{1}$$

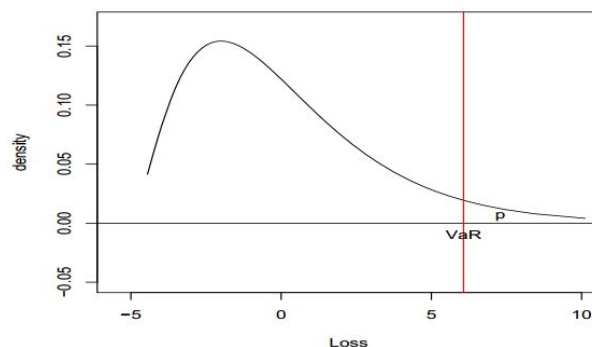where z_{1-p} denotes the $(1-p)$th quantile of $N(0,1)$.

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

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

$$\mathrm{VaR}_{1-p} = \mu_t + t_{1-p,v}\sigma_t, \qquad (2)$$

where $t_{1-p,v}$ is the $(1-p)$th quantile of t_v.

3 Standardized t_v distribution: Assume $v > 2$. If $Y_t = (L_t - \mu_t)/\sigma_t$ follows a Student-t distribution with v degrees of freedom, then

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

is a standardized Student-t variable with v degrees of freedom.

In this case, we have

$$\mathrm{VaR}_{1-p} = \mu_t + t^*_{1-p,v}\sigma_t\sqrt{v/(v-2)}, \qquad (3)$$

where $t^*_{1-p,v}$ is the $(1-p)$th quantile of standardized t_v.
If $Y_t = (L_t - \mu_t)/\sigma_t$ follows a standardized Student-t distribution with v degrees of freedom, then

$$\mathrm{VaR}_{1-p} = \mu_t + t^*_{1-p,v}\sigma_t. \qquad (4)$$

Second, VaR is coherent for a normally distributed loss. For instance, consider the subadditivity:

$$\begin{aligned}
\mathrm{Var}(X+Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2Cov(X,Y) \\
&= \sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y \\
&\leq \sigma_x^2 + \sigma_y^2 + 2\sigma_x\sigma_y \\
&= (\sigma_x + \sigma_y)^2,
\end{aligned}$$

where $\rho = \mathrm{cor}(X,Y)$. 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 $\mathrm{VaR}^x_{0.95} = 90$.

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 $F_\ell^1(X)$ of X_1 satisfies

$$\begin{aligned}
F_\ell^1(80) &= 0.9215/0.97 = 0.95 \\
F_\ell^1(90) &= 0.95/0.97 = 0.9794 \\
F_\ell^1(100) &= 0.97/0.97 = 1.
\end{aligned}$$

What is Value at Risk (VaR)?

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, $Pr(X_2 \leq 0) = P(X \leq 100) = 0.97$. 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.

What is Value at Risk (VaR)?

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

$$\text{ES}_{1-p} = E(X|X > \text{VaR}) = \frac{\int_{\text{VaR}}^{\infty} x f(x) dx}{Pr(X > \text{VaR})}. \tag{5}$$

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 (5) becomes

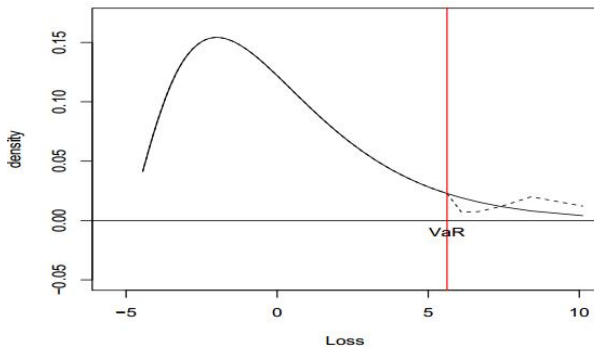$$\text{ES}_{1-p} = \frac{\int_{1-p}^{1} \text{VaR}_u du}{p}.$$

Figure 3: Density functions of two loss random variables that have the same VaR, but different loss implications.

What is Value at Risk (VaR)?

