

# Econometrics for Financial Time Series

## Chapter 11: Value at Risk

In Choi

Sogang University

- Reference:  
Chapter 7 of Tsay.
- For extreme value theory, see:  
Embrechts, P., Kuppelberg, C., and Mikosch, T. (1997), Modelling  
Extremal Events, Berlin: Springer Verlag.
- Methods for calculating VaR (value at risk) and the statistical  
theories behind these methods.

- What is VaR?
  - A measure of financial risk
  - Defined as the maximal loss of a financial position during a given time period for a given probability.
  - Mainly for market risk, but idea applies to credit risk and operational risk too.

# Value at Risk

## Definition of VaR

- We are interested in the risk of a financial position for the next  $l$  periods at time  $t$ .
- $\Delta V(l)$  : the change in the value of the assets in the financial position from time  $t$  to  $t + l$
- $F_l(x)$  : the cumulative distribution function of  $\Delta V(l)$
- The VaR of a long position over the time horizon  $l$  is defined by the relation

$$p = \Pr[\Delta V(l) \leq c_p] = F_l(c_p);$$
$$\text{VaR} = c_p \times \text{amount of position}$$

for a given probability  $p$ .

- A loss results when we observe  $\Delta V(l) < 0$ .
- VaR typically assumes a negative value when  $p$  is small.

# Value at Risk

## Definition of VaR

- The VaR of a short position is defined by the relation

$$\begin{aligned} p &= \Pr[\Delta V(I) \geq c_p] = 1 - \Pr[\Delta V(I) \leq c_p] \\ &= 1 - F_I(c_p); \end{aligned}$$

$$\text{VaR} = c_p \times \text{amount of position}$$

- A loss results when we observe  $\Delta V(I) > 0$ .
- VaR typically assumes a positive value when  $p$  is small.
- The same as the definition of VaR for a long position if  $-\Delta V(I)$  is used instead.
- VaR can be calculated once we know the distribution function.
- We use log returns  $r_t$  in calculating VaR for simplicity.

There are five methods for calculating VaR.

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

## *RiskMetrics*

- Developed by J.P. Morgan
- Assume  $r_t \mid F_{t-1} \sim N(\mu_t, \sigma_t^2)$ .
- Assume IGARCH(1,1) for  $r_t$

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

- Recall that

$$r_t[k] = r_{t+1} + \dots + r_{t+k}.$$

Then,

$$r_t[k] \mid F_t \sim N(0, k\sigma_{t+1}^2).$$

- (This part is optional.) Since  $E(r_{t+i}r_{t+j} | F_t) = 0$  <sup>1</sup>( $i, j > 0, i \neq j$ ), we have

$$\text{Var}(r_t[k] | F_t) = \sum_{i=1}^k \text{Var}(r_{t+i} | F_t).$$

Since  $E(r_t^2 | F_{t-1}) = \sigma_t^2$  by definition,  $\text{Var}(r_{t+1} | F_t) = \sigma_{t+1}^2$ .  
 Moreover, for  $i \geq 2$

$$\begin{aligned} \text{Var}(r_{t+i} | F_t) &= E(r_{t+i}^2 | F_t) \\ &= E(E(r_{t+i}^2 | F_{t+i-1}) | F_t) \quad (F_{t+i-1} \supset F_t) \\ &= E(\sigma_{t+i}^2 | F_t). \end{aligned}$$

---

<sup>1</sup>Suppose  $i > j$ . Then,  $E(r_{t+i}r_{t+j} | F_t) = E(E(r_{t+i}r_{t+j} | F_{t+i-1}) | F_t) = E(r_{t+j}\sigma_{t+i}E(\epsilon_{t+i} | F_{t+i-1}) | F_t) = 0$ .



- Thus,

$$\text{Var}(r_t[k] | F_t) = \sum_{i=1}^k E(\sigma_{t+i}^2 | F_t). \quad (1)$$

Using the relation  $r_t = \sigma_t \epsilon_t$ , rewrite the IGARCH(1,1) model as

$$\sigma_{t+i}^2 = \sigma_{t+i-1}^2 + (1 - \alpha)\sigma_{t+i-1}^2(\epsilon_{t+i-1}^2 - 1),$$

which yields

$$E(\sigma_{t+i}^2 | F_t) = E(\sigma_{t+i-1}^2 | F_t)$$

since  $E(\sigma_{t+i-1}^2(\epsilon_{t+i-1}^2 - 1) | F_t) = 0$ . This implies

$$E(\sigma_{t+k}^2 | F_t) = \dots = E(\sigma_{t+1}^2 | F_t) = \sigma_{t+1}^2. \quad (2)$$

