Econometrics for Financial Time Series

Chapter 7: 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.
- This chapter introduces methods for calculating VaR (value at risk) and the statistical theories behind these methods.

Value at Risk

- 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.

- 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

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

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.

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

$$\begin{aligned} p &= \Pr[\Delta V(l) \geq c_p] = 1 - \Pr[\Delta V(l) \leq c_p] \\ &= 1 - F_l(c_p); \\ VaR &= c_p \times \text{amount of position} \end{aligned}$$

  - A loss results when we observe $\Delta V(l) > 0$.
  - VaR typically assumes a positive value when p is small.
  - The same as the definition of VaR for a long position if $-\Delta V(l)$ is used instead.

- VaR can be calculated once we know the distribution function.
- We use log returns r_t in calculating VaR because they are the log-differences of asset values. That is, $\Delta V(l) = \ln V_{t+l} - \ln V_t$.

Value at Risk

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 (omitted here)

RiskMetrics

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

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

- Let

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

  Then,

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

- (This part is optional.) Since $E(r_{t+i}r_{t+j} \mid \mathfrak{F}_t) = 0$ [1] $(i, j > 0,\ i \neq j)$, we have

$$Var\left(r_t[k] \mid \mathfrak{F}_t\right) = \sum_{i=1}^{k} Var\left(r_{t+i} \mid \mathfrak{F}_t\right).$$

Since $E(r_t^2 \mid \mathfrak{F}_{t-1}) = \sigma_t^2$ by definition, $Var\left(r_{t+1} \mid \mathfrak{F}_t\right) = \sigma_{t+1}^2$. Moreover, for $i \geq 2$

$$\begin{aligned} Var\left(r_{t+i} \mid \mathfrak{F}_t\right) &= E(r_{t+i}^2 \mid \mathfrak{F}_t) \\ &= E(E(r_{t+i}^2 \mid \mathfrak{F}_{t+i-1}) \mid \mathfrak{F}_t)\ (\mathfrak{F}_{t+i-1} \supset \mathfrak{F}_t) \\ &= E(\sigma_{t+i}^2 \mid \mathfrak{F}_t). \end{aligned}$$

[1] Suppose $i > j$. Then, $E(r_{t+i}r_{t+j} \mid \mathfrak{F}_t) = E\left(E(r_{t+i}r_{t+j} \mid \mathfrak{F}_{t+i-1})\mathfrak{F}_t\right) = E\left(r_{t+j}\sigma_{t+i}E(\epsilon_{t+i} \mid \mathfrak{F}_{t+i-1})\mathfrak{F}_t\right) = 0$.

- Thus,

$$Var\left(r_t[k] \mid \mathfrak{F}_t\right) = \sum_{i=1}^{k} E(\sigma_{t+i}^2 \mid \mathfrak{F}_t). \tag{1}$$

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

$$E(\sigma_{t+i}^2 \mid \mathfrak{F}_t) = E(\sigma_{t+i-1}^2 \mid \mathfrak{F}_t)$$

since $E(\sigma_{t+i-1}^2(\epsilon_{t+i-1}^2 - 1) \mid \mathfrak{F}_t) = 0$. This implies

$$E(\sigma_{t+k}^2 \mid \mathfrak{F}_t) = ... = E(\sigma_{t+1}^2 \mid \mathfrak{F}_t) = \sigma_{t+1}^2. \tag{2}$$

We infer from (1) and (2)

$$Var\left(r_t[k] \mid \mathfrak{F}_t\right) = k\sigma_{t+1}^2.$$

RiskMetrics

- 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.[2] That is,

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

and

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

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

RiskMetrics

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[a] is

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

Therefore, The 5% quantile[b] of the conditional distribution $r_{9191} \mid \mathfrak{F}_{9190}$ is -1.65$\times\sqrt{0.000336 =}-0.03025$.

[a] $\hat{\sigma}_{9190}^2[1]$ is a forecast of σ_{9191}^2.

[b] The $p-$th quantile of $F_l(x)$, x_p, is defined by

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

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."

RiskMetrics

- 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

$$\begin{aligned} r_t &= \phi_0 + \sum_{i=1}^{p} \phi_i r_{t-i} + a_t - \sum_{j=1}^{q} \theta_j a_{t-j}; \\ a_t &= \sigma_t \epsilon_t; \\ \sigma_t^2 &= \alpha_o + \sum_{i=1}^{u} \alpha_i a_{t-i}^2 + \sum_{j=1}^{v} \beta_j \sigma_{t-j}^2. \end{aligned}$$

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

$$\begin{aligned} \hat{r}_t[1] &= \phi_0 + \sum_{i=1}^{p} \phi_i r_{t+1-i} - \sum_{j=1}^{q} \theta_j a_{t+1-j}; \\ \hat{\sigma}_t^2[1] &= \alpha_o + \sum_{i=1}^{u} \alpha_i a_{t+1-i}^2 + \sum_{j=1}^{v} \beta_j \sigma_{t+1-j}^2. \end{aligned}$$

Econometric modelling

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

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

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

- Alternatively, one may assume a t-distribution for ϵ_t.

