LECTURE

16 NOVEMBER

10:00-11:00AM

MATH4/68181

Estimation Methods for Expected Shortfall

a) Parametric estimation methods

b) Non-parametric estimation methods

c) Semi-parametric estimation methods

a) and b) for Level 3

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

c) Semi-parametric estimation methods

1) Heavy tail method

Suppose X = returns of losses.

Assume that

$$P(X < -x) \sim x^{-\alpha} L(x)$$

as $x \to \infty$, where $\alpha > 0$ and $L(\cdot)$ is such that

$$\frac{L(tx)}{L(t)} \longrightarrow 1$$

as $t \to \infty$.

"$\sim$" means "behaves as"

Embrechts et al (2005) showed that
(Professor at ETH)

$$\widehat{ES}_p(X) = \frac{1}{p} \int_0^p \exp\left[\left(\frac{l_{n,p}}{nq} \right)^{\frac{1}{\hat{\alpha}_{n, l_{n,p}}}} x_{l_{n,p}} \right] dq$$

$$- 1$$

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

$$l_{n,p} = [n(p + 0.05)],$$

$$\hat{\alpha}_{n,l} = \left[\frac{1}{l} \sum_{i=1}^{l} \log\left(\frac{x_{(i)}}{x_{(l)}} \right) \right]^{-1}$$

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

$$\boxed{\widehat{ES}_p(X) = \frac{1}{p}\int_0^p \exp\left[\left(\frac{l_{n,p}}{nq}\right)^{\frac{1}{\hat{\alpha}_{n,l_{n,p}}}} x_{l_{n,p}}\right] dq - 1} \quad \text{— (1)}$$

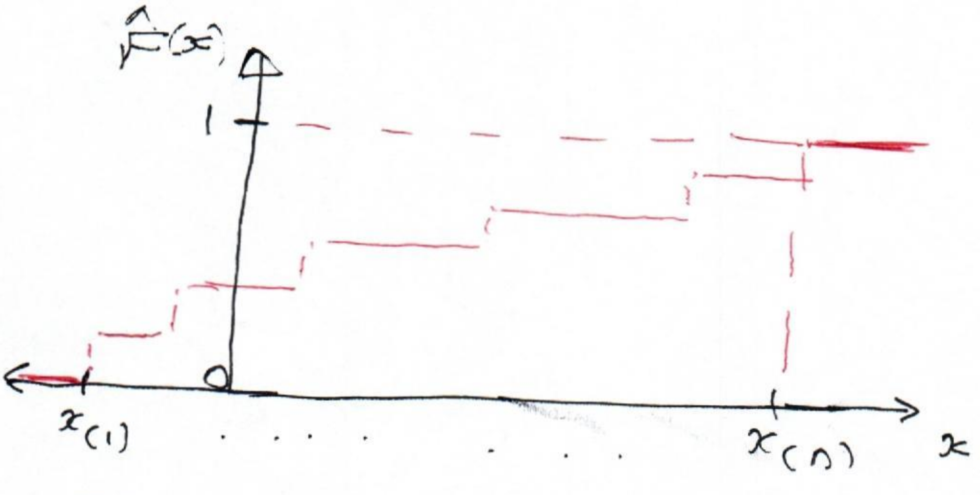
2) Method due to Necir et al (2010)

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

Ordered data : $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$

The empirical CDF

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



The estimator of ES_p is

$$\widehat{ES}_p = \frac{1}{p}\int_{\frac{k}{n}}^{p} \hat{F}(t)\, dt + \frac{k\, x_{(n-k)}}{np(1-\hat{\alpha})} \quad \text{— (2)}$$

where $1 \leq k \leq n$ and

$$\hat{\alpha} = \left[\frac{1}{k}\sum_{i=1}^{k} \log\left(\frac{x_{(i)}}{x_{(k)}}\right)\right]^{-1}.$$

(1) is semi-parametric since α is a parameter. The remaining components in (1) are based only on the data.

Similar reasons for (2).

Suppose

X_1 = stock returns for company A

X_2 = " " " " B

What is the probability $X_1 < X_2$?

Suppose

X_1 = stock returns for company A_1

X_2 = " " " " A_2

$\vdots$

X_k = " " " " A_k

What is the probability $P(X_1 < X_2 < \cdots < X_k)$

Two variables case

If X_1 and X_2 are independent RVs

$$P(X_1 < X_2) = \int_{-\infty}^{\infty} F_{X_1}(x) \, f_{X_2}(x) \, dx$$

where $F_{X_1}(\cdot)$ denotes the CDF of X_1 and $f_{X_2}(\cdot)$ denotes the PDF of X_2.

If X_1 and X_2 are dependent RVs

$$P(X_1 < X_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{x_2} f_{X_1, X_2}(x_1, x_2) \, dx_1 \, dx_2$$

where f_{X_1, X_2} denotes the joint PDF of (X_1, X_2).

Ex 1

Suppose $X_1 \sim N(\mu_1, \sigma_1^2)$

and $X_2 \sim N(\mu_2, \sigma_2^2)$

are independent RVs.

$$P(X_1 < X_2)$$

$$= P(X_1 - X_2 < 0)$$

$$= P\left(N(\mu_1, \sigma_1^2) - N(\mu_2, \sigma_2^2) < 0\right)$$

$$= P\left(N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2) < 0\right)$$

$$= P\left(\frac{N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} < \frac{0 - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)$$

$$= P\left(N(0, 1) < \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)$$

$$= \Phi\left(\frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right)$$

Suppose $x_{1,1}, x_{1,2}, \ldots, x_{1,n}$ is a random sample on X_1

$x_{2,1}, x_{2,2}, \ldots, x_{2,n}$ is a random sample on X_2

The MLEs of μ_1, μ_2, σ_1^2 and σ_2^2 are

$$\hat{\mu}_1 = \frac{1}{n}\sum_{i=1}^{n} x_{1,i}$$

$$\hat{\mu}_2 = \frac{1}{n}\sum_{i=1}^{n} x_{2,i}$$

$$\widehat{\sigma_1^2} = \frac{1}{n}\sum_{i=1}^{n} (x_{1,i} - \hat{\mu}_1)^2$$

$$\widehat{\sigma_2^2} = \frac{1}{n}\sum_{i=1}^{n} (x_{2,i} - \hat{\mu}_2)^2 .$$

Hence,

$$\widehat{P(X_1 < X_2)} = \Phi\left(\frac{\hat{\mu}_2 - \hat{\mu}_1}{\sqrt{\widehat{\sigma_1^2} + \widehat{\sigma_2^2}}}\right).$$