We infer from (1) and (2)

$$\text{Var}(r_t[k] | F_t) = k\sigma_{t+1}^2.$$

- Suppose that the financial position is a long position. If the probability is set to 5%, RiskMetrics uses  $1.65\sigma_{t+1}$  to measure the risk of the portfolio.<sup>2</sup> That is,

$$VaR = \text{Amount of position} \times 1.65\sigma_{t+1}$$

and

$$VaR[k] = \text{Amount of position} \times 1.65\sqrt{k}\sigma_{t+1}$$

---

<sup>2</sup>The actual 5% quantile is  $-1.65\sigma_{t+1}$ , but the negative sign is ignored with the understanding that it signifies a loss.

## Example

Daily log returns of IBM stock from July 3, 1962 to December 31, 1998.  
An IGARCH(1,1) fit gives

$$\sigma_t^2 = 0.9396\sigma_{t-1}^2 + (1 - 0.9396)r_{t-1}^2.$$

Since  $r_{9190} = -0.0128$  and  $\hat{\sigma}_{9190}^2 = 0.0003472$ , the 1-step ahead volatility forecast<sup>a</sup> is

$$\hat{\sigma}_{9190}^2[1] = 0.9396 \times 0.0003472 + (1 - 0.9396) \times (-0.0128)^2 = 0.000336.$$

Therefore, The 5% quantile<sup>b</sup> of the conditional distribution  $r_{9191} \mid F_{9190}$  is  $-1.65 \times \sqrt{0.000336} = -0.03025$ .

---

<sup>a</sup> $\hat{\sigma}_{9190}^2[1]$  is a forecast of  $\sigma_{9191}^2$ .

<sup>b</sup>The  $p$ th quantile of  $F_I(x)$ ,  $x_p$ , is defined by

$$x_p = \inf\{x \mid F_I(x) \geq p\}.$$

## Example

(continued) The 1-day horizon 5% VaR of a long position of \$10 million is

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

Interpretation: “With 5% chance, this financial position can lose \$302,500 tomorrow.”

- An advantage of RiskMetrics is simplicity.
- The normality assumption used often results in underestimation of VaR.
- If either the zero mean assumption or the special IGARCH(1, 1) model assumption of the log returns fails, then the rule is invalid.

## *Econometric modelling*

- Assume for  $r_t$

$$r_t = \phi_0 + \sum_{i=1}^p \phi_i r_{t-i} + a_t - \sum_{j=1}^q \theta_j a_{t-j};$$

$$a_t = \sigma_t \epsilon_t;$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^u \alpha_i a_{t-i}^2 + \sum_{j=1}^v \beta_j \sigma_{t-j}^2.$$

- The 1-step ahead forecasts of the conditional mean and conditional variance of  $r_t$  are

$$\hat{r}_t[1] = \phi_0 + \sum_{i=1}^p \phi_i r_{t+1-i} - \sum_{j=1}^q \theta_j a_{t+1-j};$$

$$\hat{\sigma}_t^2[1] = \alpha_0 + \sum_{i=1}^u \alpha_i a_{t+1-i}^2 + \sum_{j=1}^v \beta_j \sigma_{t+1-j}^2.$$

- Assume  $\epsilon_t \sim iidN(0, 1)$ . Then,

$$r_{t+1} \mid F_t \sim N(\hat{r}_t[1], \hat{\sigma}_t^2[1]).$$

The 5% quantile is  $\hat{r}_t[1] - 1.65\hat{\sigma}_t[1]$ .

- Alternatively, one may assume a t-distribution for  $\epsilon_t$ .

## Example

Daily log returns of IBM stock from July 3, 1962 to December 31, 1998.  
The fitted models is

$$\begin{aligned}r_t &= 0.00066 - 0.0247r_{t-2} + a_t, \quad a_t = \sigma_t \epsilon_t, \\ \sigma_t^2 &= 0.00000389 + 0.9073\sigma_{t-1}^2 + 0.0799r_{t-1}^2.\end{aligned}$$

Since  $r_{9189} = -0.00201$ ,  $r_{9190} = -0.0128$  and  $\hat{\sigma}_{9190}^2 = 0.00033455$ , the 1-step ahead volatility forecasts are

$$\begin{aligned}\hat{r}_{9190}[1] &= 0.00071; \\ \hat{\sigma}_{9190}^2[1] &= 0.0003211.\end{aligned}$$

