An Introduction to the Bootstrap Method

1 Introduction

The bootstrap is a resampling technique introduced by Bradley Efron in 1979. It allows statisticians to estimate the sampling distribution of an estimator by resampling with replacement from the original data. This method is particularly useful when the theoretical distribution of the estimator is complex or unknown.

2 The Basic Bootstrap Algorithm

Given a dataset $X = \{x_1, x_2, \ldots, x_n\}$ drawn independently from an unknown distribution F, and an estimator $\hat{\theta} = s(X)$ (such as the sample mean or median), the bootstrap procedure involves the following steps:

1. **Resampling**:
   - Generate B bootstrap samples $X^{*1}, X^{*2}, \ldots, X^{*B}$.
   - Each bootstrap sample X^{*b} is obtained by sampling n observations *with replacement* from the original data X.
2. **Computing Bootstrap Replicates**:
   - For each bootstrap sample X^{*b}, compute the bootstrap replicate $\hat{\theta}^{*b} = s(X^{*b})$.
3. **Estimating the Sampling Distribution**:
   - Use the distribution of the bootstrap replicates $\{\hat{\theta}^{*1}, \hat{\theta}^{*2}, \ldots, \hat{\theta}^{*B}\}$ to approximate the sampling distribution of $\hat{\theta}$.

3 Bootstrap Estimate of Standard Error

One primary use of the bootstrap is to estimate the standard error of an estimator $\hat{\theta}$. The bootstrap estimate of the standard error is calculated as:

$$\hat{\sigma}_{\text{boot}} = \sqrt{\frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{\theta}^{*b} - \bar{\hat{\theta}}^{*} \right)^2}$$

where $\bar{\hat{\theta}}^{*} = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}^{*b}$ is the mean of the bootstrap replicates.

4 Bootstrap Confidence Intervals

Bootstrap methods can construct confidence intervals for the unknown parameter θ using the bootstrap distribution of $\hat{\theta}$. Here are some common approaches:

4.1 Normal Approximation Interval

Assuming that the bootstrap distribution is approximately normal, the $(1-\alpha) \times 100\%$ confidence interval is:

$$\left[\hat{\theta} - z_{1-\alpha/2} \hat{\sigma}_{\text{boot}},\ \hat{\theta} + z_{1-\alpha/2} \hat{\sigma}_{\text{boot}} \right]$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution.

4.2 Percentile Interval

The percentile method uses the empirical quantiles of the bootstrap replicates:

$$\left[\hat{\theta}^*_{(\alpha/2)},\ \hat{\theta}^*_{(1-\alpha/2)}\right]$$

where $\hat{\theta}^*_{(q)}$ is the q-th percentile of the bootstrap replicates.

5 Asymptotic Properties

Under regularity conditions, bootstrap methods have desirable asymptotic properties:

- **Consistency**: As the sample size $n \to \infty$, the bootstrap distribution of $\hat{\theta}^*$ converges in probability to the true sampling distribution of $\hat{\theta}$.
- **Second-Order Accuracy**: Bootstrap confidence intervals often achieve higher-order accuracy compared to traditional methods, especially in small samples.

However, the bootstrap relies on the assumption that the sample is representative of the population and that observations are independent and identically distributed (i.i.d.).

6 Bootstrap Confidence Intervals for Maximum Likelihood Estimators

Maximum Likelihood Estimators (MLEs) are widely used for parameter estimation due to their desirable properties such as consistency and asymptotic normality. However, deriving the exact sampling distribution of an MLE can be complex, especially in finite samples or for complicated models. Bootstrapping provides a practical approach to estimate the sampling distribution and construct confidence intervals for MLEs.

6.1 Applying Bootstrap to MLEs

Given a dataset $X = \{x_1, x_2, \ldots, x_n\}$ and a parametric model with likelihood function $\mathcal{L}(\theta; X)$, the MLE of the parameter θ is obtained by maximizing the likelihood:

$$\hat{\theta}_{\text{MLE}} = \arg\max_{\theta} \mathcal{L}(\theta; X)$$

To use bootstrapping for constructing confidence intervals for θ:

1. **Compute the MLE** $\hat{\theta}_{\text{MLE}}$ from the original data.
2. **Generate Bootstrap Samples**:
   - Create B bootstrap samples $X^{*1}, X^{*2}, \ldots, X^{*B}$ by resampling with replacement from X.
3. **Compute Bootstrap Replicates**:
   - For each bootstrap sample X^{*b}, compute the bootstrap MLE $\hat{\theta}^{*b}_{\text{MLE}}$ by maximizing $\mathcal{L}(\theta; X^{*b})$.
4. **Estimate the Sampling Distribution**:
   - Use the distribution of $\{\hat{\theta}^{*1}_{\text{MLE}}, \hat{\theta}^{*2}_{\text{MLE}}, \ldots, \hat{\theta}^{*B}_{\text{MLE}}\}$ to approximate the sampling distribution of $\hat{\theta}_{\text{MLE}}$.
5. **Construct Confidence Intervals**:
   - Use methods such as the percentile or normal approximation method to construct confidence intervals for θ.

