The Bootstrap and Jackknife

Motivation

In scientific research, we're often interested in estimating some unknown parameter θ, e.g., mean weight of a certain strain of mice, heritability index, a genetic component of variation, a mutation rate, etc.

Two key questions need to be addressed:

1. How do we estimate θ?
2. Given an estimator for θ, how do we estimate its precision/accuracy?

• We assume Question 1 can be reasonably well specified by the researcher.

• Question 2, for our purposes, will be addressed via the estimation of the estimator's **standard error**

Standard Error

Suppose we want to estimate a parameter θ of a distribution. e.g., the mean/median

- We select a random sample and calculate $\hat{\theta}$, our estimate of *q*.
- Any function of our sample is also random.
- So our estimate, $\hat{\theta}$, is random.
- If we collect a new sample, we get a new estimate. Same for another sample, and another...

- Therefore, our estimate has a distribution, called the *sampling distribution*.

The standard deviation of that distribution is the **standard error.**

Bootstrap Motivation

Challenges

• Estimating the standard error, even for relatively simple estimators (e.g., ratios and other non-linear functions of estimators) can be quite challenging

  • Solutions to most estimators are mathematically intractable or too complicated to develop (with or without advanced training in statistical inference)

• However, great strides in computing in the last 25 years have made these calculations more feasible.

• We will investigate how the bootstrap allows us to obtain robust estimates of precision.

Limitations of the Central Limit Theorem

Estimating the precision of the sample mean

- The central limit theorem gives us the standard error of X:

$$\widehat{se}\left[\overline{X}\right] = \sqrt{\widehat{\sigma}^2 / n}$$

where

$$\widehat{\sigma}^2 = \frac{\sum_{i=1}^{n} (X_i - \overline{X})^2}{n-1}$$

- However, the CLT only applies to means -- it does not extend to other estimators. The bootstrap is a more general approach that applies to medians, diversity indices, ratios…

Bootstrap Algorithm

- (Assume the sample dataset accurately reflects the population from which it is drawn)
- Generate a large number of "bootstrap" samples by resampling (with replacement) from your dataset
- Resample with the same structure as used in the original sample
- Calculate your estimator $\hat{\theta}$ for each of the bootstrap samples
- Calculate the standard deviation of the bootstrapped estimates

Bootstrap Estimation Example: Median

What is the variance of the sample median?

=> Use the bootstrap!

	Data	Median
Original sample:	{1, 5, 8, 3, 7}	5
Bootstrap 1:	{1, 7, 1, 3, 7}	3
Bootstrap 2:	{7, 3, 8, 8, 3}	7
Bootstrap 3:	{7, 3, 8, 8, 3}	7
Bootstrap 4:	{3, 5, 5, 1, 5}	5
Bootstrap 5:	{1, 1, 5, 1, 8}	1
.		
.		
.		
Bootstrap B (=1000)	{1, 5, 7, 7, 8}	7

Bootstrap Estimation Example: Median

Bootstrapped estimates of the standard error for sample median (cont.)

- Descriptive statistics for the sample medians from 1000 bootstrap samples

B	1000
Mean	4.964
Standard Deviation	**1.914**
Median	5
Minimum, Maximum	1, 8
25th, 75th percentile	3, 7

- We estimate the standard error for the sample median as 1.914

Bootstrap Estimation Example: Relative Risk

Bootstrapped estimates of the standard error for sample relative risk

RR = P[Disease | Exposed]/P[Disease | Not exposed]

Cross-classification of Framingham Men by high systolic blood pressure and heart disease

	Heart Disease	
High Systol BP	No	Yes
No	915	48
Yes	322	44

The sample estimate of the relative risk is:

RR = (44/366)/(48/963) = 2.412

Bootstrap Estimation Example cont'd

Bootstrapped estimates of the standard error for the relative risk (cont.)

- Descriptive statistics for the sample relative risks:

B	100000
Bootstrap mean of *RR*	2.464
Bootstrap median of *RR*	2.412
Standard Deviation of *RR*	**0.507**

- The bootstrap standard error for the estimated relative risk is 0.507

Bootstrap Summary

Advantages

- All-purpose, computer intensive method useful for statistical inference.
- Bootstrap estimates of precision do not require knowledge of the theoretical form of an estimator's standard error, no matter how complicated it is.

