Jack Knife

Leave - one - out samples

$X_{-i} = \{x_1, \dots, x_{i-1}, x_{i+1} \dots, x_N\}$

$\Theta = \{\hat{\theta}_1, \dots, \hat{\theta}_N\}$

Non-parametric Bootstrap

sampling with replacement (size N)

$\Theta = \{\hat{\theta}_1, \dots, \hat{\theta}_B\}$

$B = 1000$

Parametric Bootstrap

1) Estimate the parameter set (whole sample) $\hat{\theta}$

2) Generate pseudodata from some distr.

$X_i \sim \mathrm{Law}(\hat{\theta})$

↳ subsample of size n

fit parameters to resamples

3) $\Theta = \{\hat{\theta}_1, \dots, \hat{\theta}_B\}$

$B = 1000$

(A) Assumes the model is true

Pros of Non-Parametric Bootstrap

- Easy to code
- Easily extendended to
  - estimation of covariance
  - mult. parameter models
  - predictions

Cons -//- :

- Inference limited to sample (for in-sample inference only)
- If sample size is small ($Var_{sample} < Var_{pop.}$)

Pair Bootstrap

$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$, $\varepsilon_i \sim N(\mu, \sigma^2)$

- (X_i, y_i) – pairs
  Resample data frames
  with indexes $\{1, \dots, N\}$
  drawn replacement
- $\hat{\beta}^{b}_{OLS}$

Bootstrapping Residuals (Wild)

1) $\hat{y}_i = f(X, \hat{\beta}_{OLS})$
   $\hookrightarrow \hat{\varepsilon}_i = y_i - \hat{y}_i$
   assume: $p(\hat{\varepsilon}_i) = 1/n \iff$ i.i.d
2) Subsamples: X – are fixed
   $y_i = f(X, \hat{\beta}_{OLS}) + \hat{\varepsilon}^{b}_i$ ← obtained from $\hat{\varepsilon}$ using sampling with replacement
   $\hat{\beta}^{*b}_{OLS, i}$

Problem: heteroscedasticity

e.g.:

$y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$, $\varepsilon_i \sim N(0, \sigma^2 \cdot x_i)$

Wild weighted Bootstrap:

Difference:

drawn from t

$$y_i^{*b} = f(x_i, \hat{\beta}_{OLS}) + \frac{t_i^* \cdot \hat{\varepsilon}_i}{\sqrt{1 - h_{ii}}}$$

$h_{ii} = x_i^T (X^T X) x_i$ - i-th leverage

Parametric Bootstrapping

1) $\hat{y}_i = f(X, \hat{\beta}_{OLS})$

assume: $\varepsilon_i \sim Norm(0, \hat{\sigma}^2)$

2) Subsamples: X - are fixed

$y_i = f(X, \hat{\beta}_{OLS}) + \hat{\varepsilon}_i$

drawn from $N(0, \hat{\sigma}^2)$

$\hat{\beta}^{*b}_{OLS,i}$

Jackknife for regression:

1) leave one observation out

2) Calculate $\hat{\beta}^*_{OLS}$