Lecture 10. Value at Risk

1. Risk Measure

Financial risk is classified into market risk, credit risk, and operational risk. The market risk is concerned with loss arising from unexpected changes in stock prices, interest rates, foreign exchange rates, and commodity prices.

- The loss of a position for a holding period is random, so that all inferences concerning the loss are based on the distribution of the loss random variable. Due to the difficulty of estimating the loss distribution, one employs summary statistics to quantify it in real applications, and a *risk measure* is one of these summary statistics.

Example 1.1. One buys 100 shares of stock A at \$50 per share today. The loss of the position for the next trading day is $X = 100(Y - 50)$, where Y is the tomorrow's stock price. Since Y is unknown today, the loss X for the next trading day is random.

Remark 1.2. The credit risk occurs when a borrower fails to make a payment as promised. The operational risk results from inadequate internal processes, people and systems, or external events.

1.1. Value at Risk

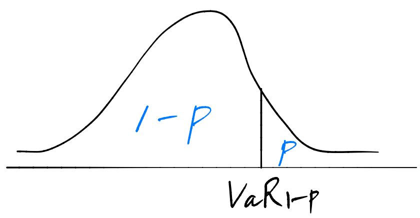
Let $x_{t+1\to t+s}$ be the loss of a position at time t for the next s periods from time $t+1$ and time $t+s$. Let v_t be the value of the position at time t. The loss $x_{t+1\to t+s}$ is either a positive or negative function of $v_{t+s} - v_{t+1}$, depending on the position being short or long.

Definition 1.3. As the most well-known measure of market risk, the *value at risk* (VaR) of a financial position from time $t+1$ to time $t+s$ with given probability p is defined as

loss
long $V_{t+s} - V_{t+1} < 0$
short $V_{t+s} - V_{t+1} > 0$

$$\mathrm{VaR}_{1-p} = \inf\{x_{t+1\to t+s} | F(x_{t+1\to t+s}|I_t) \geq 1-p\}, \tag{1.1}$$

where inf denotes the smallest real number $x_{t+1\to t+s}$ satisfying the condition and $F(\cdot|I_t)$ denotes the conditional cumulative distribution function (cdf) of $x_{t+1\to t+s}$ on I_t.



- From (1.1), one knows $F(\mathrm{VaR}_{1-p}|I_t) \geq 1-p$, which in turn means

$$\Pr(x_{t+1\to t+s} \leq \mathrm{VaR}_{1-p}|I_t) \geq 1-p$$

or

$$\Pr(x_{t+1\to t+s} \geq \mathrm{VaR}_{1-p}|I_t) \leq p.$$

$F(VaR_{1-p}|I_t) \geq 1-p$
$\Rightarrow Pr(x_{t+1\to t+s} \leq VaR_{1-p}|I_t) \geq 1-p$

So, the conditional probability that the loss in dollars is less than or equal to VaR_{1-p} is greater than or equal to $1-p$; alternatively, the conditional probability that the loss in dollars is greater than or equal to VaR_{1-p} is less than or equal to p.

Definition 1.4. For a cdf $F(x)$ and a given probability q satisfying $0 < q < 1$, the quantity

$$x_q = \inf\{x|F(x) \geq q\}$$

is called the *qth quantile* of $F(x)$.

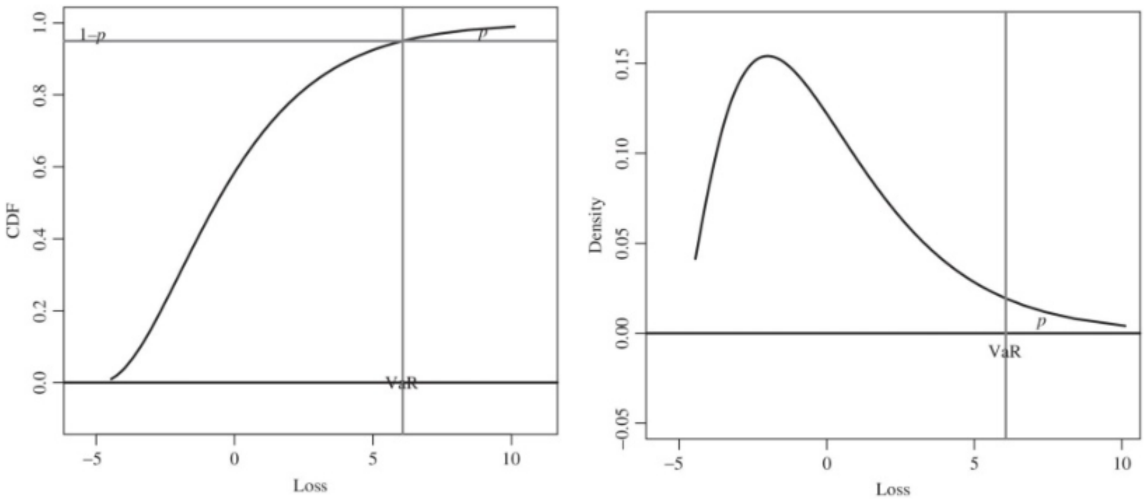
- From (1.1), VaR_{1-p} is the $(1-p)$th quantile of the loss distribution $F(\cdot|I_t)$, where p is a tail probability: i.e.,

$$\int_{-\infty}^{\text{VaR}_{1-p}} f(x_{t+1\to t+s}|I_t)dx_{t+1\to t+s} = 1-p \tag{1.2}$$

or equivalently

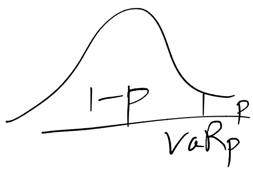
$$\int_{\text{VaR}_{1-p}}^{\infty} f(x_{t+1\to t+s}|I_t)dx_{t+1\to t+s} = p,$$

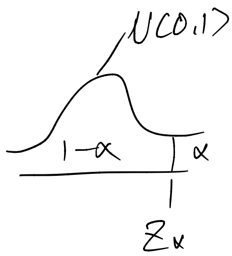
where $f(\cdot|I_t)$ denotes the conditional probability density function (pdf) of $x_{t+1\to t+s}$ on I_t.



Normal distribution. If $x_{t+1\to t+s}|I_t \sim N(E[x_{t+1\to t+s}|I_t], Var[x_{t+1\to t+s}|I_t])$, then (1.2) implies

$$\begin{aligned} 1-p &= \Pr(x_{t+1\to t+s} \le \text{VaR}_{1-p}|I_t) \\ &= \Pr\left(\frac{x_{t+1\to t+s} - E[x_{t+1\to t+s}|I_t]}{\sqrt{Var[x_{t+1\to t+s}|I_t]}} \le \frac{\text{VaR}_{1-p} - E[x_{t+1\to t+s}|I_t]}{\sqrt{Var[x_{t+1\to t+s}|I_t]}} \middle| I_t\right) \\ &= \Pr\left(Z \le \frac{\text{VaR}_{1-p} - E[x_{t+1\to t+s}|I_t]}{\sqrt{Var[x_{t+1\to t+s}|I_t]}} \middle| I_t\right). \end{aligned}$$





Since the quantity Z follows a standard normal distribution, one sees

$$\frac{\text{VaR}_{1-p} - E[x_{t+1\to t+s}|I_t]}{\sqrt{Var[x_{t+1\to t+s}|I_t]}} = z_{1-p},$$

Zp 가 맞다.

where z_{1-p} denotes the $(1-p)$th quantile of a standard normal distribution. Finally, VaR_{1-p} for the next s periods is given by

$$\text{VaR}_{1-p} = E[x_{t+1\to t+s}|I_t] + z_{1-p}\sqrt{Var[x_{t+1\to t+s}|I_t]}.$$

Remark 1.5. Denote the s-period log return from time $t+1$ to time $t+s$ by $r_{t+1\to t+s}$. The s-period log return approximates the "percentage change" in value of a financial asset from time $t+1$ and time $t+s$: i.e.,

In this class

$$r_{t+1\to t+s} = \ln(P_{t+s}) - \ln(P_{t+1}) \approx \frac{P_{t+s} - P_{t+1}}{P_{t+1}}.$$

$1-p = P(X \le z_{1-p})$

→ (1-p)th quantile of s.N

So, the loss of a long position for the next s periods in "percentage" occurs when $r_{t+1\to t+s}$ is negative, while the loss of a short position for the next s periods in "percentage" occurs when $r_{t+1\to t+s}$ is positive; that is, the loss in percentage can be represented as

$$x_{t+1\to t+s} \text{ in } \% = \begin{cases} -r_{t+1\to t+s} & \text{if the position is long} \\ r_{t+1\to t+s} & \text{if the position is short.} \end{cases}$$

Provided that the conditional distribution of $r_{t+1\to t+s}$ on I_t is known, one can compute VaR_{1-p} in %. Consequently, VaR_{1-p}, defined in dollars in (1.1), is computed as the cash value of the position times VaR_{1-p} in %.