Therefore, The 5% quantile of the conditional distribution  $r_{9191} \mid F_{9190}$  is

$$0.00071 - 1.65 \times \sqrt{0.0003211} = -0.02877.$$



- The  $k$ -step ahead forecast of  $r_t$  is

$$\hat{r}_t[k] = r_t(1) + \dots + r_t(k).$$

- Using the MA representation of  $r_t$

$$r_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots,$$

we have<sup>3</sup>

$$\begin{aligned} r_t(l) &= E(r_{t+l} | F_t) \\ &= E(\mu + a_{t+l} + \psi_1 a_{t+l-1} + \psi_2 a_{t+l-2} + \dots | F_t) \\ &= \mu + \psi_l a_t + \psi_{l+1} a_{t-1} + \dots \end{aligned}$$

---

<sup>3</sup> $E(a_{t+l} | F_t) = E[E(a_{t+l} | F_{t+l}) | F_t] = E[\sigma_{t+l} E(\epsilon_{t+l} | F_{t+l}) | F_t] = 0$

- Thus, the  $l$ -step ahead forecast error at the forecast origin  $t$  as

$$\begin{aligned}e_t(l) &= r_{t+l} - r_t(l) \\ &= \mu + a_{t+l} + \psi_1 a_{t+l-1} + \psi_2 a_{t+l-2} + \dots \\ &\quad - (\mu + \psi_l a_t + \psi_{l+1} a_{t-1} + \dots) \\ &= a_{t+l} + \psi_1 a_{t+l-1} + \psi_2 a_{t+l-2} + \dots + \psi_{l-1} a_{t+1}.\end{aligned}$$

- The forecast error of the expected  $k$ -period return  $\hat{r}_t[k]$  is the sum of 1-step to  $k$ -step forecast errors of  $r_t$  at the forecast origin  $t$ . It is

$$\begin{aligned}e_t[k] &= r_t[k] - \hat{r}_t[k] \\ &= r_{t+1} + \dots + r_{t+k} - (r_t(1) + \dots + r_t(k)) \\ &= e_t(1) + \dots + e_t(k) \\ &= a_{t+k} + (1 + \psi_1)a_{t+k-1} + \dots + \left( \sum_{i=0}^{k-1} \psi_i \right) a_{t+1}\end{aligned}$$

with  $\psi_0 = 1$ .

- The conditional mean of  $r_t[k]$  given  $F_t$  is  $\hat{r}_t[k]$ . Thus, its conditional variance is the conditional variance of  $e_t[k]$  given  $F_t$ . This is

$$\text{Var}(e_t[k] | F_t) = \sigma_t^2(k) + (1 + \psi_1)^2 \sigma_t^2(k-1) + \dots + \left( \sum_{i=0}^{k-1} \psi_i \right)^2 \sigma_t^2(1).$$

## Example

Let

$$\begin{aligned}r_t &= \mu + a_t; a_t = \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2.\end{aligned}$$

First,

$$\hat{r}_t[k] = r_t(1) + \dots + r_t(k) = k\mu.$$

Next, since  $\psi_i = 0$  for all  $i > 0$ ,

$$\begin{aligned}e_t[k] &= r_t[k] - \hat{r}_t[k] \\ &= a_{t+k} + a_{t+k-1} + \dots + a_{t+1}\end{aligned}$$

and

$$\text{Var}(e_t[k] | F_t) = \sigma_t^2(k) + \sigma_t^2(k-1) + \dots + \sigma_t^2(1).$$

## Example

Using the forecasting method of GARCH(1,1) models, we obtain

$$\begin{aligned}\sigma_t^2(1) &= \alpha_0 + \alpha_1 a_t^2 + \beta_1 \sigma_t^2 \\ \sigma_t^2(l) &= \alpha_0 + (\alpha_1 + \beta_1) \sigma_t^2(l-1), l \geq 2.\end{aligned}$$

These relations give

$$\text{Var}(e_t[k] | F_t) = \frac{\alpha_0}{1-\phi} \left[ k - \frac{1-\phi^k}{1-\phi} \right] + \frac{1-\phi^k}{1-\phi} \sigma_t^2(1),$$

where  $\phi = \alpha_1 + \beta_1$ . If we assume normality for  $\epsilon_t$ , we have

$$r_{t+1} | F_t \sim N(k\mu, \text{Var}(e_t[k] | F_t)).$$

## Quantile estimation

- No distributional assumption is required.
- Let  $r_1, \dots, r_n$  be the returns of a portfolio in the sample period. The order statistics of the sample are these values arranged in increasing order. We use the notation

$$r_{(1)} \leq \dots \leq r_{(n)}.$$

- For  $n$  large,

$$r_{(l)} \sim N \left( x_p, \frac{p(1-p)}{n[f(x_p)]^2} \right), \quad l = np,$$

where  $x_p$  is the  $p$ th quantile of  $F(x)$  [ $x_p = F^{-1}(p)$ ], and  $f(\cdot)$  is the pdf of  $r_t$ .

