LECTURE

19 NOVEMBER

9:00-10:00AM

MATH3/4/68181

b) Non-parametric estimation methods

1) Historical method

Data $-$ $x_1, x_2, \dots, x_n$

Order the data from smallest to largest

$$\underbrace{\boxed{x_{(1)}}}_{\text{smallest}} \leq x_{(2)} \leq \cdots \leq \underbrace{\boxed{x_{(n)}}}_{\text{largest}}$$

Then

$$\widehat{VaR}_p(X) = x_{(i)}$$

if $p \in \left(\frac{i-1}{n}, \frac{i}{n}\right]$.

Ex X = Losses

Data on losses as

8 $\quad$ -2 $\quad$ 2 $\quad$ 0 $\quad$ 5

Order the data

$$\begin{array}{ccccc} -2 & 0 & 2 & 5 & 8 \\ \| & \| & \| & \| & \| \\ x_{(1)} & x_{(2)} & x_{(3)} & x_{(4)} & x_{(5)} \end{array}$$

$$\widehat{VaR}_{0.2}(X) = x_{(1)} = -2$$

$$\widehat{VaR}_{0.9}(X) = x_{(5)} = 8$$

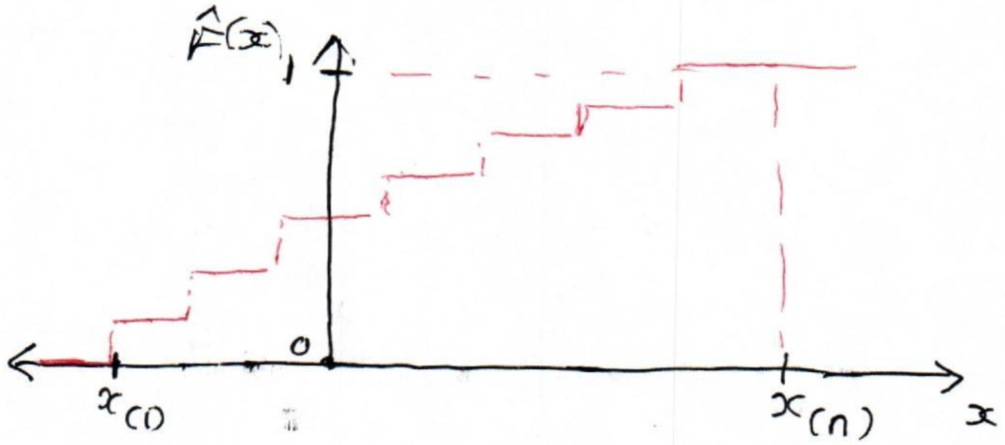
2) <u>Bootstrap</u> <u>Method</u>

due to <u>Efron</u>

Data $x_1, x_2, \ldots, x_n$

Compute the empirical CDF

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



i) simulate B samples from $\hat{F}$

ii) estimate VaR_p for each of the samples by the historical method, yielding $\widehat{VaR}_p^{(j)}$, $j = 1, 2, \ldots, B$

iii) Compute $\widehat{VaR}_p$ as either

$$\widehat{VaR}_p = \text{Mean}\left(\widehat{VaR}_p^{(1)}, \ldots, \widehat{VaR}_p^{(B)}\right)$$

or

$$\widehat{VaR}_p = \text{Median}\left(\widehat{VaR}_p^{(1)}, \ldots, VaR_p^{(B)}\right)$$

3) Jackknife method

Data $x_1, x_2, \dots, x_n$

(i) estimate VaR_p by the historical method for $x_2, x_3, \dots, x_n$.
Call this $\widehat{VaR}_p^{(1)}$.

(ii) estimate VaR_p by the historical method for $x_1, x_3, \dots, x_n$.
Call this $\widehat{VaR}_p^{(2)}$.