Remark 1.6. Since VaR_{1-p} is a quantile with upper tail probability p, it does not completely describe the distribution of the loss random variable $x_{t+1\to t+s}$. The actual loss, if it occurs, can be greater than VaR_{1-p}, and as a result, VaR_{1-p} may underestimate the actual loss. To overcome this drawback, *expected shortfall* (ES) is introduced.

1.2. Expected Shortfall

Definition 1.7. *Expected shortfall* (ES) is defined as

$$\mathrm{ES}_{1-p} = E[x_{t+1\to t+s} | x_{t+1\to t+s} > \mathrm{VaR}_{1-p}]. \tag{1.3}$$

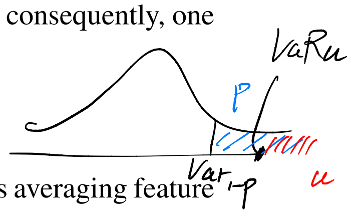
Put differently, ES_{1-p} is the expected loss of a position given that the loss exceeds VaR_{1-p}.

- One writes (1.3) as

$f(x|y) = \frac{f(x,y)}{f(y)}$

$$\begin{aligned} \mathrm{ES}_{1-p} &= \int_{-\infty}^{\infty} x_{t+1\to t+s} f(x_{t+1\to t+s} | x_{t+1\to t+s} > \mathrm{VaR}_{1-p}, I_t) dx_{t+1\to t+s} \\ &= \frac{\int_{-\infty}^{\infty} x_{t+1\to t+s} f(x_{t+1\to t+s}, x_{t+1\to t+s} > \mathrm{VaR}_{1-p} | I_t) dx_{t+1\to t+s}}{\Pr(x_{t+1\to t+s} > \mathrm{VaR}_{1-p} | I_t)} \\ &= \frac{\int_{\mathrm{VaR}_{1-p}}^{\infty} x_{t+1\to t+s} f(x_{t+1\to t+s} | I_t) dx_{t+1\to t+s}}{\Pr(x_{t+1\to t+s} > \mathrm{VaR}_{1-p} | I_t)} \end{aligned} \tag{1.4}$$

- Let $u = F(x_{t+1\to t+s}|I_t)$ for $\mathrm{VaR}_{1-p} \le x_{t+1\to t+s} \le \infty$. Then, one obtains that $F(\mathrm{VaR}_{1-p}|I_t) = 1-p$, $F(\infty|I_t) = 1$, $du = f(x_{t+1\to t+s}|I_t)dx_{t+1\to t+s}$, and $x_{t+1\to t+s} = F^{-1}(u|I_t) = \mathrm{VaR}_u$; consequently, one writes (1.4) as



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

which in turn means that ES_{1-p} is the average of all VaR_u for $1-p \le u \le 1$. This averaging feature enables ES_{1-p} to better reflect the tail behavior of the loss random variable.

Theorem 1.8. *For $X \sim N(\mu, \sigma^2)$, it shows*

truncated?

$$E[X|X > a] = \mu + \frac{\phi(\alpha)}{1-\Phi(\alpha)}\sigma,$$

ES_{1-p} should be greater than VaR_{1-p}.

where a is a constant, $\alpha = (a-\mu)/\sigma$, $\phi(\cdot)$ is a standard normal pdf, and $\Phi(\cdot)$ is a standard normal cdf.

In this class,

Def. $1-p = P(X \le \underline{z_{1-p}})$.

$100(1-p)$th quantile of $N(0,1)$.

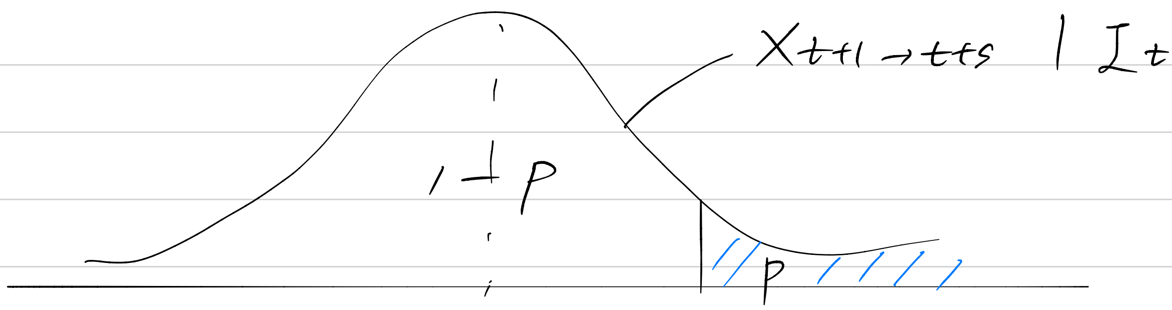
VaR in % $= 0.001$ $\quad -r_{t+1\to t+s} \mid I_t$

$x_{t+1\to t+s} \mid I_t$ $\left\langle \begin{array}{ll} -r_{t+1\to t+s} & \text{long} \\ r_{t+1\to t+s} & \text{short} \end{array} \right.$

\$ 1,000,000

VaR in \$

$= 0.001 \times 1{,}000{,}000$



$$ES_{1-p} = E[x_{t+1\to t+s} \mid x_{t+1\to t+s} > VaR_{1-p}]$$
$$= \int_{-\infty}^{\infty} x_{t+1\to t+s} f(x_{t+1\to t+s} \mid x_{t+1\to t+s} > VaR_{1-p}, I_t)\, dx_{t+1\to t+s}$$
$$= \frac{\int_{-\infty}^{\infty} x_{t+1\to t+s} f(x_{t+1\to t+s}, x_{t+1\to t+s} > VaR_{1-p} \mid I_t)\, dx_{t+1\to t+s}}{Pr(x_{t+1\to t+s} > VaR_{1-p})}$$
$$= \frac{\int_{VaR_{1-p}}^{\infty} x_{t+1\to t+s} f(x_{t+1\to t+s} \mid I_t)\, dx_{t+1\to t+s}}{Pr(x_{t+1\to t+s} > VaR_{1-p}}.$$

For $VaR_{1-p} \le x_{t+1\to t+s} \le \infty$,

let $u = F(x_{t+1\to t+s} \mid I_t)$.

$F(VaR_{1-p} \mid I_t) = 1-p$.

$F(\infty \mid I_t) = 1$

$du = f(x_{t+1\to t+s} \mid I_t)\, dx_{t+1\to t+s}$

and $x_{t+1\to t+s} = F^{-1}(u \mid I_t) = VaR_u$

consequently, $ES_{1-p} = \dfrac{\int_{1-p}^{1} VaR_u \, du}{p}$

ES_{1-p} is the average of all VaR_u for $1-p \le u \le 1$

$X \sim N(\mu, \sigma^2)$.

$$E[X \mid X > a] = \mu + \frac{\phi(\alpha)}{1-\Phi(\alpha)}\sigma$$

$$\alpha = \frac{a-\mu}{\sigma}$$

If $x_{t+1\to t+s} \sim N(E[x_{t+1\to t+s} \mid I_t], Var[x_{t+1\to t+s} \mid I_t])$.

then,
$$ES_{1-p} = E[x_{t+1\to t+s} \mid x_{t+1\to t+s} > VaR_{1-p}]$$
$$= E[x_{t+1\to t+s} \mid I_t] + \frac{\phi(\alpha)}{1-\Phi(\alpha)}\sqrt{Var(x_{t+1\to t+s} \mid I_t)}$$

where $\alpha = \dfrac{VaR_{1-p} - E[x_{t+1\to t+s} \mid I_t]}{\sqrt{Var(x_{t+1\to t+s} \mid I_t)}}$

and
$$1-p = Pr(x_{t+1\to t+s} < VaR_{1-p} \mid I_t)$$
$$= Pr\left(\frac{x_{t+1\to t+s} - E[x_{t+1\to t+s} \mid I_t]}{\sqrt{Var(x_{t+1\to t+s} \mid I_t)}} = Z < \frac{VaR - E}{\sqrt{Var}} \mid I_t\right)$$
$$= Pr(Z < Z_{1-p})$$
$$\therefore Z_{1-p} = \frac{VaR - E}{\sqrt{Var}}$$

$$\Rightarrow ES_{1-p} = E[x_{t+1\to t+s} \mid I_t] + \frac{\phi(Z_{1-p})}{1-\Phi(Z_{1-p})}\sqrt{Var(x_{t+1\to t+s} \mid I_t)}$$
$$= E[\qquad] + \frac{\phi(Z_{1-p})}{p}\sqrt{Var}$$

only

Normal distribution. ~~If~~ $x_{t+1\to t+s}|I_t \sim N(E[x_{t+1\to t+s}|I_t], Var[x_{t+1\to t+s}|I_t])$, one obtains

$$\mathrm{ES}_{1-p} = E[x_{t+1\to t+s}|I_t] + \frac{\phi(\alpha)}{1-\Phi(\alpha)}\sqrt{Var[x_{t+1\to t+s}|I_t]}, \tag{1.5}$$

where

$$\alpha = \frac{\mathrm{VaR}_{1-p} - E[x_{t+1\to t+s}|I_t]}{\sqrt{Var[x_{t+1\to t+s}|I_t]}}. \quad = z_{1-p}$$

