Bootstrapping

Given data $((x_i, y_i))_{i=1}^{n}$ and we want to estimate

$S(z)$

Construct the empirical distribution function $\hat{f}$.

Idea: For $i=1,\dots,n$

- Resample $z_i^* \sim \hat{F}$
- Calculate $S_i \equiv S(z_i^*)$

Return distribution $\{S_i\}_{i=1}^{B}$

$$\widehat{Err}_{Boot} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{B}\sum_{b=1}^{B} L(y_i, \hat{f}_b^*(x_i))$$

Bagging – Bootstrap Aggregation

Define an estimator

$$\hat{f}_{AVE}(x) = \sum_{i=1}^{B} \underbrace{w_i}_{\text{weights}} \underbrace{\hat{f}_i^*(x)}_{\text{Bootstrapped estimate}}$$

General Aggregation models

Suppose we have models $\{m_i\}_{i=1}^{M}$ that we wish to

Combine to reduce uncertainty.

Assume that m_i is parameterized by ζ_i

$$\mathbb{P}(\zeta \mid z) = \frac{\mathbb{P}(\zeta, z)}{z} = \sum_{i=1}^{m} \frac{\mathbb{P}(\zeta, z, m_i)}{\mathbb{P}(z)} = \sum_{i=1}^{m} \frac{\mathbb{P}(\zeta, z, m_i)}{\mathbb{P}(z, m_i)} \frac{\mathbb{P}(z, m_i)}{\mathbb{P}(z)}$$

$$= \sum_{i=1}^{m} \underbrace{\mathbb{P}(\zeta \mid z, m_i)}_{\text{model param posterior}} \underbrace{\mathbb{P}(m_i \mid z)}_{\text{Post. Prob of } m}$$

$$\mathbb{E}[\zeta \mid z] = \sum_{i=1}^{m} \mathbb{E}[\zeta \mid z, m_i]\, \mathbb{P}(m_i \mid z)$$