Thus, ES can be seen to average all 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: If $L_t \sim N(\mu_t, \sigma_t^2)$, then

$$\text{ES}_{1-p} = \mu_t + \frac{f(z_{1-p})}{p}\sigma_t, \tag{6}$$

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

What is Value at Risk (VaR)?

- Student-t_v loss distribution: If $(L_t - \mu_t)/\sigma_t$ follows a Student-t distribution with v degrees of freedom, then

$$\text{ES}_{1-p} = \mu_t + \sigma_t \frac{f_v(x_{1-p})}{p} \left(\frac{v + x_{1-p}^2}{v - 1} \right), \quad (7)$$

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

- Standardized Student-t_v loss. Assume $v > 2$. If $(L_t - \mu_t)/[\sqrt{v/(v-2)}\sigma_t]$ follows a standardized Student-t distribution with v degrees of freedom, then

$$\text{ES}_{1-p} = \mu_t + \sigma_t \sqrt{v/(v-2)} \frac{f_v^*(x_{1-p}^*)}{p} \left(\frac{(v-2) + [x_{1-p}^*]^2}{v-1} \right), \quad (8)$$

What is Value at Risk (VaR)?

where $f_v^*(x)$ is the pdf a Standardized Student-t_v and x_{1-p}^* is the $(1-p)$th quantile of $f_v^*(x)$. If $(L_t - \mu_t)/\sigma_t$ follows directly a standardized Student-t distribution with v degrees of freedom, then

$$\text{ES}_{1-p} = \mu_t + \sigma_t \frac{f_v^*(x_{1-p}^*)}{p} \left(\frac{(v-2) + [x_{1-p}^*]^2}{v-1} \right). \quad (9)$$

Calculation of VaR involves several factors:

- Tail probability $p : p = 0.01$ for risk management and $p = 0.001$ in stress testing.

- The time horizon ℓ : 1 day or 10 days for market risk and 1 year or 5 years for credit risk.

What is Value at Risk (VaR)?

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

- 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

$$x_t = \begin{cases} r_t, & \text{if the position is short,} \\ -r_t, & \text{if the position is long.} \end{cases} \quad (10)$$

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

Methods for calculating financial risk

- RiskMetrics

- Econometric modeling

- Quantile and quantile regression

- Extreme value theory: traditional & Peaks over Thresholds

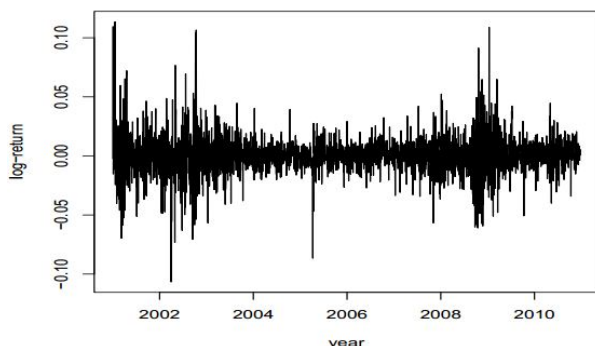- Serially correlated data: extremal index

Figure 4: Daily log returns of IBM stock from January 2, 2001 to December 31, 2010.

Methods for calculating financial risk

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 σ_t^2 follows the simple model:

$$\sigma_t^2 = \alpha \sigma_{t-1}^2 + (1-\alpha)x_{t-1}^2, \quad 1 > \alpha > 0. \tag{11}$$

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 α is often in the interval (0.9, 1) with a typical value of 0.94.

Methods for calculating financial risk

Example 2. Demonstration with IBM stock returns. We fit the special IGARCH(1,1) model of Eq. (11) to obtain an estimate of the parameter α 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,

$$\text{VaR}_{0.95} = \$1,000,000\frac{1.173}{100} = \$11,730, \quad \text{VaR}_{0.99} = \$16,590.$$

Finally, for 15 days holding period, we have

$$\text{VaR}_{0.95}(15) = \sqrt{15} \times \$11730 = \$45,430$$
$$\text{ES}_{0.95}(15) = \sqrt{15} \times \$14710 = \$56,972.$$

Methods for calculating financial risk

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.