Since it shows

$$\begin{aligned}
1-p &= \Pr\left(x_{t+1\to t+s} \le \mathrm{VaR}_{1-p}|I_t\right) \\
&= \Pr\left(\frac{x_{t+1\to t+s} - E[x_{t+1\to t+s}|I_t]}{\sqrt{Var[x_{t+1\to t+s}|I_t]}} \le \frac{\mathrm{VaR}_{1-p} - E[x_{t+1\to t+s}|I_t]}{\sqrt{Var[x_{t+1\to t+s}|I_t]}} \middle| I_t\right) \\
&= \Pr\left(Z \le \alpha|I_t\right),
\end{aligned}$$

α is the $(1-p)$th quantile of the standard normal distribution: i.e., $\alpha = z_{1-p}$. Consequently, (1.5) changes to

$$\begin{aligned}
\mathrm{ES}_{1-p} &= E[x_{t+1\to t+s}|I_t] + \frac{\phi(z_{1-p})}{1-\Phi(z_{1-p})}\sqrt{Var[x_{t+1\to t+s}|I_t]} \\
&= E[x_{t+1\to t+s}|I_t] + \frac{\phi(z_{1-p})}{p}\sqrt{Var[x_{t+1\to t+s}|I_t]}.
\end{aligned}$$

2. RiskMetrics

For VaR calculation, the RiskMetrics methodology (developed by J.P. Morgan) specifies a model of the form

$\mu_t = E[r_t | I_{t-1}] \qquad \sigma_t^2 = Var[r_t | I_{t-1}]$

$$r_t = \mu_t + \varepsilon_t \quad = \varepsilon_t \qquad \text{ARMA}(0,0) - \text{IGARCH}(1,1).$$

$$\mu_t = 0$$

$$\sigma_t^2 = b_1\sigma_{t-1}^2 + (1-b_1)\varepsilon_{t-1}^2, \quad a_0 = 0 \tag{2.1}$$

shock

where $\varepsilon_t = z_t\sigma_t$ and $z_t \overset{iid}{\sim} N(0,1)$.

$$\sigma_t^2 = (1-b_1)\varepsilon_{t-1}^2 + b_1\sigma_{t-1}^2 = (1-b_1)(\varepsilon_{t-1}^2 + b_1\varepsilon_{t-2}^2 + b_1^2\varepsilon_{t-3}^2 + \cdots)$$

- With the information set at time $t-1$ (denoted by I_{t-1}), the RiskMetrics methodology implies that

$$r_t|I_{t-1} \sim N(0, \sigma_t^2)$$

and the conditional volatility σ_t is governed by an IGARCH(1, 1) model.

Theorem 2.1. *Suppose that X and Y are independent. Let g be a measurable function defined over a Borel set. Then, $g(X)$ and $g(Y)$ are also independent.*

Proof. By definition, X and Y are independent if

$$F(x,y) = F(x)F(y)$$

or equivalently

$$\Pr(X \in A, Y \in B) = \Pr(X \in A)\Pr(Y \in B)$$

for any Borel sets A and B. One shows

$$\begin{aligned} F(g(X),g(Y)) &= \Pr(g(X) \leq x, g(Y) \leq y) \\ &= \Pr(X \in g^{-1}(-\infty,x], Y \in g^{-1}(-\infty,y]) \\ &= \Pr(X \in g^{-1}(-\infty,x])\Pr(Y \in g^{-1}(-\infty,y]) \\ &= \Pr(g(X) \leq x)\Pr(g(Y) \leq y) \\ &= F(g(X))F(g(Y)). \end{aligned}$$

□

- For instance, consider $X = U^2$, $Y = V^2$, and $g(x) = \sqrt{x}$. If X and Y are independent, then $\sqrt{X}$ and $\sqrt{Y}$ are also independent; that is, if U^2 and V^2 are independent, then U and V are also independent.

Proposition 2.2. *The RiskMetrics methodology implies* k step ahead forecast of variance.

$$r_{t+k}|I_t \sim N(0, E[\sigma^2_{t+k}|I_t])$$

$r_{t+k} = \varepsilon_{t+k} = \sigma_t z_t$ $\quad \varepsilon_t = \sigma_t z_t$

for $k > 0$.

Proof. From the IGARCH property that $\sigma^2_{t+1} = (1-b_1)\left[\varepsilon^2_t + b_1\varepsilon^2_{t-1} + b_1^2\varepsilon^2_{t-2} + \cdots\right]$, one knows that σ^2_{t+1} is a function of $\{z^2_t, z^2_{t-1}, z^2_{t-2}\cdots\}$, which implies $E[\sigma^2_{t+1}z^2_{t+1}|I_t] = E[\sigma^2_{t+1}|I_t]E[z^2_{t+1}|I_t]$. By repeated substitutions, one obtains that $E[\sigma^2_{t+k}z^2_{t+k}|I_t] = E[\sigma^2_{t+k}|I_t]E[z^2_{t+k}|I_t]$ and thus $E[\sigma_{t+k}z_{t+k}|I_t] = E[\sigma_{t+k}|I_t]E[z_{t+k}|I_t]$ for $k > 0$. Notice that

$$\begin{aligned} E[r_{t+k}|I_t] &= E[\varepsilon_{t+k}|I_t] \qquad (\because\ r_t = \mu_t + \varepsilon_t,\ \mu_t = 0). \\ &= E[z_{t+k}\sigma_{t+k}|I_t] \\ &= E[\sigma_{t+k}|I_t]E[z_{t+k}|I_t] \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} Var[r_{t+k}|I_t] &= E[r^2_{t+k}|I_t] \\ &= E[z^2_{t+k}\sigma^2_{t+k}|I_t] \\ &= E[\sigma^2_{t+k}|I_t]E[z^2_{t+k}|I_t] \\ &= E[\sigma^2_{t+k}|I_t]. = \hat{\sigma}^2_t[k] \end{aligned}$$

Since z_t follows a normal distribution, the conditional distribution of $z_{t+k}\sigma_{t+k}$ on I_t is also normal. □

Proposition 2.3. *For* $r_{t+1\to t+s} = r_{t+1} + \cdots + r_{t+s}$, *the RiskMetrics methodology implies*

$r_{t+1} = \log R_{t+1}$

$$r_{t+1\to t+s}|I_t \sim N\left(0, s\hat{\sigma}^2_t[1]\right),$$

loss distribution in percentage VaR

$r_t = \mu_t + \varepsilon_t$

$\mu_t = 0$

$\sigma_t^2 = b_1 \sigma_{t-1}^2 + (1-b_1)\varepsilon_{t-1}^2 \quad \Rightarrow \quad \text{IGARCH}(1,1).$

$\varepsilon = z_t \sigma_t \qquad z_t \overset{iid}{\sim} N(0,1) \quad \Rightarrow \quad \sigma_t^2 = (1-b_1)(\varepsilon_{t-1}^2 + b_1 \varepsilon_{t-2}^2 + \cdots)$

$r_t \mid I_{t-1} \sim N(0, \sigma_t^2)$ $\quad$ $\mu_t = 0$ 이기 때문 가정!