⋮

(iii) estimate VaR_p by the historical method for $x_1, x_2, \dots, x_{n-1}$.
Call this $\widehat{VaR}_p^{(n)}$

(iv) Compute $\widehat{VaR}_p$ as either

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

or

$$\widehat{VaR}_p = \text{Median}\left(\widehat{VaR}_p^{(1)}, \dots, \widehat{VaR}_p^{(n)}\right).$$

4) Kernel method

Let $F(\cdot)$ denote the CDF of $X = \text{Loss}$. The kernel estimate of $F(\cdot)$ is

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

where $x_1, x_2, \ldots, x_n$ is a random sample on X

h = bandwidth

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

$K(\cdot)$ = kernel PDF usually chosen as the PDF of $N(0,1)$

The estimate of $VaR_p(X)$ can be computed as

- solve $\hat{F}(x) = p$ for x

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

where $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ are the ordered data.

Expected shortfall

If X is a continuous RV

$$ES_p(X) = \frac{1}{p}\int_0^p VaR_t(X)\,dt$$

Properties

(i) $ES_p(0) = 0$

(ii) $X < Y \Rightarrow ES_p(X) < ES_p(Y)$

(iii) $ES_p(X+c) = ES_p(X) + c$

(iv) $ES_p(cX) = c\,ES_p(X)$

(v) $ES_p(X+Y) \leq ES_p(X) + ES_p(Y)$

Proof of (ii)

We know if $X < Y$ then

$$VaR_p(X) < VaR_p(Y) \quad \forall p$$

$$\Rightarrow \int_0^p VaR_t(X)\,dt < \int_0^p VaR_t(Y)\,dt \quad \forall p$$

$$\Rightarrow \frac{1}{p}\int_0^p VaR_t(X)\,dt < \frac{1}{p}\int_0^p VaR_t(Y)\,dt \quad \forall p$$

$$\Rightarrow ES_p(X) < ES_p(Y) \quad \forall p$$

Proof of (iii)

We know

$$VaR_p(X+c) = VaR_p(X) + c \quad \forall p$$

$$\Rightarrow \int_0^p VaR_t(X+c)\,dt = \int_0^p [VaR_t(X)+c]\,dt \quad \forall p$$

$$= \int_0^p VaR_t(X)\,dt + cp \quad \forall p$$

$$\Rightarrow \frac{1}{p}\int_0^p VaR_t(X+c)\,dt = \frac{1}{p}\int_0^p VaR_t(X)\,dt + c \quad \forall p$$

$$\Rightarrow ES_p(X+c) = ES_p(X) + c \quad \forall p.$$

Proof of (iv)

We know

$$VaR_p(cX) = c\,VaR_p(X) \quad \forall p$$

$$\Rightarrow \int_0^p VaR_t(cX)\,dt = c\int_0^p VaR_t(X)\,dt \quad \forall p$$

$$\Rightarrow \frac{1}{p}\int_0^p VaR_t(cX)\,dt = \frac{c}{p}\int_0^p VaR_t(X)\,dt \quad \forall p$$

$$\Rightarrow ES_p(cX) = c\,ES_p(X).$$

Estimation methods for expected shortfall

a) Parametric estimation methods

b) Non-parametric " "

c) Semi-parametric " "

a), b) for Level 3

a), b), c) for Levels 4, 6

a) Parametric estimation methods

1) Normal distribution method

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

$$ES_p(X) = \mu + \frac{\sigma}{p} \int_0^p \Phi^{-1}(t)\, dt$$

If $x_1, x_2, \ldots, x_n$ is a random sample on X then

$$\widehat{ES}_p(X) = \bar{x} + \frac{\hat{\sigma}}{p} \int_0^p \Phi^{-1}(t)\, dt$$

where $\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$.

$\widehat{ES}_p(X)$ is a biased estimator but consistent.

