Resampling Methods for statistical inference

Bootstrap methods

Eric Gilleland
Research Applications Laboratory,
National Center for Atmospheric Research

What is a bootstrap?

From the expression ...

pull oneself up by one's bootstraps

Thought to derive from the
"Adventures of Baron Munchausen"
by Rudolph Erich Raspe

Efron and Tibshirani (1998), p. 5

What is bootstrapping in statistics?

Basic Principle

Recreate the relation between the *population* and the *sample* by considering the sample as an epitome of the underlying population.

Resample (with replacement) from the given sample to obtain the bootstrap sample $\longrightarrow$ analog of the given sample (Lahiri, 2003).

Compute statistic of interest from the resampled data, and aggregate over multiple realizations (resamples) of the data.

What is bootstrapping in statistics?

Numerous textbooks on the subject. Efron and Tibshirani (1998); Davison and Hinkley (1997) both give easy-to-comprehend introductions to bootstrapping. Lahiri (2003) gives a thorough treatment of dealing with dependent data with the bootstrap.

The seminal paper by Singh (1981) gives a theoretical proof that under iid situations, the bootstrap outperforms the classic normal results, but also shows that when the iid assumption fails, so does the bootstrap approach.

References in Lahiri (2003) cited for when to use $m < n$ instead of $m = n$ in the bootstrap samples.

What is bootstrapping in statistics?

IID Bootstrap

Assume $X_1, \ldots, X_n \overset{iid}{\sim} F$.

Resample from $\mathcal{X}_n = \{X_1, \ldots, X_n\}$ *with replacement* multiple times, say R times.

For example, one resample from $\mathcal{X}_3$ might be $\mathcal{X}_3^* = \{X_1, X_1, X_3\}$

Non-parametric bootstrap:
Sample from $\mathcal{X}_n$ using the EDF $\tilde{F}(z) = \frac{\{\# z_i \leq z\}}{n}$

Parametric bootstrap:
Sample from $\hat{F}(z)$, where a distribution is assumed for F.

What is bootstrapping in statistics?

- The common distribution of the X_i^*'s is given by the empirical distribution of the original sample.
- Estimate the statistic(s) of interest, θ, from $\mathcal{X}_i^*$ for each iteration. Call it $\hat{\theta}^*$.
- Estimate the unknown distribution of θ, or a function thereof. (e.g., using the EDF of $\hat{\theta}^*$).

Notation: Let G be the distribution function for θ, and $\hat{G}$ its bootstrap estimate.

Example: Standard Error of ETS

Suppose we are interested in the Equitable Threat Score, and want to estimate its standard error. For each bootstrap sample, calculate $\hat{\theta}^* = \text{ETS}^*$. The standard error is then given by

$$\hat{\text{se}}_{\text{boot}} = \{\sum_{r=1}^{R}[\hat{\theta}^{*(r)} - \hat{\theta}^{*(\cdot)}]^2/(R-1)\}^{1/2}$$

where $\hat{\theta}^{*(\cdot)} = \sum_{r=1}^{R} \hat{\theta}^{*(r)}/R$.

Bootstrap Confidence Intervals

- Normal approximation (using μ^* and σ^*)
- Percentile Interval
- Bootstrap-t Intervals
- BC_a
- ABC

Bootstrap Confidence Intervals

Percentile Interval

$1 - 2\alpha$ *percentile interval* is defined by

$$[\hat{\theta}^{\%,\text{low}}, \hat{\theta}^{\%,\text{up}}] = [\hat{G}^{-1}(\alpha), \hat{G}^{-1}(1-\alpha)]$$

$$= [\hat{\theta}^{*(\alpha)}, \hat{\theta}^{*(1-\alpha)}]$$

Bootstrap Confidence Intervals

Percentile Interval: Pros and Cons

Pros

- Easy to understand
- Easy to compute
- Transformation-respecting
  (i.e., $[\hat{\theta}^{\%,\text{low}}, \hat{\theta}^{\%,\text{up}}] = [\hat{m}(\hat{\theta}^{\%,\text{low}}), m(\hat{\theta}^{\%,\text{up}})]$,
  for m monotone increasing).
- range-preserving

Bootstrap Confidence Intervals

Percentile Interval: Pros and Cons

Cons: Does not work well in general

- Coverage performance is poor
  (i.e., for a 95% interval, the target miscoverage is 2.5%)
- Cannot account for a biased estimate
  (i.e., $\hat{\theta} \sim N(\theta + \text{bias}, \hat{\text{se}}^2)$)
- Cannot account for non-constant variance

Bootstrap Confidence Intervals

Bias-corrected and accelerated – BC_a

$1-2\alpha$ interval of *intended* coverage is given by

$$(\hat{\theta}^{\text{low}}, \hat{\theta}^{\text{up}}) = (\hat{\theta}^{*(\alpha_1)}, \hat{\theta}^{*(\alpha_2)})$$

$$\alpha_1 = \boldsymbol{\Phi}\left(\hat{z}_0 + \frac{\hat{z}_0 + z^{(\alpha)}}{1-\hat{a}(\hat{z}_0 + z^{(\alpha)})}\right)$$

α_2 is defined similarly, but with $1-\alpha$ in place of α.