$r_{t+k} \mid I_t \sim N(0, E[\sigma_{t+k}^2 \mid I_t])$

$$E[r_{t+k} \mid I_t] = E[\varepsilon_{t+k} \mid I_t] = E[z_{t+k}\sigma_{t+k} \mid I_t]$$
$$= E[z_{t+k} \mid I_t]\, E[\sigma_{t+k} \mid I_t] = 0$$
$$Var[r_{t+k} \mid I_t] = E[r_{t+k}^2 \mid I_t] = E[\varepsilon_{t+k}^2 \mid I_t] = E[z_{t+1}^2 \sigma_{t+1}^2 \mid I_t]$$
$$= E[z_{t+k}^2 \mid I_t] \cdot E[\sigma_{t+k}^2 \mid I_t] = E[\sigma_{t+k}^2 \mid I_t]$$

$r_{t+1 \to t+s} \sim N(0, s\hat{\sigma}_t^2[1])$

$$E[r_{t+1 \to t+s} \mid I_t] = E[r_{t+1} + r_{t+2} + \cdots + r_{t+s} \mid I_t]$$
$$= E[r_{t+1} \mid I_t] + \cdots + E[r_{t+s} \mid I_t] = 0.$$
$$Cov[r_{t+m}, r_{t+n} \mid I_t] = E[r_{t+m} r_{t+n} \mid I_t]$$
$$= E[z_{t+m}\sigma_{t+m} z_{t+n}\sigma_{t+n} \mid I_t]$$
$$= E[z_{t+m} z_{t+n} \sigma_{t+m} \sigma_{t+n} \mid I_t]$$
$$= E[z_{t+m} z_{t+n} \mid I_t]\, E[\sigma_{t+m}\sigma_{t+n} \mid I_t]$$
$$= 0.$$

$$E[\sigma_{t+k}^2 \mid I_t] = E[b_1 \sigma_{t+k-1}^2 + (1-b_1)\sigma_{t+k-1}^2 z_{t+k-1}^2 \mid I_t]$$
$$= b_1 E[\sigma_{t+k-1}^2 \mid I_t] + (1-b_1) E[\sigma_{t+k-1}^2 \mid I_t]$$
$$= E[\sigma_{t+k-1}^2 \mid I_t] = \hat{\sigma}_t^2[k-1] = \hat{\sigma}_t^2[1]$$
$$Var[r_{t+1} + \cdots + r_{t+s} \mid I_t] = Var[r_{t+1} \mid I_t] + \cdots + Var[r_{t+s} \mid I_t]$$
$$= E[\sigma_{t+1}^2 \mid I_t] + \cdots + E[\sigma_{t+s}^2 \mid I_t]$$
$$= \hat{\sigma}_t^2[1] + \cdots + \hat{\sigma}_t^2[1] = s\hat{\sigma}_t^2[1]$$

$$VaR_{1-p} = E[x_{t+1 \to t+s} \mid I_t] + z_{1-p} \cdot \sqrt{Var[x_{t+1 \to t+s} \mid I_t]}$$
$$ES_{1-p} = E[x_{t+1 \to t+s} \mid I_t] + \frac{\phi(z_{1-p})}{p} \cdot \sqrt{Var(x_{t+1 \to t+s} \mid I_t)}$$

$$VaR_{1-p} \text{ in } \% = E[r_{t+1 \to t+s} \mid I_t] + z_{1-p} \sqrt{Var[r_{t+1 \to t+s} \mid I_t)}$$
$$= z_{1-p} \cdot \sqrt{s \, \hat{\sigma}_t^2 [1]}$$

$$ES_{1-p} \text{ in } \% = E[r_{t+1 \to t+s} \mid I_t] + \frac{\phi(z_{1-p})}{p} \cdot \sqrt{Var[r_{t+1 \to t+s} \mid I_t)}$$
$$= \frac{\phi(z_{1-p})}{p} \cdot \sqrt{s \, \hat{\sigma}_t^2 [1]}$$

where $\hat{\sigma}_t^2[1]$ *is the 1-step ahead forecast from (2.1).*

Proof. Since $E[r_{t+k}|I_t] = 0$ for $k > 0$, one shows

$$\underline{E[r_{t+1\to t+s}|I_t]} = E[r_{t+1}|I_t] + \cdots + E[r_{t+s}|I_t] = \underline{0.}$$

Using the independent assumption of z_t, one obtains for $m \neq n > 0$

$$\begin{aligned}
\underline{Cov[r_{t+m}, r_{t+n}|I_t]} &= E[(r_{t+m} - E[r_{t+m}|I_t])(r_{t+n} - E[r_{t+n}|I_t])|I_t] \\
&= E[r_{t+m}r_{t+n}|I_t] \\
&= E[z_{t+m}\sigma_{t+m}z_{t+n}\sigma_{t+n}|I_t] \\
&= E[\sigma_{t+m}\sigma_{t+n}|I_t]E[z_{t+m}z_{t+n}|I_t] \\
&= \underline{0.}
\end{aligned}$$

From (2.1), one shows

$$\begin{aligned}
\underline{E[\sigma_{t+k}^2|I_t]} &= b_1E[\sigma_{t+k-1}^2|I_t] + (1-b_1)E[\varepsilon_{t+k-1}^2|I_t] \\
&= b_1E[\sigma_{t+k-1}^2|I_t] + (1-b_1)E[\sigma_{t+k-1}^2 z_{t+k-1}^2|I_t] \\
&= b_1E[\sigma_{t+k-1}^2|I_t] + (1-b_1)E[\sigma_{t+k-1}^2|I_t] \\
&= E[\sigma_{t+k-1}^2|I_t] = \cdots = E[\sigma_{t+1}^2|I_t] = \underline{\hat{\sigma}_t^2[1]}
\end{aligned}$$

for $k > 1$. By repeated substitutions, one obtains that $E[\sigma_{t+k}^2|I_t] = E[\sigma_{t+1}^2|I_t] = \hat{\sigma}_t^2[1]$ for $k > 1$. Consequently, it shows

$$\begin{aligned}
Var[r_{t+1\to t+s}|I_t] &= Var[r_{t+1}|I_t] + \cdots + Var[r_{t+s}|I_t] \\
&= E[\sigma_{t+1}^2|I_t] + \cdots + E[\sigma_{t+s}^2|I_t] \\
&= s\hat{\sigma}_t^2[1].
\end{aligned}$$

Since the conditional distribution of r_{t+k} is normal and a linear combination of normal random variables is normal, the conditional distribution of $r_{t+1\to t+s}$ is normal. □

- With $r_{t+1\to t+s}|I_t \sim N(0, s\hat{\sigma}^2[1])$, one obtains

$$\text{VaR}_{1-p} \text{ in } \% = z_{1-p}\sqrt{s\hat{\sigma}_t^2[1]}$$

and

$$\text{ES}_{1-p} \text{ in } \% = \frac{\phi(z_{0.95})}{p}\sqrt{s\hat{\sigma}_t^2[1]}.$$

z_{1-p} $x_{t+1\to t+s} \to r_{t+1\to t+s}$ in %

The dollar amount of VaR_{1-p} is computed as the amount of a position times VaR_{1-p} in %, and the dollar amount of ES_{1-p} is computed as the amount of a position times ES_{1-p} in %.

Example 2.4. Consider daily returns of IBM stock from January 2, 2001 to December 31, 2010 for 2515 observations. One has a long position of \$1 million on the stock. So, loss is represented as negative daily log returns.

```
> mydat <- read.table("data1.txt", header = T)
> head(mydat)
      date    return
1 20010102 -0.002206
2 20010103  0.115696
3 20010104 -0.015192
4 20010105  0.008719
5 20010108 -0.004654
6 20010109 -0.010688
> loss <- -log(mydat$return + 1)
```

∵ long position

```
> library(rugarch)
> spec <- ugarchspec(mean.model = list(armaOrder = c(0, 0), include.mean = F),
                     variance.model = list(model = "iGARCH",
                                           garchOrder = c(1, 1)),
                     distribution.model = "norm",
                     fixed.pars = list(omega = 0))
```

$\mu_t = 0$

```
> (fit <- ugarchfit(spec = spec, data = loss))
       Estimate Std. Error t value Pr(>|t|)
omega  0.000000         NA      NA       NA
alpha1 0.057143   0.007172  7.9675        0
beta1  0.942857         NA      NA       NA
```

$\sigma_t^2 = b_1 \sigma_{t-1}^2 + (1-b_1)\varepsilon_{t-1}^2$

$\Rightarrow$ no intercept. (a_0)

$b_1 = 1 - a_1$

- One obtains

$$\sigma_t^2 = 0.942857\sigma_{t-1}^2 + 0.057143\varepsilon_{t-1}^2.$$

```
> pred <- ugarchforecast(fit, n.ahead = 1)
> (sig1 <- as.numeric(sigma(pred)))
[1] 0.007133031
```

- One obtains $\hat{\sigma}_t[1] = 0.007133031$. $= \sqrt{\hat{\sigma}_t^2[1]}$

VaR_{1-p} in %

$= z_{1-p}\sqrt{s\,\hat{\sigma}_t^2[1]}$ $(S=1)$

```
> (var_0.95 <- qnorm(0.95)*sig1)
[1] 0.01173279
> (var_0.99 <- qnorm(0.99)*sig1)
[1] 0.01659391
> round(1000000*var_0.95, digits = 0)
[1] 11733
> round(1000000*var_0.99, digits = 0)
[1] 16594
```