2) Uniform distribution method

Suppose $X = \text{Loss} \sim \text{Uni}[a, b]$

$$ES_p(X) = a + \frac{p}{2}(b-a)$$

If $x_1, x_2, \ldots, x_n$ is a random sample on X then

$$\widehat{ES}_p(X) = \min(x_1, \ldots, x_n) + \frac{p}{2}\left[\max(x_1, \ldots, x_n) - \min(x_1, \ldots, x_n)\right]$$

$\widehat{ES}_p(X)$ is biased but consistent for $ES_p(X)$.

3) Power function distribution method

Suppose $X = \text{Loss}$ has PDF $f(x) = a x^{a-1}$, $0 < x < 1$.

$$ES_p(X) = \frac{p^{\frac{1}{a}}}{\frac{1}{a} + 1}$$

If $x_1, x_2, \ldots, x_n$ is a random sample on X then

$$\widehat{ES_p}(X) = \frac{p^{\frac{1}{\hat{a}}}}{\frac{1}{\hat{a}} + 1}$$

where $\hat{a} = -\frac{n}{\sum_{i=1}^{n} \log x_i}$.

4) Weibull distribution method

Suppose $X = \text{Loss}$ has CDF $F(x) = 1 - e^{-\left(\frac{x}{\theta}\right)^{\beta}}$.

We know

$$VaR_p(X) = \theta\left[-\log(1-p)\right]^{\frac{1}{\beta}}.$$

So,

$$ES_p(X) = \frac{\theta}{p}\int_0^p \left[-\log(1-t)\right]^{\frac{1}{\beta}} dt$$

$$\boxed{\begin{aligned} \text{Set } \quad y &= -\log(1-t) \\ t &= 1 - e^{-y} \\ \frac{dt}{dy} &= e^{-y} \end{aligned}}$$

$$= \frac{\theta}{p}\int_0^{-\log(1-p)} y^{\frac{1}{\beta}} e^{-y}\, dy$$

$$= \frac{\theta}{p}\,\gamma\left(1+\frac{1}{\beta},\ -\log(1-p)\right)$$

$$\boxed{\begin{aligned} \gamma(a, x) = \int_0^x t^{a-1} e^{-t}\, dt \\ \text{Incomplete gamma function} \end{aligned}}$$

Hence,

$$\widehat{ES}_p(X) = \frac{\hat{\theta}}{p}\,\gamma\left(1+\frac{1}{\hat{\beta}},\ -\log(1-p)\right)$$

where $\hat{\theta}$ and $\hat{\beta}$ are the solutions of

$$\hat{\theta} = \frac{\bar{x}}{\Gamma\left(1 + \frac{1}{\hat{\beta}}\right)}$$

and

$$\frac{(\bar{x})^2}{s^2} = \frac{\left[\Gamma\left(1+\frac{1}{\hat{\beta}}\right)\right]^2}{\Gamma\left(1+\frac{2}{\hat{\beta}}\right) - \left[\Gamma\left(1+\frac{1}{\hat{\beta}}\right)\right]^2}$$

where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2.$$

b) Non-parametric estimation methods

1) Historical method

Data : $x_1, x_2, \ldots, x_n$

Order the data as

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

Then

$$\widehat{ES}_p(X) = \frac{1}{[np]} \sum_{i=1}^{[np]} x_{(i)}$$

where $[x]$ denotes the largest integer less than or equal to x.

Ex $[4.4] = 4, \quad [5.1] = 5$

Ex

$$-2 \quad 8 \quad 9 \quad -10 \quad 1$$

$$\begin{array}{ccccc} -10 & -2 & 1 & 8 & 9 \\ \| & \| & \| & \| & \| \\ x_{(1)} & x_{(2)} & x_{(3)} & x_{(4)} & x_{(5)} \end{array}$$

$$\widehat{ES}_{0.4}(X) = \frac{1}{2} \sum_{i=1}^{2} x_{(i)}$$

$$= \frac{1}{2} (-10 - 2)$$

$$= -6.$$