- Use this result to estimate the quantile  $x_p$ .

# Quantile estimation

- In practice,  $np$  may not be a positive integer. In this case, one can use simple interpolation to obtain quantile estimates. More specifically, for noninteger  $np$ , let  $l_1$  and  $l_2$  be the two neighboring positive integers such that

$$l_1 < np < l_2.$$

Define  $p_i = l_i/n$ . Then,

$$x_{p_1} < x_p < x_{p_2}$$

and

$$r_{(l_1)} \simeq x_{p_1} \quad \text{and} \quad r_{(l_2)} \simeq x_{p_2}.$$

Therefore, the quantile  $x_p$  can be estimated by

$$\hat{x}_p = \frac{p_2 - p}{p_2 - p_1} r_{(l_1)} + \frac{p - p_1}{p_2 - p_1} r_{(l_2)}.$$



## Example

Daily log returns of IBM stock from July 3, 1962 to December 31, 1998. Using all the 9190 observations, the empirical 5% quantile can be obtained as

$$\left( r_{(459)} + r_{(460)} \right) / 2 = -0.021603.,$$

( $np = 9190 \times 0.05 = 459.5$ ). The VaR of a long position of \$10 million is \$216,030.

- Advantages: simplicity and no distributional assumptions
- Disadvantages:
  - 1 The distribution of the return  $r_t$  remains unchanged.
  - 2 For extreme quantiles (i.e., when  $p$  is close to zero or unity), the empirical quantiles are not efficient estimates of the theoretical quantiles.
  - 3 The direct quantile estimation fails to take into account the effect of explanatory variables that are relevant to the portfolio under study.

## *Extreme value theory*

- Focus on the minimum return  $r_{(1)}$ . This is highly relevant to VaR calculation for a long position.
- For the maximum return  $r_{(n)}$ , use the identity

$$\max(r_1, \dots, r_n) = -\min(-r_1, \dots, -r_n).$$

- Assume that the returns  $r_t$  are serially independent with a common cumulative distribution function  $F(x)$  and that the range of the return  $r_t$  is  $[l, u]$ .

- The CDF of  $r_{(1)}$  is given by

$$\begin{aligned}F_{n,1}(x) &= \Pr[r_{(1)} \leq x] = 1 - \Pr[r_{(1)} > x] \\&= 1 - \Pr[r_1 > x, \dots, r_n > x] \\&= 1 - \prod_{j=1}^n \Pr[r_j > x] \\&= 1 - \prod_{j=1}^n (1 - \Pr[r_j \leq x]) \\&= 1 - \prod_{j=1}^n (1 - F(x)) \\&= 1 - (1 - F(x))^n.\end{aligned}$$

- $F_{n,1}(x) \rightarrow 0$  if  $x \leq l$  and  $F_{n,1}(x) \rightarrow 1$  if  $x > l$ . (degenerate distributions; not useful).

- The extreme value theory is concerned with finding two sequences  $\beta_n$  and  $\alpha_n$ , where  $\alpha_n > 0$ , such that the distribution of

$$r_{(1^*)} = \frac{r_{(1)} - \beta_n}{\alpha_n}$$

converges to a nondegenerate distribution as  $n$  goes to infinity.  
( $\beta_n$  : location parameter,  $\alpha_n$  : scale parameter).

# Extreme value theory

- Let the limiting distribution of  $r_{(1^*)}$  be  $F_*(x)$ . The limiting distribution belongs to one of the following three distributions (Fisher-Tippett theorem, 1928)

Type I: Gumbel family ( $k = 0$ )

$$F_*(x) = 1 - \exp[-\exp(x)], \quad -\infty < x < \infty.$$

Type II: Fréchet family ( $k < 0$ )

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

Type III: Weibull family ( $k > 0$ )

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

- $k$  : shape parameter  $\alpha = -1/k$  : tail index
- $\alpha$  is coming from the tail property of the underlying distribution.

- For risk management, we are mainly interested in the Fréchet family.
- If the underlying distributions are, for example, Pareto, Cauchy and stable with exponent  $\alpha < 2$ , the Fréchet family distribution follows. That is, if the underlying distributions have thick tails, the Fréchet family distribution is obtained.
- The pdf of  $r_{(1^*)}$  for the Fréchet family distribution is for  $k < 0$

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

This is used for the maximum likelihood estimation.