- The 1-day $\mathrm{VaR}_{0.95}$ is \$11,733 and the 1-day $\mathrm{VaR}_{0.99}$ is \$16,594. Put differently, the conditional probability that the loss of a long position of \$1 million for the next trading day is less than or equal to \$11,733 (or \$16,594) is greater than or equal to 95% (or 99%).

```
> (var_0.95 <- qnorm(0.95)*sqrt(15)*sig1)
[1] 0.04544091
> (var_0.99 <- qnorm(0.99)*sqrt(15)*sig1)
[1] 0.06426794
> round(1000000*var_0.95, digits = 0)
[1] 45441
> round(1000000*var_0.99, digits = 0)
[1] 64268
```

15-day VaR_{1-p} = 45,441

- The 15-day $\text{VaR}_{0.95}$ is \$45,441 and the 15-day $\text{VaR}_{0.99}$ is \$64,268.

```
> (es_0.95 <- sqrt(15)*sig1*dnorm(qnorm(0.95))/0.05)
[1] 0.05698473
> (es_0.99 <- sqrt(15)*sig1*dnorm(qnorm(0.99))/0.01)
[1] 0.0736295
> round(1000000*es_0.95, digits = 0)
[1] 56985
> round(1000000*es_0.99, digits = 0)
[1] 73630
```

pdf cdf

ES_{1-p} in %

: $\frac{\phi(z_{1-p})}{p}\sqrt{15}\hat{\sigma}_t(l)$

- The 15-day $\text{ES}_{0.95}$ is \$56,985 and the 15-day $\text{ES}_{0.99}$ is \$73,630. Put differently, the expected loss of the long position for the next 15 days given that the loss exceeds \$45,441 (or \$64,268) is \$56,985 (or \$73,630).

Remark 2.5. The RiskMetrics methodology is simple and easy to understand. However, there are some disadvantages as well. First, the normality assumption used often results in underestimation of VaR, as returns tend to have heavy tails. Second, the zero conditional mean assumption is not valid for returns of some stocks.

3. Econometric Approach to VaR Calculation

3.1. *Single Period*

A general time series model for the log return r_t is given by

$$r_t = \mu_t + \varepsilon_t$$

$$\mu_t = \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \tag{3.1}$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^{m} a_i \varepsilon_{t-i}^2 + \sum_{j=1}^{n} b_j \sigma_{t-j}^2, \quad \text{IGARCH.} \tag{3.2}$$

where $\varepsilon_t = \sigma_t z_t$ and $z_t \sim WN(0,1)$.

- The mean equation follows a "stationary" ARMA(p, q) model, and the variance equation follows a GARCH(m, n) model. Other volatility models can also be used.

Proposition 3.1. *The conditional mean of r_{t+1} is given by $\hat{\mu}_t[1]$, the 1-step ahead forecast of the mean equation (3.1). The conditional variance of r_{t+1} is given by $\hat{\sigma}_t^2[1]$, the 1-step ahead forecast of the variance equation (3.2).*

$\mu_t = E(r_t | I_{t-1})$

Proof. Using $r_{t+1} = \mu_{t+1} + \varepsilon_{t+1}$, one shows

$$\begin{aligned} E[r_{t+1}|I_t] &= E[\mu_{t+1} + \varepsilon_{t+1}|I_t] \\ &= E[\mu_{t+1}|I_t] + E[z_{t+1}\sigma_{t+1}|I_t] \\ &= E[\mu_{t+1}|I_t] + E[z_{t+1}|I_t]E[\sigma_{t+1}|I_t] \\ &= \hat{\mu}_t[1]. \end{aligned}$$

The Wold theorem implies that the MA representation of the stationary ARMA model for r_t is given by

$$r_t = \mu + \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + \cdots;$$

$E[z_{t+1}\sigma_{t+1}|I_t] = E[z_{t+1}|I_t]E[\sigma_{t+1}|I_t] = 0.$

consequently, one obtains

$$\begin{aligned} r_{t+1} - E[r_{t+1}|I_t] &= (\mu + \varepsilon_{t+1} + \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \cdots) - (\mu + E[\varepsilon_{t+1}|I_t] + \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \cdots) \\ &= (\mu + \varepsilon_{t+1} + \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \cdots) - (\mu + \psi_1\varepsilon_t + \psi_2\varepsilon_{t-1} + \cdots) \\ &= \varepsilon_{t+1}. \end{aligned}$$

Therefore, the conditional variance of r_{t+1} is

$$\begin{aligned} Var[r_{t+1}|I_t] &= E[(r_{t+1} - E[r_{t+1}|I_t])^2|I_t] \\ &= E[\varepsilon_{t+1}^2|I_t] \\ &= E[z_{t+1}^2|I_t]E[\sigma_{t+1}^2|I_t] \\ &= E[\sigma_{t+1}^2|I_t]. = \hat{\sigma}_t^2[1] \end{aligned}$$

If
$z_t \overset{iid}{\sim} N(0,1)$
$r_{t+1} | I_t \sim N(\hat{\mu}_t[1], \hat{\sigma}_t^2[1])$

□

일반화.

Normal distribution. If $z_t \overset{iid}{\sim} N(0,1)$, then $r_{t+1}|I_t \sim N(\hat{\mu}_t[1], \hat{\sigma}_t^2[1])$, where $\hat{\mu}_t[1]$ comes from (3.1) and $\hat{\sigma}_t^2[1]$ comes from (3.2). So, the 1-period VaR_{1-p} in % is

$$\mathrm{VaR}_{1-p} \text{ in } \% = \hat{\mu}_t[1] + z_{1-p}\hat{\sigma}_t[1]$$

and the 1-period ES_{1-p} in % is

$$\mathrm{ES}_{1-p} \text{ in } \% = \hat{\mu}_t[1] + \frac{\phi(z_{1-p})}{p}\hat{\sigma}_t[1].$$

non-normal

Standard Student t distribution. If $z_t \overset{iid}{\sim} t_\nu$, then

$$\frac{r_{t+1} - \hat{\mu}_t[1]}{\hat{\sigma}_t[1]} \sim t_\nu.$$

So, the 1-period VaR_{1-p} in % is

$$\text{VaR}_{1-p} \text{ in } \% = \hat{\mu}_t[1] + t_{1-p,\nu}\hat{\sigma}_t[1],$$

where $t_{1-p,\nu}$ denotes the $(1-p)$th quantile of a standard Student t distribution with ν degrees of freedom.

Proposition 3.2. *If the conditional distribution of r_{t+1} follows a standard Student t distribution with ν degrees of freedom, it shows*

$r_{t+1} \sim t_\nu$ difficult.

$$ES_{1-p} \text{ in } \% = \hat{\mu}_t[1] + \frac{f_\nu(t_{1-p,\nu})}{p}\left(\frac{\nu + t^2_{1-p,\nu}}{\nu - 1}\right)\hat{\sigma}_t[1],$$

where $f_\nu(\cdot)$ denotes the pdf of a Student t-distribution with ν degrees of freedom.

Example 3.3. Consider an ARMA(0, 0)-GARCH(1, 1) model for daily returns of IBM stock from January 2, 2001 to December 31, 2010. One has a long position of $1 million on the stock.

$\mu_t \neq 0.$

```
> spec <- ugarchspec(mean.model = list(armaOrder = c(0, 0)),
                     variance.model = list(garchOrder = c(1, 1)),
                                          distribution.model = "norm")
> (fit <- ugarchfit(spec = spec, data = loss))
       Estimate Std. Error t value Pr(>|t|)
mu     -0.000601   0.000239 -2.5089 0.012110
omega   0.000004   0.000002  2.2594 0.023861
alpha1  0.099878   0.015262  6.5442 0.000000
beta1   0.885105   0.016966 52.1693 0.000000
> pred <- ugarchforecast(fit, n.ahead = 1)
> mu1 <- as.numeric(fitted(pred))
> sig1 <- as.numeric(sigma(pred))
> var_0.95 <- mu1 + qnorm(0.95)*sig1
> round(1000000*var_0.95, digits = 0)
[1] 12286
```

mu1 ⊕

```
> es_0.95 <- sig1*dnorm(qnorm(0.95))/0.05
> round(1000000*es_0.95, digits = 0)
[1] 16160
```

- With normal innovations, the fitted model is

$$\mu_t = -0.000601$$
$$\sigma_t^2 = 0.000004 + 0.099878\epsilon^2_{t-1} + 0.885105\sigma^2_{t-1}.$$