Methods for calculating financial risk

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 ρ_{ij} is the cross-correlation coefficient between returns of the ith and jth instruments and VaR_i is the VaR of the ith instrument.

A simple derivation for the prior formula for two assets

Basic setup: Two assets with log returns r_{1t} and r_{2t}. The portfolio consists of w_1 and w_2 amounts invested in asset 1 and asset 2, respectively.

Under RiskMetrics, we have

$$r_{1t}|F_{t-1} \sim N(0, \sigma_{1t}^2), \quad \sigma_{1t}^2 = \beta\sigma_{1,t-1}^2 + (1-\beta)r_{1,t-1}^2$$
$$r_{2t}|F_{t-1} \sim N(0, \sigma_{2t}^2), \quad \sigma_{2t}^2 = \beta\sigma_{2,t-1}^2 + (1-\beta)r_{2,t-1}^2$$

VaR for the two assets are $w_1\text{VaR}_1$ and $w_2\text{VaR}_2$, respectively. For instance, for tail probability 0.05, we have VaR for the two assets as $1.645w_1\sigma_{1t}$ and $1.645w_2\sigma_{2t}$, respectively.

A simple derivation for the prior formula for two assets

Let p_t be the log return of the portfolio. Then, we have

$$p_t \approx wr_{1t} + (1-w)r_{2t},$$

where $w = \frac{w_1}{w_1+w_2}$ and $1-w = \frac{w_2}{w_1+w_2}$. The prior approximation becomes equality for simple returns.

Remark. $(w_1 + w_2)w = w_1$ and $(w_1 + w_2)(1-w) = w_2$.

Under RiskMetrics, we have

$$p_t|F_{t-1} \sim N(0, \sigma_{pt}^2),$$

where

$$\sigma_{pt}^2 = \text{Var}(r_{pt}|F_{t-1}) = w^2\sigma_{1t}^2 + (1-w)^2\sigma_{2t}^2 + 2w(1-w)\rho_t\sigma_{1t}\sigma_{2t}.$$

A simple derivation for the prior formula for two assets

The VaR for the portfolio is $(w_1 + w_2)\,\text{VaR}_p$. For tail probability of 0.05, we have $\text{VaR}_p = 1.645(w_1 + w_2)\sigma_{pt}$.
Therefore, the square of VaR for the portfolio with tail probability 0.05 is

$$
\begin{aligned}
(\text{VaR}_p)^2 &= (1.645)^2(w_1 + w_2)^2\sigma_{pt}^2 \\
&= (1.645)^2(w_1 + w_2)^2 \\
&\quad \times [w^2\sigma_{1t}^2 + (1-w)^2\sigma_{2t}^2 + 2w(1-w)\rho_t\sigma_{1t}\sigma_{2t}] \\
&= (1.645)^2[w_1^2\sigma_{1t}^2 + w_2^2\sigma_{2t}^2 + 2w_1w_2\rho_t\sigma_{1t}\sigma_{2t}] \\
&= \text{VaR}_1^2 + \text{VaR}_2^2 + 2\rho_t\text{VaR}_1\text{VaR}_2.
\end{aligned}
$$

This is exactly similar to

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\rho \times \text{std}(X)\text{std}(Y).$$

A simple derivation for the prior formula for two assets

Remarks:

1. The result can be generalized to more than two assets.

2. The formula continues to hold for expected shortfall provided that the mean returns of the two assets are zero. [Which is assumed under RiskMetrics.] If the means are not zero, then some adjustments are need for the portfolio.

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.

A simple derivation for the prior formula for two assets

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

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

A simple derivation for the prior formula for two assets

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

$$\text{VaR}_{0.95}^e = 0.01173 \times 0.6 = 0.00704, \quad \text{VaR}_{0.95}^b = 0.00705 \times 0.4 = 0.00282,$$

where the superscripts e and b denote equity and bond returns, respectively. The $\text{VaR}_{0.95}$ 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.

