

< BRR >

Stratum	Weight	Population mean	Population variance	Sample observations	Sample mean
1	W_1	\bar{Y}_1	S_1^2	y_{11}, y_{12}	\bar{y}_1
2	W_2	\bar{Y}_2	S_2^2	y_{21}, y_{22}	\bar{y}_2
.
.
.
h	W_h	\bar{Y}_h	S_h^2	y_{h1}, y_{h2}	\bar{y}_h
.
.
.
L	W_L	\bar{Y}_L	S_L^2	y_{L1}, y_{L2}	\bar{y}_L

$$(\sum W_h = 1)$$

$$d_h := y_{h1} - y_{h2}$$

$\{y_{h1}, y_{h2}\}$ 의 sample variance

$$= \frac{1}{2-1} [(y_{h1} - \bar{y}_h)^2 + (y_{h2} - \bar{y}_h)^2] = \frac{1}{2} d_h^2$$

∴ S_h^2 의 추정량으로 쓰일 수 있음. But, 매우 부정확하다.

(i.e. $\text{Var}(\frac{1}{4} \sum_{h=1}^L W_h^2 d_h^2)$ 이 매우 크다)

우리의 추정량

$$: \bar{y}_{st} = \sum_{h=1}^L W_h \bar{y}_h$$

$$\text{Var}(\bar{y}_{st}) = \frac{1}{2} \sum_{h=1}^L W_h^2 S_h^2$$

∴ 추정량의 분산은

최대한 정확히

추정해본다.

Stratum	Weight	Population mean	Population variance	Sample observations	Sample mean
1	W_1	\bar{Y}_1	S_1^2	y_{11}, y_{12}	\bar{y}_1
2	W_2	\bar{Y}_2	S_2^2	y_{21}, y_{22}	\bar{y}_2
.
.
.
h	W_h	\bar{Y}_h	S_h^2	y_{h1}, y_{h2}	\bar{y}_h
.
.
.
L	W_L	\bar{Y}_L	S_L^2	y_{L1}, y_{L2}	\bar{y}_L

$$\bar{y}_{si} = \sum_{h=1}^L W_h y_{hi(i)}$$

2는 1, 2, $i(h) = 1 \text{ or } 2$

즉, 각 층에서 두 줄

하나를 임의로 골라

weighted sum을 해줄 것.

\bar{y}_{si} 의 값으로 2L 가수가 존재.

$$\bar{y}_{si} - \bar{y}_{st} = \frac{1}{2} (\pm W_1 d_1 \pm \dots \pm W_L d_L) \quad (\text{각 항의 } \frac{1}{2})$$

$$\Rightarrow E[(\bar{y}_{si} - \bar{y}_{st})^2]$$

$$= E[E[(\bar{y}_{si} - \bar{y}_{st})^2 | d_1, \dots, d_L]]$$

$$= E\left[\frac{1}{4} \sum W_h d_h^2\right] = \frac{1}{2} \sum W_h^2 S_h^2 = \text{Var}(\bar{y}_{st})$$

$$\Rightarrow \text{Var}(\bar{y}_{st}) \text{의 추정량} = \frac{1}{K} \sum_{i=1}^K (\bar{y}_{si} - \bar{y}_{st})^2 \text{을}$$

사용할 수 있다.

2L 개 층에서 적당히 K개
골라서 평균.

< Bootstrap >