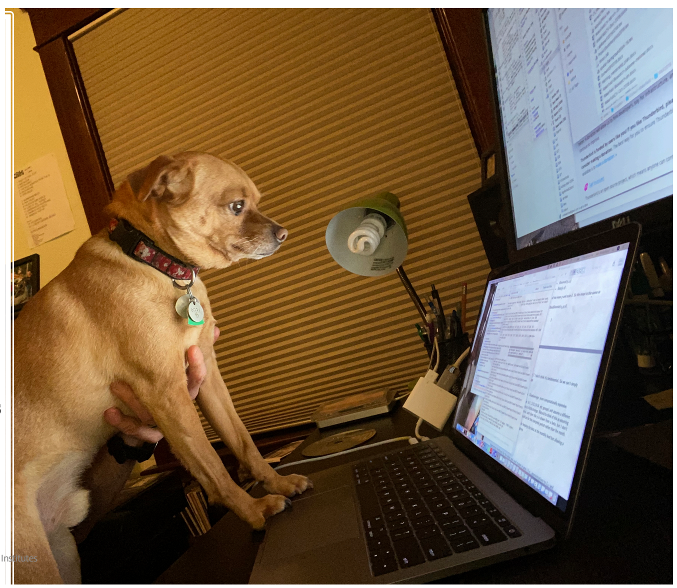
Disadvantages

- Typically not useful for correlated (dependent) data.
- Missing data, censoring, data with outliers are also problematic
- Often used incorrectly

Note that there are many different types of bootstraps: we have only discussed one

Paws-

break time then work on exercises



Question 1

Suppose you are interested in the number of times an experiment works before it fails.
Suppose it failed on the first try, then on the sixth try, then on the first try.
Therefore, the data are: (0, 5, 0).

What is a bootstrapped estimate of the median number of successes before failures, and the standard error in your estimate?

To help, here are 10 bootstrap samples:
{0, 0, 5}
{5, 0, 0}
{5, 0, 5}
{0, 5, 0}
{0, 0, 5}
{5, 5, 0}
{0, 5, 0}
{0, 5, 5}
{0, 5, 5}
{5, 0, 5}

Jackknife

Jackknife Estimation

• The jackknife (or leave one out) method, invented by Quenouille (1949), is an alternative resampling method to the bootstrap.

• The method is based upon sequentially deleting one observation from the dataset, recomputing the estimator, here, $\hat{\theta}_{(i)}$, *n* times. That is, there are exactly *n* jackknife estimates obtained in a sample of size *n*.

• Like the bootstrap, the jackknife method provides a relatively easy way to estimate the precision of an estimator, *q*.

• The jackknife is generally less computationally intensive than the bootstrap

Jackknife Algorithm

Jackknifing

• For a dataset with *n* observations, compute *n* estimates by sequentally omitting each observation from the dataset and estimating $\hat{\theta}$ on the remaining *n* – 1 observations.

• Using the *n* jackknife estimates, $\hat{\theta}_{(1)}, \hat{\theta}_{(2)}, \ldots, \hat{\theta}_{(n)}$,

we estimate the standard error of the estimator as

$$\widehat{se}_{jack} = \sqrt{\frac{n-1}{n}\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \bar{\hat{\theta}}_{(.)})^2}$$

• Unlike the bootstrap, the jackknife standard error estimate will not change for a given sample

Jackknife Summary

Advantages

• Useful method for estimating and compensating for bias in an estimator.
• Like the bootstrap, the methodology does not require knowledge of the theoretical form of an estimator's standard error.
• Is generally less computationally intensive compared to the bootstrap method.

Disadvantages

• The jackknife method is more conservative than the bootstrap method, that is, its estimated standard error tends to be slightly larger.
• Performs poorly when the the estimator is not sufficiently smooth, i.e., a non-smooth statistic for which the jackknife performs poorly is the median.