The 1-day $\text{VaR}_{0.95}$ is $12,286 and the 1-day $\text{ES}_{0.95}$ is $16,160.

```
> spec <- ugarchspec(mean.model = list(armaOrder = c(0, 0)),
                     variance.model = list(garchOrder = c(1, 1)),
```

```
                                       distribution.model = "std")
```

$z_t \overset{iid}{\sim} t(\nu)$

```
> (fit <- ugarchfit(spec = spec, data = loss))
       Estimate Std. Error t value Pr(>|t|)
mu     -0.000411   0.000225 -1.8228 0.068328
omega   0.000002   0.000001  1.8363 0.066307
alpha1  0.064092   0.010700  5.9897 0.000000
beta1   0.929029   0.011362 81.7697 0.000000
shape   5.756441   0.597132  9.6401 0.000000
> pred <- ugarchforecast(fit, n.ahead = 1)
> mu1 <- as.numeric(fitted(pred))
> sig1 <- as.numeric(sigma(pred))
> x <- qt(0.95, 5.756441)
```

$t_{0.95}(5.756441)$.

```
> var_0.95 <- mu1 + x*sig1
> round(1000000*var_0.95, digits = 0)
[1] 15448
```

mu1 +

```
> es_0.95 <- sig1*(dt(x, 5.756441)/0.05)*((5.756441 + x^2)/(5.756441 - 1))
> round(1000000*es_0.99, digits = 0)
[1] 20881
```

- With standard Student t innovations, the fitted model is

$$
\begin{aligned}
\mu_t &= -0.000411 \\
\sigma_t^2 &= 0.000002 + 0.064092\varepsilon_{t-1}^2 + 0.929029\sigma_{t-1}^2 \\
z_t &\sim t_{5.756441}.
\end{aligned}
$$

The 1-day $\text{VaR}_{0.95}$ is \$15,448 and the 1-day $\text{ES}_{0.95}$ is \$20,881. Compared with the results from the normal innovations, the heavy-tailed innovations give rise to higher risk measures. In other words, if the normality assumption for r_t is rejected, then the VaR under normality is likely to underestimate the true risk.

3.2. Multiple Periods

2 // complicate.

Proposition 3.4. *Consider $r_{t+1\to t+s} = r_{t+1} + \cdots + r_{t+s}$. The conditional mean of $r_{t+1\to t+s}$ on I_t is*

$$E[r_{t+1\to t+s}|I_t] = \hat{\mu}_t[1] + \cdots + \hat{\mu}_t[s], \tag{3.3}$$

where $\hat{\mu}_t[s]$ is the s-step ahead forecast from the ARMA(p, q) model in (3.1). The conditional variance of $r_{t+1\to t+s}$ on I_t is

$$
\begin{aligned}
Var[r_{t+1\to t+s}|I_t] &= (1+\psi_1+\cdots+\psi_{s-1})^2\hat{\sigma}_t^2[1] + (1+\psi_1+\cdots+\psi_{s-2})^2\hat{\sigma}_t^2[2] + \cdots \\
&\quad + (1+\psi_1)^2\hat{\sigma}_t^2[s-1] + \hat{\sigma}_t^2[s],
\end{aligned} \tag{3.4}
$$

where $\hat{\sigma}_t^2[s]$ is the s-step ahead forecast from the GARCH(m, n) model in (3.2) and ψ_i is the MA coefficient implied by the stationary ARMA model for r_t, i.e.,

$$r_t = \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots . \tag{3.5}$$

Proof. Since $E[r_{t+s}|I_t] = \hat{\mu}_t[s]$ for $s > 0$, the conditional mean of $r_{t+1\to t+s}$ is

$$\begin{aligned} E[r_{t+1\to t+s}|I_t] &= E[r_{t+1}|I_t] + \cdots + E[r_{t+s}|I_t] \\ &= \hat{\mu}_t[1] + \cdots + \hat{\mu}_t[s]. \end{aligned}$$

From (3.5), one obtains that

$$\begin{aligned} r_{t+s} - E[r_{t+s}|I_t] &= (\mu + \varepsilon_{t+s} + \psi_1 \varepsilon_{t+s-1} + \cdots + \psi_{s-1}\varepsilon_{t+1} + \psi_s \varepsilon_t + \cdots) \\ &\quad -(\mu + E[\varepsilon_{t+s}|I_t] + \psi_1 E[\varepsilon_{t+s-1}|I_t] + \cdots + \psi_{s-1}E[\varepsilon_{t+1}|I_t] + \psi_s \varepsilon_t + \cdots) \\ &= \varepsilon_{t+s} + \psi_1 \varepsilon_{t+s-1} + \cdots + \psi_{s-1}\varepsilon_{t+1} \end{aligned}$$

and

$$\begin{aligned} r_{t+1\to t+s} - E[r_{t+1\to t+s}|I_t] &= (r_{t+1} + r_{t+2} + \cdots + r_{t+s}) - (E[r_{t+1}|I_t] + \cdots + E[r_{t+s}|I_t]) \\ &= (r_{t+1} - E[r_{t+1}|I_t]) + (r_{t+2} - E[r_{t+2}|I_t]) + \cdots + (r_{t+s} - E[r_{t+s}|I_t]) \\ &= \varepsilon_{t+1} + (\varepsilon_{t+2} + \psi_1 \varepsilon_{t+1}) + \cdots + (\varepsilon_{t+s} + \psi_1 \varepsilon_{t+s-1} + \cdots + \psi_{s-1}\varepsilon_{t+1}) \\ &= (1 + \psi_1 + \cdots + \psi_{s-1})\varepsilon_{t+1} + (1 + \psi_1 + \cdots + \psi_{s-2})\varepsilon_{t+2} + \cdots \\ &\quad + (1 + \psi_1)\varepsilon_{t+s-1} + \varepsilon_{t+s} \\ &= \left(\sum_{i=0}^{s-1} \psi_i\right) \varepsilon_{t+1} + \left(\sum_{i=0}^{s-2} \psi_i\right) \varepsilon_{t+2} + \cdots + (1+\psi_1)\varepsilon_{t+s-1} + \varepsilon_{t+s}, \quad // \end{aligned}$$

where $\psi_0 = 1$. Since z_{t+i} is independent and $E[\varepsilon_{t+i}^2|I_t] = E[\sigma_{t+i}^2|I_t] = \hat{\sigma}_t^2[i]$ for $i = 1, \ldots, s$, the conditional variance of $r_{t+1\to t+s}$ is

$E[z_{t+i}^2|I_t] = 1$

$$\begin{aligned} E[(r_{t+1\to t+s} - E[(r_{t+1\to t+s}|I_t)^2|I_t] &= E\left[\left(\left(\sum_{i=0}^{s-1} \psi_i\right) \varepsilon_{t+1} + \left(\sum_{i=0}^{s-2} \psi_i\right) \varepsilon_{t+2} \cdots + \varepsilon_{t+s}\right)^2 \middle| I_t\right] \\ &= \left(\sum_{i=0}^{s-1} \psi_i\right)^2 E[\varepsilon_{t+1}^2|I_t] + \left(\sum_{i=0}^{s-2} \psi_i\right)^2 E[\varepsilon_{t+2}^2|I_t] \cdots + E[\varepsilon_{t+s}^2|I_t] \\ &= \left(\sum_{i=0}^{s-1} \psi_i\right)^2 \hat{\sigma}_t^2[1] + \left(\sum_{i=0}^{s-2} \psi_i\right)^2 \hat{\sigma}_t^2[2] + \cdots + \hat{\sigma}_t^2[s]. \end{aligned}$$

∎

Normal distribution. If $z_t \overset{iid}{\sim} N(0,1)$, then the conditional distribution of r_{t+k} on I_t is normal. Since a linear combination of normal random variables is normal, the conditional distribution of $r_{t+1\to t+s}$ is normal with the conditional mean in (3.3) and the conditional variance in (3.4). As a result, the s-period

VaR_{1-p} in % is

$$\text{VaR}_{1-p} \text{ in } \% = E[r_{t+1\to t+s}|I_t] + z_{1-p}\sqrt{Var[r_{t+1\to t+s}|I_t]}$$

and the s-period ES_{1-p} in % is

$$\text{ES}_{1-p} \text{ in } \% = E[r_{t+1\to t+s}|I_t] + \frac{\phi(z_{1-p})}{p}\sqrt{Var[r_{t+1\to t+s}|I_t]}.$$

Remark 3.5. If $z_t \overset{iid}{\sim} t_\nu$, then the conditional distribution of r_{t+k} on I_t follows a standard Student t distribution with ν degrees of freedom. However, a linear combination of standard Student t random variables does not follow a standard Student t distribution; i.e., the conditional distribution of $r_{t+1\to t+s}$ is not a standard Student t distribution with ν degrees of freedom. In this case, one uses “simulation” to compute multiperiod VaR_{1-p} and ES_{1-p}.