A simple derivation for the prior formula for two assets

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
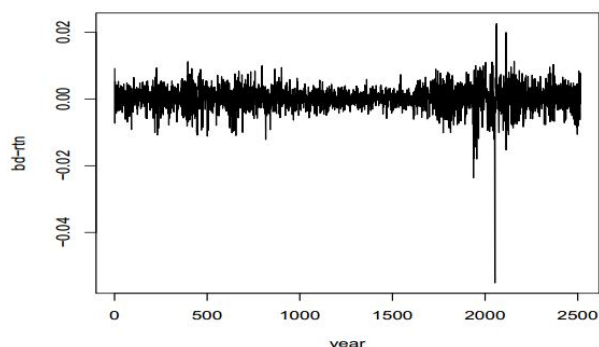
Figure 5: Daily log returns of bond index from January 2, 2001 to December 31, 2010. The bond index is the Bank of America Merrill Lynch U.S. Corp AAA total return index.

A simple derivation for the prior formula for two assets

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

$$x_t = -6.01 \times 10^{-4} + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0,1)$$
$$\sigma_t^2 = 4.378 \times 10^{-6} + 0.101 a_{t-1}^2 + 0.884 \sigma_{t-1}^2.$$

The 1-step ahead predictions at $T = 2515$ are -6.01×10^{-4} and 7.82×10^{-3}, respectively, for mean and volatility.

A simple derivation for the prior formula for two assets

Consequently, we have

$$\text{VaR}_{0.95} = 0.01227, \quad \text{ES}_{0.95} = 0.01554,$$
$$\text{VaR}_{0.99} = 0.01760, \quad \text{ES}_{0.99} = 0.02025.$$

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 innovations. The fitted model is

$$x_t = -4.113 \times 10^{-4} + a_t, \quad a_t = \sigma_t \epsilon_t, \quad \epsilon_t \sim t_{5.751}^*$$
$$\sigma_t^2 = 1.922 \times 10^{-6} + 0.0645 a_{t-1}^2 + 0.9286 \sigma_{t-1}^2.$$

The 1-step ahead predictions of the model for the mean and volatility at $T = 2515$ are -4.113×10^{-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

$$\hat{x}_{0.95} = 0.75 x_{(2389)} + 0.25 x_{(2390)} = 0.02654,$$

$x_{(i)}$ is the ith 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 $\widehat{ES}_{0.95} = \$39,949$ 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

$$\hat{x}_q|F_{t-1} \equiv \inf\{\beta_o'z|R_q(\beta_o) = \min\}, \tag{12}$$

where "$R_q(\beta_o) = \min$" means that β_o is obtained by

$$\beta_o = \operatorname{argmin}_\beta \sum_{t=1}^n w_q(x_t - \beta'z_t),$$

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

$$Q(q|z_t) = \sum_{t=2}^{2515} w_q(x_t - \beta_0 - \beta_1 s_{t-1} - \beta_2 v_{t-1}), \tag{13}$$

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.
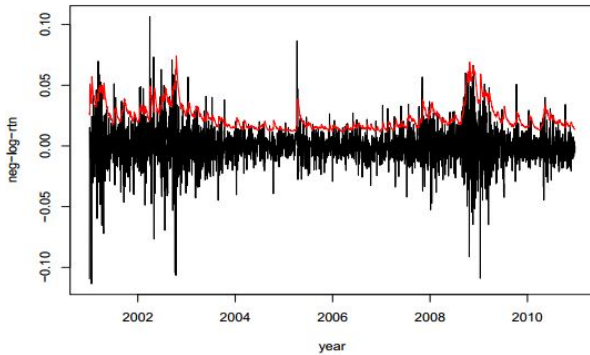
Figure 6: Time plot of the negative daily log returns of IBM stock from January 3, 2001 to December 31, 2010. The upper line shows the 95th quantiles obtained by the quantile regression of Equation (13)

Applying the quantile regression in (13) with $q = 0.95$, we obtain $\hat{\beta}_0 = -0.001(0.003)$, $\hat{\beta}_1 = 1.17724(0.22268)$, $\hat{\beta}_2 = 0.02809(0.01615)$, 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 $\widehat{Q}(0.95|z_{2515}) = 0.013385$. This implies that VaR$_{0.95}$ = \$13,385 for the financial position.