## Empirical estimation

- Estimate  $k$ ,  $\beta_n$  and  $\alpha_n$ .
- Assume  $T = ng$ . Divide the data as

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

and write the observed returns as  $r_{in+j}$  ( $1 \leq j \leq n$  and  $i = 0, \dots, g-1$ ).

- Let

$r_{n,i} = \min_{1 \leq j \leq n} \{r_{(i-1)n+j}\}$ ,  $i = 1, \dots, g$ . (the minimum of the  $i$ th sample)



- The collection of subsample minima  $\{r_{n,i}\}$  are the data we use to estimate the unknown parameters of the extreme value distribution.
- Letting  $x_i = (r_{n,i} - \beta_n)/\alpha_n$ , the pdf of  $r_{n,i}$  can be obtained by  $f_*(x)$ . Denoting this as  $f(r_{n,i})$ , the likelihood function is written as.

$$l(r_{n,1}, \dots, r_{n,g} \mid k_n, \alpha_n, \beta_n) = \prod_{i=1}^g f(r_{n,i}).$$

Nonlinear estimation procedures can then be used to obtain maximum likelihood estimates of  $k_n$ ,  $\beta_n$  and  $\alpha_n$ .

# An extreme value approach to VaR

- Suppose that the MLEs  $k_n, \beta_n$  and  $\alpha_n$  are available.
- $p^*$  : a small probability that indicates the potential loss of a long position.
- $r_n^*$  : the  $p^*$ th quantile of the subperiod minimum under the limiting generalized extreme value distribution.
- Then,

$$p^* = \begin{cases} 1 - \exp\left[-\left(1 + \frac{k_n(r_n^* - \beta_n)}{\alpha_n}\right)^{1/k_n}\right] & \text{if } k_n \neq 0 \\ 1 - \exp\left[-\exp\left(\frac{r_n^* - \beta_n}{\alpha_n}\right)\right] & \text{if } k_n = 0 \end{cases}$$

or

$$\ln(1 - p^*) = \begin{cases} -\left(1 + \frac{k_n(r_n^* - \beta_n)}{\alpha_n}\right)^{1/k_n} & \text{if } k_n \neq 0 \\ -\exp\left(\frac{r_n^* - \beta_n}{\alpha_n}\right) & \text{if } k_n = 0 \end{cases} .$$

# An extreme value approach to VaR

- Solving the latter equation with respect to  $r_n^*$ , we obtain the quantile as

$$r_n^* = \begin{cases} \beta_n - \frac{k_n}{\alpha_n} (1 - [-\ln(1 - p^*)])^{k_n} & \text{if } k_n \neq 0 \\ \beta_n + \alpha_n \ln[-\ln(1 - p^*)] & \text{if } k_n = 0 \end{cases}$$

The quantile  $r_n^*$  is the VaR based on the extreme value theory for the subperiod minima. This is used to obtain VaR for the original asset return series  $r_t$ .

# An extreme value approach to VaR

- Relationship between subperiod minima and the observed return  $r_t$ .

$$\begin{aligned} p^* &= \Pr[r_{n,i} \leq r_n^*] \\ &= 1 - [1 - \Pr[r_{n,i} \leq r_n^*]] \\ &= 1 - \Pr[r_{n,i} > r_n^*] \\ &= 1 - \Pr[r_{(i-1)n+1} > r_n^*, \dots, r_{in} > r_n^*] \\ &= 1 - \prod_{t=1}^n \Pr[r_t > r_n^*] \\ &= 1 - [1 - \Pr[r_t \leq r_n^*]]^n \end{aligned}$$

or

$$1 - p^* = [1 - \Pr[r_t \leq r_n^*]]^n.$$

# An extreme value approach to VaR

- If we choose  $p$  such that

$$p = \Pr[r_t \leq r_n^*],$$

then

$$\ln(1 - p^*) = n \ln(1 - p).$$

Consequently, for a given small probability  $p$ , the VaR of holding a long position in the asset underlying the log return  $r_t$  is

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

## Example

Daily log returns of IBM stock from July 3, 1962 to December 31, 1998.  
For  $n = 63$ ,

$$\alpha_n = 0.945, \beta_n = -2.583 \text{ and } k_n = -0.335.$$

Thus, for  $p = 0.05$ ,

$$\text{VaR} = \beta_n - \frac{k_n}{\alpha_n} (1 - [-n \ln(1 - 0.05)]^{k_n}) = -1.66641.$$

If one holds a long position on the stock worth \$10 million, then the estimated VaR with probability 5% is  $\$10,000,000 \times 0.0166641 = \$166,641$ .  
If we choose  $n = 21$ , the estimated VaR is \$184,127.