6.2 Example: Estimating the Mean of a Normal Distribution

Suppose we have a sample $X = \{x_1, x_2, \ldots, x_n\}$ from a normal distribution $N(\mu, \sigma^2)$ with unknown mean μ and known variance σ^2.

1. **Compute the MLE**:

$$\hat{\mu}_{\text{MLE}} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

2. **Generate Bootstrap Samples**:
   - Resample n observations with replacement from X to create B bootstrap samples $X^{*1}, X^{*2}, \ldots, X^{*B}$.

3. **Compute Bootstrap MLEs**:
   - For each bootstrap sample X^{*b}, compute:

$$\hat{\mu}^{*b}_{\text{MLE}} = \frac{1}{n}\sum_{i=1}^{n} x_i^{*b}$$

   where x_i^{*b} are the observations in X^{*b}.

4. **Construct Confidence Interval**:
   - Use the percentile method to find the $(\alpha/2)$ and $(1-\alpha/2)$ quantiles of the bootstrap MLEs $\hat{\mu}^{*b}_{\text{MLE}}$.

$$\text{CI}_{\text{boot}} = \left[\hat{\mu}^*_{(\alpha/2)},\ \hat{\mu}^*_{(1-\alpha/2)}\right]$$

6.3 Example: Estimating the Rate Parameter of an Exponential Distribution

Suppose $X = \{x_1, x_2, \ldots, x_n\}$ is a sample from an exponential distribution with density $f(x;\lambda) = \lambda e^{-\lambda x}$ for $x \geq 0$, where $\lambda > 0$ is the rate parameter.

1. **Compute the MLE**:

$$\hat{\lambda}_{\text{MLE}} = \frac{n}{\sum_{i=1}^{n} x_i}$$

2. **Generate Bootstrap Samples**:
   - Resample n observations with replacement from X to create B bootstrap samples.

3. **Compute Bootstrap MLEs**:
   - For each bootstrap sample X^{*b}, compute:

$$\hat{\lambda}^{*b}_{\text{MLE}} = \frac{n}{\sum_{i=1}^{n} x_i^{*b}}$$

4. **Construct Confidence Interval**:
   - Use the bootstrap replicates $\hat{\lambda}^{*b}_{\text{MLE}}$ to construct a confidence interval for λ using the desired method (e.g., percentile method).

6.4 Remarks

- **Non-i.i.d. Data**: For MLEs involving non-i.i.d. data or complex models (e.g., time series, clustered data), adjustments to the bootstrap procedure may be necessary, such as block bootstrapping.
- **Computational Considerations**: Computing MLEs for each bootstrap sample can be computationally intensive, especially for large datasets or complex models. Efficient algorithms or parallel computing can mitigate this issue.
- **Bias Correction**: MLEs can be biased, particularly in small samples. Bootstrapping can help assess and correct for this bias.

7 Example

Suppose we have a sample of size $n = 10$:

$$X = \{2, 4, 5, 6, 7, 8, 9, 10, 12, 14\}$$

We assume that X comes from an exponential distribution with rate parameter λ. We want to estimate λ using MLE and construct a 95% confidence interval using the bootstrap method.

1. **Compute the MLE**:

$$\hat{\lambda}_{\text{MLE}} = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{10}{77} \approx 0.1299$$

2. **Generate** $B = 1000$ **Bootstrap Samples**:
   - Resample $n = 10$ observations with replacement from X to create 1000 bootstrap samples.

3. **Compute Bootstrap MLEs**:
   - For each bootstrap sample X^{*b}, compute:

$$\hat{\lambda}^{*b}_{\text{MLE}} = \frac{10}{\sum_{i=1}^{10} x_i^{*b}}$$

4. **Construct the 95% Confidence Interval**:
   - Use the percentile method to find the 2.5% and 97.5% quantiles of the bootstrap MLEs $\hat{\lambda}^{*b}_{\text{MLE}}$.
   - The confidence interval is:

$$\text{CI}_{\text{boot}} = \left[\hat{\lambda}^{*}_{(2.5\%)},\ \hat{\lambda}^{*}_{(97.5\%)}\right]$$

References

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1), 1-26.