13) Non-parametric estimation methods for VaR

1) Historical method

Suppose $x_1, x_2, \ldots, x_n$ are the losses.

Arrange the data from the smallest to the largest:

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

smallest loss   largest loss

Then

$$\widehat{VaR}_p(X) = x_{(i)} \quad \text{if } p \in \left(\frac{i-1}{n}, \frac{i}{n}\right]$$

Ex 1

Data : $+5, -2, 3, 0, 1$

$$\begin{array}{ccccc} -2 & 0 & 1 & 3 & 5 \\ \| & \| & \| & \| & \| \\ x_{(1)} & x_{(2)} & x_{(3)} & x_{(4)} & x_{(5)} \end{array}$$

$\widehat{VaR}_{0.4}(X) = 0$ because $0.4 \in \left(\frac{1}{5}, \frac{2}{5}\right]$

$\widehat{VaR}_{0.9}(X) = 5$ because $0.9 \in \left(\frac{4}{5}, \frac{5}{5}\right]$

2) Bootstrap method

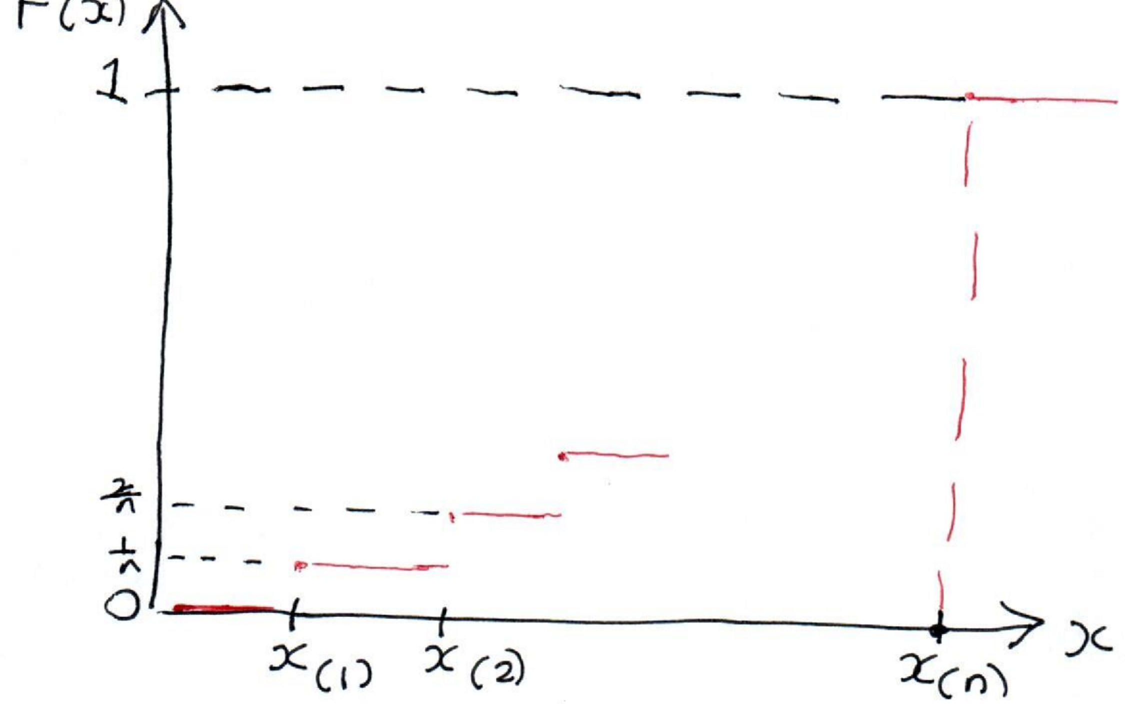
(Bootstrap is due to B. Efron at Stanford)

Suppose $x_1, \dots, x_n$ are the losses.

Compute the empirical CDF of data

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^{n} I\{x_i \le x\}$$

Where $I\{\cdot\}$ denotes the indicator function.



The procedure for computing VaR:

- simulate B independent random samples each of size n from $\hat{F}$

- Compute $\widehat{VaR}_p$ by the historical method for each of the B samples, resulting in $\widehat{VaR}_p^{(1)}, VaR_p^{(2)}, \ldots, VaR_p^{(B)}$.

- Take $\widehat{VaR}_p(X) = \text{Mean}(\widehat{VaR}_p^{(1)}, \ldots, VaR_p^{(B)})$
  or $\widehat{VaR}_p(X) = \text{Median}(\widehat{VaR}_p^{(1)}, \ldots, VaR_p^{(B)})$.

3) Jackknife method

Suppose $x_1, \dots, x_n$ are the data on losses. The procedure is as follows:

- compute VaR by the historical method for $x_2, x_3, \dots, x_n$. Let $VaR_p^{(1)}$ denote the estimate.
- compute VaR by the historical method for $x_1, x_3, \dots, x_n$. Let $VaR_p^{(2)}$ denote the estimate.

⋮

- compute VaR by the historical method for $x_1, x_2, \dots, x_{n-1}$. Let $VaR_p^{(n)}$ denote the estimate.
- Take

$$\widehat{VaR_p}(X) = \text{Mean}\left(VaR_p^{(1)}, \dots, VaR_p^{(n)}\right)$$

or $$\widehat{VaR_p}(X) = \text{Median}\left(VaR_p^{(1)}, \dots, VaR_p^{(n)}\right).$$

4) Kernel method

Suppose $x_1, x_2, \ldots, x_n$ are data on losses.

The kernel CDF of the data is

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^{n} G\left(\frac{x - x_i}{h}\right)$$

where

h = bandwidth,

$$G(x) = \int_{-\infty}^{x} K(u)\,du$$

and $K(\cdot)$ = kernel PDF (usually chosen as $K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$).

h controls the smoothness of $\hat{F}$. Larger the value of h more smooth $\hat{F}$ will become.

The $\widehat{VaR}_p$ can be computed as

- the root of $\hat{F}(x) = p$

- $$\widehat{VaR}_p(X) = \frac{\sum_{i=1}^{n} \hat{F}\left(\frac{i-\frac{1}{2}}{n} - p\right) x_{(i)}}{\sum_{i=1}^{n} \hat{F}\left(\frac{i-\frac{1}{2}}{n} - p\right)}$$

where $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ are the data arranged in increasing order.