Figure 6 shows the negative IBM log returns $x_t = -r_t$ and 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

$$F_*(x) = \begin{cases} \exp[-(1 + \xi x)^{-1/\xi}] & \text{if } \xi \neq 0 \\ \exp[-\exp(-x)] & \text{if } \xi = 0 \end{cases} \quad (14)$$

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

The parameter ξ 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.

Methods based on extreme value theory

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

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

$$F_*(x) = \exp[-\exp(-x)], \quad -\infty < x < \infty. \quad (15)$$

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

$$F_*(x) = \begin{cases} \exp[-(1 + \xi x)^{-1/\xi}] & \text{if } x > -1/\xi, \\ 0 & \text{otherwise.} \end{cases} \quad (16)$$

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

$$F_*(x) = \begin{cases} \exp[-(1 + \xi x)^{-1/\xi}] & \text{if } x < -1/\xi, \\ 1 & \text{otherwise.} \end{cases}$$
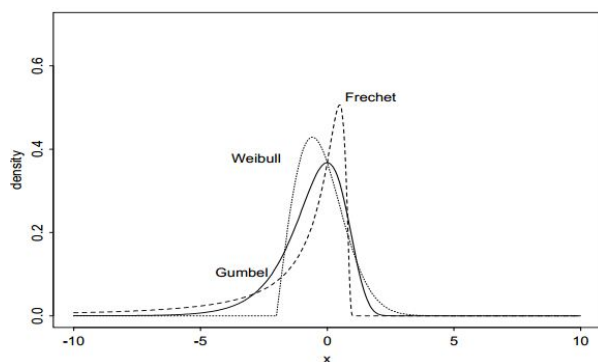
Figure 7: Probability density functions of extreme value distributions for normalized maximum. The solid line is for a Gumbel distribution, the dotted line is for the Weibull distribution with $\xi = -0.5$, and the dashed line is for the Fréchet distribution with $\xi = 0.9$.

Methods based on extreme value theory

Two important implications of EVT.

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 ξ 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
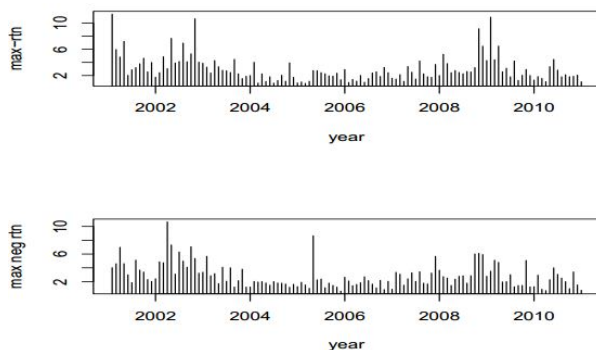
2. Maximum Likelihood Method

Figure 8: Block maximum of daily log returns of IBM stock, in percentages, when the subperiod 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.

Methods based on extreme value theory

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.

q	110	130	150
r_t	0.380(0.036)	0.399(0.035)	0.398(0.032)
$-r_t$	0.356(0.034)	0.383(0.034)	0.405(0.033)

⑤ The Nonparametric Approach: shape parameter ξ Hill estimator or Pickands estimator.

Methods based on extreme value theory

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 position with loss variable x_t is

$$\text{VaR} = \begin{cases} \mu_n - \frac{\sigma_n}{\xi_n}\{1 - [-n\ln(1-p)^{-\xi_n}]\} & \text{if} \quad \xi_n \neq 0 \\ \mu_n - \sigma_n \ln[-n\ln(1-p)] & \text{if} \quad \xi_n = 0, \end{cases} \quad (17)$$
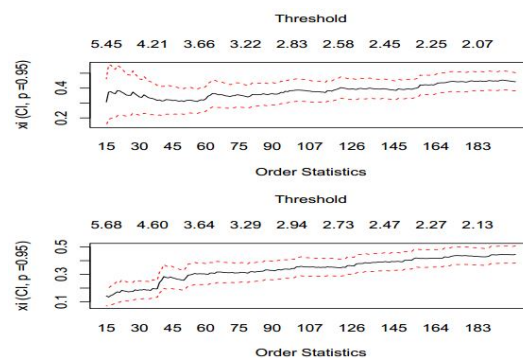
where n is the length of sub-periods.