Example 3.6. Consider an ARMA(0, 0)-GARCH(1, 1) model with a Gaussian error for daily returns of IBM stock from January 2, 2001 to December 31, 2010 for 2515 observations. One has a long position of 1 million on the stock, and is interested in evaluating the loss of the position for the next 15 days.

```
> mydat <- read.table("data1.txt", header = T)
> loss <- -log(mydat$return + 1)
> spec <- ugarchspec(mean.model = list(armaOrder = c(0, 0)),
                     variance.model = list(garchOrder = c(1, 1)),
                     distribution.model = "norm")
> fit <- ugarchfit(spec = spec, data = loss)
> pred <- ugarchforecast(fit, n.ahead = 15)
> (mu <- sum(as.numeric(fitted(pred))))
[1] -0.009012482
> sig2 <- as.numeric(sigma(pred)^2)
> psi <- c(1, ARMAtoMA(ar = coef(fit)[1], ma = 0, lag.max = 14))
> (Sig2 <- sum(rev(cumsum(psi))*sig2))
[1] 0.001256978
```

$\hat{\sigma}_t^2[S]$ $\quad S=1,\cdots,15$

ψ coefficient //

- With normal innovations, one obtains

$$r_{t+1\to t+s}|I_t \sim N(-0.009012482, 0.001256978). \quad //$$

```
> var_0.95 <- mu + qnorm(0.95)*sqrt(Sig2)
> round(1000000*var_0.95, digits = 0)
[1] 49304
> es_0.95 <- mu + sqrt(Sig2)*dnorm(qnorm(0.95))/0.05
> round(1000000*es_0.95, digits = 0)
[1] 64119
```

- The 15-day $\text{VaR}_{0.95}$ of a long position of \$1 million is \$49,304 and the 15-day $\text{ES}_{0.95}$ of the same position is \$64,119.

3.3. Computing Mutiperiod VaR and ES with Simulation

From the fitted model with a standard Student t distribution, one can simulate $r_{t+1\to t+s}|I_t$ many times and thus construct an empirical distribution of $r_{t+1\to t+s}|I_t$. Then, the s-period VaR$_{1-p}$ in % is obtained as the $(1-p)$th quantile of the empirical distribution and the s-period ES$_{1-p}$ in % is obtained as the average of all simulated $r_{t+1\to t+s}|I_t$ satisfying that $r_{t+1\to t+s}|I_t$ is greater than the s-period VaR$_{1-p}$.

Step 1: Simulating $r_{t+k}|I_t$ for $k>0$. From the fitted model with a standard Student t distribution with ν^* degrees of freedom, one computes the k-step ahead forecasts of mean and variance, $\hat{\mu}_t[k]$ from (3.1) and $\hat{\sigma}_t^2[k]$ from (3.2). With the forecasts, one simulates $r_{t+k}|I_t$ by implementing

$$\begin{aligned} r_{t+k} &= \hat{\mu}_t[k] + \varepsilon^*_{t+k} \\ \varepsilon^*_{t+k} &= \hat{\sigma}_t[k] z^*_{t+k}, \end{aligned}$$

Assumption $\varepsilon_{t+k} = \sigma_{t+k} z_{t+k}$

$r_t = \mu_t + \varepsilon_t$ $\hat{\sigma}_t$

$\varepsilon_t = \sigma_t z_t$ $\varepsilon^*_{t+k} = \hat{\sigma}_t^2[k] \cdot z^*_{t+k}$

$z^*_{t+k} \sim t(\nu^*)$

where z^*_{t+k} is a "random" draw form the standard Student t distribution with ν^* degrees of freedom.

Step 2: Constructing an empirical distribution of $r_{t+1\to t+s}|I_t$. For $k = 1,2,\ldots,s$, one implements step 1 to simulate $r_{t+1}|I_t, r_{t+2}|I_t,\ldots,r_{t+s}|I_t$. The sum of these simulated returns provide the simulated value of $r_{t+1\to t+s}|I_t$. This procedure is repeated for many times, say 30,000 iterations, so that one has 30,000 simulated values of $r_{t+1\to t+s}|I_t$. These values are used to construct an empirical distribution of $r_{t+1\to t+s}$.

Example 3.7. Consider an ARMA(0, 0)-GARCH(1, 1) model with a standard Student t error for daily returns of IBM stock from January 2, 2001 to December 31, 2010. One has a long position of 1 million on the stock, and is interested in evaluating the loss of the position for the next 15 days.

```
> spec <- ugarchspec(mean.model = list(armaOrder = c(0, 0)),
                     variance.model = list(garchOrder = c(1, 1)),
                                           distribution.model = "std")
> (fit <- ugarchfit(spec = spec, data = loss))
       Estimate Std. Error t value Pr(>|t|)
mu    -0.000411   0.000225 -1.8228 0.068328
omega  0.000002   0.000001  1.8363 0.066307
alpha1 0.064092   0.010700  5.9897 0.000000
beta1  0.929029   0.011362 81.7697 0.000000
shape  5.756441   0.597132  9.6401 0.000000
```

ARMA(0,0) GARCH(1,1) $\sim t$

- The fitted model with standard Student t innovations is

$$\begin{aligned} \mu_t &= -0.000411 \\ \sigma_t^2 &= 0.000002 + 0.064092\varepsilon_{t-1}^2 + 0.929029\sigma_{t-1}^2 \\ z_t &\sim t_{5.756441}. \end{aligned}$$

```
> pred <- ugarchforecast(fit, n.ahead = 15)
> mu <- as.numeric(fitted(pred))
> sig <- as.numeric(sigma(pred))
```

```
> result <- NULL
> for (i in 1:30000){
     result1 <- NULL
     for (j in 1:15){
          z_t <- rt(n = 1, df = 5.756441)
          r_t <- mu[j] + sig[j]*z_t
          result1[j] <- r_t
          }
     result[i] <- sum(result1)
     }
> var_0.95 <- as.numeric(quantile(result, 0.95))
> round(1000000*var_0.95, digits = 0)
[1] 62551
> idx <- which(result > var_0.95)
> es_0.95 <- mean(result[idx])
> round(1000000*es_0.95, digits = 0)
[1] 80240
```

- The 15-day $\mathrm{VaR}_{0.95}$ is \$62,551 and the 15-day $\mathrm{ES}_{0.95}$ is \$80,240. Compared with the results from the normal innovations, the heavy-tailed innovations give rise to higher risk measures.

4. Quantile Estimation

Quantile estimation provides a nonparamteric approach to VaR calculation. It makes no assumption on the specific distribution of the loss variable, but assumes that the previous distribution continues to hold within the prediction period.

Simple. best advantage ↕ disadvantage

4.1. Quantile and Order Statistics

Definition 4.1. The *order statistics* of the sample $x_1, \ldots, x_n$ are the sample values placed in increasing order and satisfy

$$x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)},$$

where $x_{(i)}$ is the ith order statistic of the sample. In particular, $x_{(1)}$ is the sample minimum and $x_{(n)}$ is the sample maximum.

Theorem 4.2. *Assume that the order statistics are independent and identically distributed and have a continuous distribution with pdf $f(x)$ and cdf $F(x)$. Let x_q be the qth quantile of $F(x)$. Provided that the pdf $f(x)$ is not zero at x_q, it shows for the order statistic $x_{(k)}$, where $k = nq$ with $0 < q < 1$,*

$$x_{(k)} \approx N\left(x_q, \frac{q(1-q)}{n[f(x_q)]^2}\right).$$

- The asymptotic mean of $x_{(k)}$ is x_q, where $k = nq$, which means that $x_{(k)}$ is a consistent estimate of x_q; i.e., $x_{(k)}$ converges x_q as the sample size n increases. So, one can use $x_{(k)}$ to estimate the quantile x_q for large n.

$$F_{x_{(k)}}(t) = \sum_{i=k}^{n} \binom{n}{i} [F_X(t)]^i [1-F_X(t)]^{n-i}$$

$$f_{x_{(k)}}(t) = \frac{n!}{(k-1)!(n-k)!} [F_X(t)]^{k-1} [1-F_X(t)]^{n-k} f_X(t)$$

Definition 4.3. *Linear interpolation* is a method of curve fitting using linear polynomials to construct new data points with the range of a "discrete" set of known data points.