Bootstrap Confidence Intervals

Bias-corrected and accelerated – BC_a: Pros and Cons

Pros

- transformation-respecting
- Much better coverage than percentile: 2^{nd} order accurate.

Cons

- Computationally burdensome!!
- Calculation of $\hat{a}$ does not use the bootstrap replicates.
- R must be very large in order to reduce the sampling error sufficiently.

Bootstrap Confidence Intervals

Approximate Bootstrap Confidence Intervals – ABC

Approximates the BC_a endpoints analytically (without resorting to further Monte Carlo simulations).

Pros

- Computationally more efficient than BC_a
- 2^{nd}-order accurate (good coverage)
- transformation-respecting

Cons

Requires $\hat{\theta} = s(\boldsymbol{x})$ to be defined smoothly in $\boldsymbol{x}$
(e.g., will not work for the median!).

Bootstrap and Dependent data

Bootstrap Assumptions: Less *stringent* than most other approaches.

However, care still needs to be taken. For example,

Singh (1981) showed that for $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} F$, the IID Bootstrap approximation for $\Pr\{T_n \leq \cdot\}$ is far more accurate than the classical normal approximation.

But, in the face of dependent data

$$\lim_{n \longrightarrow \infty} [\Pr_*(T_n^* \leq x) - \Pr(T_n \leq x)] \neq 0$$

Bootstrap and Dependent data

Numerous approaches for handling dependent data. A few are . . .

- Parametric Bootstrap (Efron and Tibshirani, 1998, pp. 53–55)
  - Bootstrap based on IID Innovations (Lahiri, 2003, pp. 23-24)
- Block Bootstrap methods (Wilks, 1997)
  - Nonoverlapping block (NBB) (Carlstein, 1986)
  - Moving Block (MBB) (Kunsch, 1989; Liu and Singh, 1992)
  - Generalized Block (GBB) (Lahiri, 2003, pp. 31–33)
  - Circular Block (CBB) (Politis and Romano, 1992)
  - Stationary Block (SB) (Politis and Romano, 1994)
- Subsampling (Carlstein, 1986)
- Transformation-based (TBB) (Hurvich and Zeger, 1987)
- Sieve Bootstrap (Lahiri, 2003, pp. 41–43)

Block Bootstrap Methods

Main idea is to restore the dependence structure in the data by resampling contiguous blocks of data instead of individual points.

Conclusions

I have a lot of reading to do.

That's all. Complete references on next slide.

References

Carlstein, E., 1986. The use of subseries methods for estimating the variance of a general statistic from a stationary time series. Annals of Statistics 14, 1171–1179.

Davison, A., Hinkley, D., 1997. Bootstrap Methods and Their Application. Cambridge University Press.

Efron, B., Tibshirani, R., 1998. An Introduction to the Bootstrap. Chapman and Hall, Boca Raton.

Hurvich, C., Zeger, S., 1987. Frequency domain bootstrap methods for time series. Statistics and Operations Research Working Paper, New York University, New York.

Kunsch, H., 1989. The jacknife and the bootstrap for general stationary observations. Annals of Statistics 17, 1217–1261.

Lahiri, S., 2003. Resampling Methods for Dependent Data. Springer-Verlag, New York.

Liu, R., Singh, K., 1992. Exploring the Limits of Bootstrap. Wiley, New York, Ch. Moving blocks jacknife and bootstrap capture weak dependence. eds. Lepage, R. and Billard, L., pp. 225–248.

Politis, D., Romano, J., 1992. Exploring the Limits of Bootstrap. Wiley, New York, Ch. A circular block resampling procedure for stationary data. eds. Lepage, R. and Billard, L., pp. 263–270.

Politis, D., Romano, J., 1994. The stationary bootstrap. J. American Statistic. Assoc. 1989, 1303–1313.

Singh, K., 1981. On the asymptotic accuracy of efron's bootstrap. Annals of Statistics 9 (6), 1187–1195.

Wilks, D., 1997. Resampling hypothesis tests for autocorrelated fields. J. Climate 10, 65–82.