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.

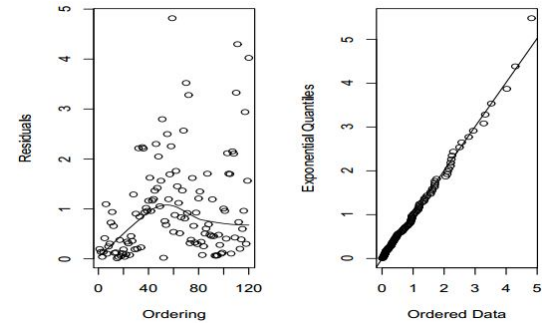| Length of sub-period	Shape Par. ξ	Scale σ	Location μ
Maximal positive returns			
1 mon.$(n = 21, g = 120)$	0.278(0.087)	1.046(0.092)	2.046(0.111)
2 mon.$(n = 42, g = 60)$	0.315(0.109)	1.168(0.145)	2.622(0.170)
Maximal negative returns			
1 mon.$(n = 21, g = 120)$	0.251(0.088)	1.029(0.090)	1.966(0.109)
2 mon.$(n = 42, g = 60)$	0.287(0.142)	1.100(0.143)	2.489(0.170)

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.

Methods based on extreme value theory

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 μ_n, σ_n, and ξ_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. (17) to calculate VaR.

Methods based on extreme value theory

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

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

Thus, for negative daily log returns of the stock, the upper 1% quantile is 1.8902%.

Methods based on extreme value theory

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

$$\text{VaR} = 2.489 - \frac{1.1}{0.287}\{1 - [-42\ln(1 - 0.01)]^{-0.287}\} = 3.5655.$$

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.

Methods based on extreme value theory

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.

Methods based on extreme value theory

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 $n = 21$.

Methods based on extreme value theory

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. (14).
The conditional distribution of $x \leq y + \eta$ given $x > \eta$ is

$$\begin{aligned}\Pr(x \leq y + \eta | x > \eta) &= \frac{\Pr(\eta \leq x \leq y + \eta)}{\Pr(x > \eta)} \\ &= \frac{\Pr(x \leq y + \eta) - \Pr(x \leq \eta)}{1 - \Pr(x \leq \eta)}. \quad (18)\end{aligned}$$

Using the CDF $F_*(.)$ of Eq. (14) and the approximation $e^{-z} \approx 1 - z$ and after some algebra, we obtain that

Methods based on extreme value theory

$$\Pr(x \le y + \eta | x > \eta) = \frac{F_*(y + \eta) - F_*(\eta)}{1 - F_*(\eta)}$$

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

$$\approx 1 - \left(1 + \frac{\xi y}{\sigma + \xi(\eta - \mu)}\right)^{-1/\xi}, \qquad (19)$$

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

$$\Pr(x \le y + \eta | x > \eta) \approx 1 - \exp(-y/\sigma).$$

Methods based on extreme value theory

Generalized Pareto distribution:
The probability distribution with cumulative distribution function

$$G_{\xi,\psi(\eta)}(y) = \begin{cases} 1 - \left[1 + \frac{\xi y}{\psi(\eta)}\right]^{-1/\xi} & \text{for} \quad \xi \ne 0, \\ 1 - \exp[-y/\psi(\eta)] & \text{for} \quad \xi = 0, \end{cases} \qquad (20)$$

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

Mean Excess Function

$$e_T(\eta) = \frac{1}{N_\eta} \sum_{i=1}^{N_\eta} (x_{t_i} - \eta), \qquad (21)$$

where N_η is the number of returns that exceed η 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

$$\text{VaR}_{0.95} = \$26,115, \quad \text{ES}_{0.95} = \$39,603.$$

Mean Excess Function

Finally, for threshold of 0.8%, we have

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

Thr. η	n.exceed	Shape ξ	Scale σ	Location μ	$\psi(\eta)$
1	504	0.107(0.042)	0.009(0.001)	-0.006(0.001)	0.011
1.2	410	0.075(0.044)	0.010(0.001)	-0.007(0.002)	0.011
0.8	610	0.106(0.039)	0.009(0.001)	-0.006(0.001)	0.010

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.