- The simple linear interpolation is the straight line between two points. Suppose that two points, (x_1, y_1) and (x_2, y_2), are known. Then, for a value $x \in [x_1, x_2]$, the value y along the straight line is given by

$$\begin{aligned} y &= y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) \\ &= \frac{y_1(x_2 - x) + y_2(x - x_1)}{x_2 - x_1}. \end{aligned}$$

Let $r_1, \ldots, r_T$ be daily returns in the sample period. Assume that the returns are *iid* and the return distribution for tomorrow is the same as that in the sample period.

- For "integer" $k = Tq$, where $q = 1 - p$ and p is a tail probability, the 1-day VaR_{1-p} in % is computed as

$$\text{VaR}_{1-p} \text{ in } \% = r_{(Tq)}.$$

  This is because VaR_{1-p} in % is the qth quantile of the return distribution in the prediction period.

- For "noninteger" $k = Tq$, let k_1 and k_2 be the two neighboring integers such that $k_1 < k < k_2$. Define $q_i = k_i/T$. Given that $(q_1, r_{(k_1)})$ and $(q_2, r_{(k_2)})$ are known, one employs the simple linear interpolation to obtain the 1-day VaR_{1-p} in % as

$$\text{VaR}_{1-p} \text{ in } \% = \frac{q_2 - q}{q_2 - q_1} r_{(k_1)} + \frac{q - q_1}{q_2 - q_1} r_{(k_2)}.$$

- The 1-day ES_{1-p} in % is computed as the sample average of returns that are greater than the 1-day VaR_{1-p} in %.

Example 4.4. Consider daily returns of IBM stock from January 2, 2001 to December 31, 2010 for 2515 observations. One has a long position of 1 million on the stock.

```
> mydat <- read.table("data1.txt", header = T)
> loss <- -log(mydat$return + 1)
> T <- length(loss)
> (k <- n*0.95)
[1] 2389.25
> (k1 <- floor(k))
[1] 2389
> (k2 <- ceiling(k))
[1] 2390
```

- As $k = 2,389.25$, one sets $k_1 = 2,389$ and $k_2 = 2,390$.

```
> q1 <- k1/T
> q2 <- k2/T
```

```
> loss1 <- sort(loss, decreasing = F)
> (var_0.95 <- (q2 - 0.95)/(q2 - q1)*loss1[k1] + (0.95 - q1)/(q2 - q1)*loss1[k2]
[1] 0.02653547
> round(1000000*var_0.95, digits = 0)
[1] 26535
> idx <- which(loss1 > var_0.95)
> (es_0.95 <- mean(loss1[idx]))
[1] 0.03994857
> round(1000000*es_0.95, digits = 0)
[1] 39949
```

- The 1-day $\text{VaR}_{0.95}$ of a long position of \$1 million is \$26,535 and the $\text{ES}_{0.95}$ of the same position is \$39,949.

Remark 4.5. This simple approach requires no specific distributional assumption but instead assumes that the loss distribution remains unchanged from the sample period to the prediction period. Consequently, the predicted loss cannot be greater than that of the historical loss, which is definitely not so in practice.

4.2. *Quantile Regression*

unconditional $X_t \sim F(x)$ $X_{t+1} \sim F(x)$
conditional $X_t \sim F(x)$ $X_{t+1}|I_t \nsim F(x)$

In many applications, the 1-day loss x_{t+1} is affected by explanatory variables included in I_t, so that the quantile of the conditional distribution of x_{t+1} on I_t is interesting. Such a quantile is referred to as a *regression quantile* or a *conditional quantile.*

Example 4.6. The action taken by Federal Reserve Banks on interest rates today could have impacts on the returns (i.e., the loss in percentage) of U.S. stocks tomorrow.

Consider a linear regression of the form

financial loss tomorrow

$$x_{t+1} = \beta' z_t + \varepsilon_{t+1},$$

List matrix
general Econometric
quantile.

where z_t is a vector of explanatory variables that are elements of the information set I_t.

Theorem 4.7. *The (1-p)th conditional quantile, denoted by $x_{1-p}|I_t$, is estimated as*

$$\hat{x}_{1-p}|I_t = \inf\{b' z_t | R_{1-p}(b) = \min\},$$

where "$R_{1-p}(b) = \min$" means that b is the estimate of a quantile regression

weight

$$b = argmin_\beta \sum_{t=1}^{T-1} w_{1-p}(x_{t+1} - \beta' z_t)$$

$$w_{1-p}(x_{t+1} - \beta' z_t) = \begin{cases} (1-p)(x_{t+1} - \beta' z_t) & \text{if } x_{t+1} - \beta' z_t \geq 0 \\ -p(x_{t+1} - \beta' z_t) & \text{if } x_{t+1} - \beta' z_t < 0. \end{cases}$$

quantile을 나타내는 또 하나의 대표적 방법.

Example 4.8. Consider daily returns of IBM stock from January 2, 2001 to December 31, 2010 for 2515 observations. One has a long position of 1 million on the stock. Given that the return distribution at time $t+1$ is related to the stock volatility at time t and the market volatility at time t, one adopts the following quantile regression

$$Q(1-p|z_t) = \sum_{t=1}^{2514} w_{1-p}(x_{t+1} - \beta_0 - \beta_1 s_t - \beta_2 v_t), \tag{4.1}$$

where $x_{t+1} = -r_{t+1}$, s_t is the lag-1 daily volatility of r_t obtained from an ARMA(0, 0)-GARCH(1, 1) model with Gaussian innovations, and v_t is the lag-1 VIX index obtained from CBOE.

⇒ volatility of U.S stock market.

```
> mydat1 <- read.table("data1.txt", header = T)
> head(mydat1)
      date    return
1 20010102 -0.002206
2 20010103  0.115696
3 20010104 -0.015192
4 20010105  0.008719
> mydat1 <- cbind(mydat1, loss = -log(mydat1$return + 1))
> library(rugarch)
> spec <- ugarchspec(mean.model = list(armaOrder = c(0, 0)),
                     variance.model = list(garchOrder = c(1, 1)),
                     distribution.model = "norm")
> fit <- ugarchfit(spec = spec, data = mydat1$loss)
> vol <- as.numeric(sigma(fit))
> mydat2 <- read.csv("data2.csv", header = T)
> head(mydat2)
      Date Adj.Close
1 20010102     29.99
2 20010103     26.60
3 20010104     26.97
4 20010105     28.67
> names(mydat2)[2] <- "vix"
> mydat2$vix <- mydat2$vix/100
> mydat <- cbind(mydat1, vol, vix = mydat2$vix)
> vol_lag <- c(NA, mydat$vol[-dim(mydat)[1]])
> vix_lag <- c(NA, mydat$vix[-dim(mydat)[1]])
> mydat <- cbind(mydat, vol_lag, vix_lag)
> head(mydat)
      date    return         loss        vol    vix    vol_lag vix_lag
1 20010102 -0.002206  0.002208437 0.01699781 0.2999         NA      NA
2 20010103  0.115696 -0.109478425 0.01615122 0.2660 0.01699781  0.2999
3 20010104 -0.015192  0.015308581 0.03767259 0.2697 0.01615122  0.2660
```

```
4 20010105  0.008719 -0.008681209 0.03585786 0.2867 0.03767259  0.2697
5 20010108 -0.004654  0.004664864 0.03389573 0.2984 0.03585786  0.2867
6 20010109 -0.010688  0.010745527 0.03200044 0.2799 0.03389573  0.2984
> library(quantreg)
> fit <- rq(loss ~ vol_lag + vix_lag, tau = 0.95, data = mydat)
> summary(fit)
Coefficients:
               Value  Std. Error  t value  Pr(>|t|)
(Intercept) -0.00103    0.00259 -0.39899  0.68993
vol_lag      1.18416    0.22686  5.21979  0.00000
vix_lag      0.02777    0.01625  1.70954  0.08748
```

- Applying the quantile regression in (4.1) with $1-p=0.95$, one obtains that $\beta_0=-0.00103$, $\beta_1=1.18416$, and $\beta_2=0.02777$.

```
                                                   St                          Vt
> var_0.95 <- coef(fit)[1] + coef(fit)[2]*mydat[2515, 6] + coef(fit)[3]*mydat[2515,
7]                                                 Zt                          Zt
> round(1000000*as.numeric(var_0.95), digits = 0)
[1] 13590
> idx <- which(mydat$loss > var_0.95)
> es_0.95 <- mean(mydat$loss[idx])
> round(1000000*es_0.95, digits = 0)
[1] 25994
```

small than unconditional quantile.

- The 1-day $\mathrm{VaR}_{0.95}$ of a long position of \$1 million is \$13,590 and the 1-day $\mathrm{ES}_{0.95}$ of the same position is \$25,994.

EVT > general Econometric

Peaks - Over - Threshell

$Y = X - \eta$

$Y \mid X \geq \eta$