B) Non-parametric estimation methods for ES

1) Historical method

Suppose $x_1, \ldots, x_n$ are the losses.
Order the losses as

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

smallest loss

largest loss

Then

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

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

Ex

$$[6.1] = 6$$
$$[5.9] = 5$$
$$[0.5] = 0$$

Ex 1

losses are : 9, -10, 2, 0, 5, -5

Order them as

$$\begin{array}{cccccc} -10 & -5 & 0 & 2 & 5 & 9 \\ \| & \| & \| & \| & \| & \| \\ x_{(1)} & x_{(2)} & x_{(3)} & x_{(4)} & x_{(5)} & x_{(6)} \end{array}$$

$$\widehat{ES}_{0.2}(X) = \frac{1}{[1.2]} \sum_{i=1}^{[1.2]} x_{(i)}$$

$$= \sum_{i=1}^{1} x_{(i)} = -10$$

$$\widehat{ES}_{0.9}(X) = \frac{1}{[5.4]} \sum_{i=1}^{[5.4]} x_{(i)}$$

$$= \frac{1}{5} \sum_{i=1}^{5} x_{(i)}$$

$$= \frac{-10-5+0+2+5}{5}$$

$$= -\frac{8}{5}.$$

2) Bootstrap method

Suppose $x_1, x_2, \ldots, x_n$ are the losses. Compute the empirical CDF as

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

The procedure for estimating ES_p is as follows:

- simulate B independent random samples each of size n from $\hat{F}$
- Use the historical method to estimate ES for each of B samples, resulting in $\widehat{ES}_p^{(1)}, \widehat{ES}_p^{(2)}, \ldots, \widehat{ES}_p^{(B)}$.
- estimate $\widehat{ES}_p$ as

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

or $\widehat{ES}_p = \text{median}\left(\widehat{ES}_p^{(1)}, \ldots, \widehat{ES}_p^{(B)}\right)$.

3) Jackknife method

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

- Use historical method to estimate ES for $x_2, x_3, \ldots, x_n$, resulting in $\widehat{ES}_p^{(1)}$
- Use historical method to estimate ES for $x_1, x_3, \ldots, x_n$, resulting in $\widehat{ES}_p^{(2)}$

$\vdots$

- Use historical method to estimate ES for $x_1, x_2, \ldots, x_{n-1}$, resulting in $\widehat{ES}_p^{(n)}$.
- estimate $\widehat{ES}_p$ as

$$\widehat{ES}_p = \text{mean}\left(\widehat{ES}_p^{(1)}, \ldots, \widehat{ES}_p^{(n)}\right)$$
$$\widehat{ES}_p = \text{median}\left(\widehat{ES}_p^{(1)}, \ldots, \widehat{ES}_p^{(n)}\right).$$

4) Kernel method

Suppose $x_1, \ldots, x_n$ are the losses.
The kernel estimator of ES is

$$\widehat{ES}_p = \frac{1}{np} \sum_{i=1}^{n} x_i \, A_h\left(\hat{q}(p) - x_i\right)$$

where

$$\hat{q}(p) = \sum_{i=1}^{n} \left[\int_{\frac{i-1}{n}}^{\frac{i}{n}} K_h(t-p)\,dt \right] x_{(i)},$$

$$A_h(u) = A\left(\frac{u}{h}\right),$$

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

$$K_h(u) = \frac{1}{h} K\left(\frac{u}{h}\right),$$

$K(\cdot) =$ the kernel PDF (usually, $K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$),

$h =$ bandwith

and $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ are losses arranged in increasing order.

5) Richardson's method

Richardson was professor at Univ of Manchester. His photo can be found at the ground floor of ATB

Suppose $x_1, x_2, \ldots, x_n$ are the losses. The procedure is as follows

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

(ii) simulate $x_1, \ldots, x_N$ from $\hat{F}$

(iii) Use the historical method to estimate ES for $x_1, \ldots, x_N$

(iv) repeat (ii) & (iii) 1000 times

(v) compute

$$m_N = \frac{1}{1000} \sum_{i=1}^{1000} \widehat{ES}_{N,i}$$

where $\widehat{ES}_{N,i}$ denotes the historical estimate of ES at the i^{th} iteration.

(vi) Set

$$s_j = m_{N_j}$$

for $j = 1, 2, \dots, k+1$ for some k and $N_1, N_2, \dots, N_{k+1}$ [<u>eg</u> $k = 1$, $N_1 = 100$, $N_2 = 1000$].

(vii) Compute

$$\widehat{ES}_p = \frac{\begin{vmatrix} s_1 & s_2 & \cdots & s_{k+1} \\ 1 & \frac{1}{2} & \cdots & \frac{1}{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ 1^k & \left(\frac{1}{2}\right)^k & \cdots & \left(\frac{1}{k+1}\right)^k \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & \frac{1}{2} & \cdots & \frac{1}{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ 1^k & \left(\frac{1}{2}\right)^k & \cdots & \left(\frac{1}{k+1}\right)^k \end{vmatrix}}.$$