An Alternative Parameterization: Use $\psi(\eta) = \alpha + \xi(\eta - \beta)$. The parameter of GPD becomes $(\xi, \psi(\eta))$.
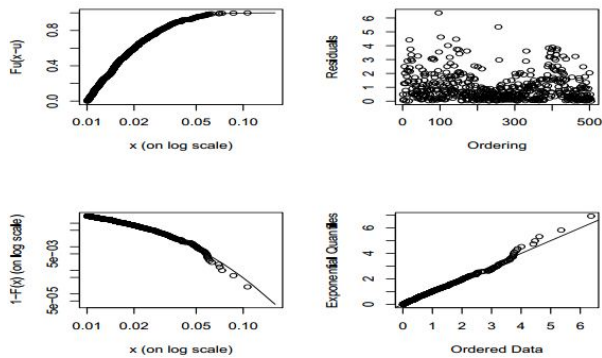
Figure 11: Diagnostic plots for GPD fit to the daily negative log returns of IBM stock from January 2, 2001 to December 31, 2010.

Mean Excess Function

Credit Risk:

Reference: *Credit Risk Measurement* :
New Approaches to Value at Risk and Other Paradigms, 2nd Edition, by Anthony Saunders and Linda Allen, Wiley, 2002.
Some techniques for credit risk measurement:

1. Long-term credit rating (High to Low)

S&P	Moody	Fitch
AAA	Aaa	AAA
AA	Aa	AA
A	A	A
BBB	Baa	BBB
BB	Ba	BB
B	B	B
CCC	Caa	CCC
CC	Ca	CC
C	C	C
D	D	D

Mean Excess Function

2. Credit quality over time (transition)

S&P One-year transition matrix
(Source: Standard & Poor's, Feb. 1997)

Ini.	Rating at year-end(%)							
Rat.	AAA	AA	A	BBB	BB	B	CCC	Default
AAA	88.5	8.05	0.72	0.06	0.11	0.00	0.00	0.00
AA	0.76	88.3	7.47	0.56	0.05	0.13	0.02	0.00
A	0.08	2.32	87.6	5.02	0.65	0.22	0.01	0.05
BBB	0.03	0.29	5.54	82.5	4.68	1.02	0.11	0.17
BB	0.02	0.11	0.58	7.01	73.8	7.4	0.89	0.98
B	0.00	0.09	0.21	0.39	5.98	72.8	3.42	4.92
CCC	0.17	0.00	0.34	1.02	2.20	9.64	53.13	19.21

Mean Excess Function

3. CreditMetrics (JP Morgan & other sponsors)
   Now, there is a R package **CreditMetrics** available.

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

   $Z = 3.3$(Earnings before Interest and Taxes [EBIT]/Totl Assets)
   $+ 1.0$(Sales/Total Assets)
   $+ 0.6$(Market Value of Equity/Book Value of Debt)
   $+ 1.4$(Retained Earnings/Total Assets)
   $+ 1.2$(Working Capital/Total Assets)

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

Mean Excess Function

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 imporant role in VaR calculation. Application of VaR methodology to *nontrabable 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.

Mean Excess Function

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 deafulted 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.

Mean Excess Function

Rating migration: One-year transition probabilities for BBBrated borrower

AAA	AA	A	BBB	BB	B	CCC	Default
0.02	0.33	5.95	86.93	5.30	1.17	0.12	0.18

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 $\rightarrow$ credit spread premium rises $\rightarrow$ present value of the loan should fall.
Upgrade has the opposite effect.
return to the example. (after one-year and a credit rating change)

$$P = 6 + \frac{6}{1 + r_{1,1} + s_1} + \frac{6}{(1 + r_{1,2} + s_2)^2} + \frac{6}{(1 + r_{1,3} + s_3)^3} + \frac{106}{(1 + r_{1,4} + s_4)^4},$$

Mean Excess Function

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:

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

Mean Excess Function

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):

Rating	AAA	AA	A	BBB	BB	B	CCC	Default
value	109.37	109.19	108.66	107.55	102.02	98.10	83.64	51.13

Calculation of VaR

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

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).

Calculation of VaR

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).