Bootstrap Confidence Intervals

MIT 18.443

Dr. Kempthorne

Spring 2015

Outline

Bootstrap Confidence Intervals

Bootstrap Framework

- Data Model : $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ i.i.d. sample with pdf/pmf $f(x_1, \ldots, x_n \mid \theta) = \prod_{i=1}^{n} f(x_i \mid \theta)$
- Data Realization: $\mathbf{X}_n = \mathbf{x}_n = (x_1, \ldots, x_n)$
- $\hat{\theta}_n$: Estimate of θ given $\mathbf{x}_n = (x_1, \ldots, x_n)$
  ($\hat{\theta}_n$ can be MLE, MOM, or any well-defined estimate)
- θ_0: the true value of the parameter θ.

Exact Confidence Interval

- **Estimate Error:** $\Delta = \hat{\theta}_n - \theta_0 = g(\mathbf{X}_n, \theta_0)$
- **Sampling Distribution of Δ:** $\Delta \sim P_\Delta$, induced by $(\mathbf{X} \mid \theta_0)$.
- Exact confidence interval using Δ as a *pivotal*.
  - Set $\underline{\delta}$ and $\overline{\delta}$ as the $\alpha/2$ and $(1-\alpha/2)$ quantiles of P_Δ
  - $$\begin{aligned} P_\Delta(\underline{\delta} \le \Delta \le \overline{\delta}) &= P_{\mathbf{X}_n|\theta_0}(\underline{\delta} \le \hat{\theta}_n - \theta_0 \le \overline{\delta}) \\ &= P(\hat{\theta}_n - \overline{\delta} \le \theta_0 \le \hat{\theta}_n - \underline{\delta}) \\ &= 1 - \alpha \end{aligned}$$

Bootstrap Confidence Intervals

Approximating P_Δ: Sampling Distribution of

$$\Delta = \hat{\theta}_n - \theta_0 = g(\mathbf{X}_n, \theta_0)$$

- If θ_0 known, then
  - Simulate $\mathbf{X}_n^* \sim \mathbf{X}_n \mid \theta_0$
  - Use simulation distribution of $\Delta^* = g(\mathbf{X}_n^*, \theta_0)$
- θ_0 unknown, then
  - Simulate $\mathbf{X}_n^* \sim \mathbf{X}_n \mid \hat{\theta}_n$
  - Use simulation distribution of $\Delta^* = g(\mathbf{X}_n^*, \hat{\theta}_n)$

Bootstrap Confidence Interval

- Generate B samples from the distribution of $[\mathbf{X}_n \mid \hat{\theta}_n]$
- Compute estimate $\hat{\theta}_j^*$ for each sample j, $j = 1, \ldots, B$.
- Compute sample values: $\Delta_j^* = (\hat{\theta}_j^* - \hat{\theta}_j)$, $j = 1, \ldots, B$.
- Approximate $\underline{\delta}$ and $\overline{\delta}$ with appropriate quantiles of $\{\Delta_j^*\}$
- Plug $\hat{\theta}_n$, $\hat{\underline{\delta}}$, and $\hat{\overline{\delta}}$ into *pivotal* confidence interval formula

18.443 Statistics for Applications
Spring 2015