Econometric modelling

Example

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

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

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

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

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

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

Econometric modelling

- The k-step ahead forecast of r_t is

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

- Using the MA representation of r_t

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

we have[3]

$$\begin{aligned} r_t(l) &= E(r_{t+l} \mid \mathfrak{F}_t) \\ &= E(\mu + a_{t+l} + \psi_1 a_{t+l-1} + \psi_2 a_{t+l-2} + ... \mid \mathfrak{F}_t) \\ &= \mu + \psi_l a_t + \psi_{l+1} a_{t-1} + ... \end{aligned}$$

[3] $E(a_{t+l} \mid \mathfrak{F}_t) = E\left[E(a_{t+l} \mid \mathfrak{F}_{t+l}) \mid \mathfrak{F}_t\right] = E\left[\sigma_{t+l} E(\epsilon_{t+l} \mid \mathfrak{F}_{t+l}) \mid \mathfrak{F}_t\right] = 0.$

Econometric modelling

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

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

- 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

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

with $\psi_0 = 1$.

Econometric modelling

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

$$Var(e_t[k] \mid \mathfrak{F}_t) = \sigma_t^2(k) + (1+\psi_1)^2\sigma_t^2(k-1) + ... + \left(\sum_{i=0}^{k-1}\psi_i\right)^2 \sigma_t^2(1),$$

where $\sigma_t^2(k) = E(\sigma_{t+k}^2 \mid \mathfrak{F}_t)$.

Econometric modelling

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,

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

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

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

and

$$Var(e_t[k] \mid \mathfrak{F}_t) = \sigma_t^2(k) + \sigma_t^2(k-1) + ... + \sigma_t^2(1).$$

Econometric modelling

Example

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

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

These relations give

$$Var(e_t[k] \mid \mathfrak{F}_t) = \frac{\alpha_0}{1-\phi}\left[k - \frac{1-\phi^k}{1-\phi}\right] + \frac{1-\phi^k}{1-\phi}\sigma_t^2(1),$$

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

$$r_{t+k} \mid \mathfrak{F}_t \sim N(k\mu, Var(e_t[k] \mid \mathfrak{F}_t)).$$

Quantile estimation

- No distributional assumption is required.
- Let $r_1, \ldots, r_n$ 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

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

- Assume $r_1, \ldots, r_n$ are i.i.d. with pdf $f(\cdot)$. 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 pth 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, for large n,

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

Therefore, the quantile x_p can be estimated by

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

Quantile estimation

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.

Quantile estimation

- Advantages: simplicity and no distributional assumptions
- Disadvantages:

  1. The distribution of the return r_t remains unchanged.
  2. The returns are independent.
  3. For extreme quantiles (i.e., when p is close to zero or unity), the empirical quantiles are not efficient estimates of the theoretical quantiles.
  4. 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

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

- Assume that the returns r_t are i.i.d. 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

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

- $F_{n,1}(x) \rightarrow 0$ if $x \leq l$ because $F(x) = 0$.
- $F_{n,1}(x) \rightarrow 1$ if $x > l$ because $F(x) > 0$ $(1 - F(x) < 1)$.
- These are degenerate distributions, not useful for statistical analysis.

- The extreme value theory is concerned with finding two sequences β_n and α_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. (β_n : location parameter, α_n : scale parameter).

Extreme value theory

- Let the limiting distribution of $r_{(1*)}$ be $F_*(x)$. It is given by

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

  for $x < -1/k$ if $k < 0$ and for $x > -1/k$ if $k > 0$.
- k : shape parameter $\alpha = -1/k$: tail index
- α is coming from the tail property of the underlying distribution.
- Depending on $k = 0$, $k < 0$ and $k > 0$, we obtain three types of distributions (Gumbel, Fréchet and Weibull).

Empirical estimation

- Estimate k, β_n and α_n.
- Assume $T = ng$. Divide the data as

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

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

$$\begin{aligned} r_{n,i} &= min_{1 \leq j \leq n}\{r_{(i-1)n+j}\},\ i = 1, ..., g. \\ &\quad \text{(the minimum of the } i-\text{th group sample)} \end{aligned}$$

- 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 the pdf of $r_{(1*)}$, $f_*(x)$. Denoting this as $f(r_{n,i})$, the likelihood function is written as.

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

Nonlinear estimation procedures can then be used to obtain maximum likelihood estimates of k_n, β_n and α_n.

- Suppose that the MLEs k_n, β_n and α_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,

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

or

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

An extreme value approach to VaR

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

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

- The quantile r_n^* is the VaR that is based on the extreme value theory for the subperiod minima. This is used to 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.

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

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

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

- The VaR is r_n^*.
- Consequently, for a given small probability p, the VaR of holding a long position in the asset underlying the log return r_t is

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

An extreme value approach to VaR

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 $p = 0.05$,

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

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.