Notes and Exercises Chapter 12 – Confidence Intervals with Bootstrapping
Math 217
Saidi

Chapter 12

Bootstrapping is another simulation method that is best suited for modeling studies where the data have been generated through random sampling from a population. Bootstrapping provides a computational approach for providing interval estimates for almost any population parameter.

As with randomization tests, our goal with bootstrapping is to understand variability of a statistic. Unlike randomization tests (which modeled how the statistic would change if the treatment had been allocated differently), the bootstrap will model how a statistic varies from one sample to another taken from the population. This will provide information about how different the statistic is from the parameter of interest.

Quantifying the variability of a statistic from sample to sample is a hard problem.

Example 12.1.1 Observed data

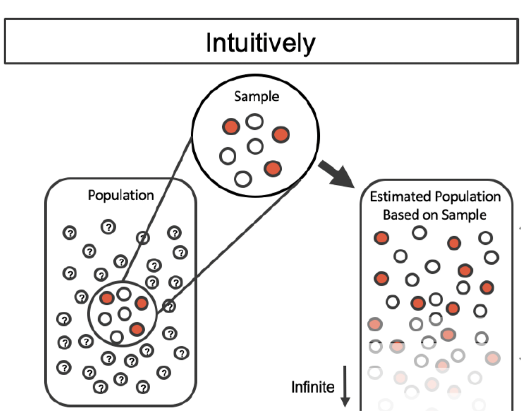
One consultant tried to attract patients by noting the average complication rate for liver donor surgeries in the US is about 10%, but her clients have had only 3 complications in the 62 liver donor surgeries she has facilitated. She claims this is strong evidence that her work meaningfully contributes to reducing complications (and therefore she should be hired!).

Let $\boldsymbol{p}$ **represent the true complication** rate for liver donors working with this consultant. (The "true" complication rate will be referred to as the parameter.) We **estimate $\boldsymbol{p}$ using the data**, and label the **estimate $\boldsymbol{\hat{p}}$.**

The sample proportion for the complication rate is 3 complications divided by the 62 surgeries the consultant has worked on: $\hat{p} = \frac{3}{62} = 0.048$

> Is it possible to assess the consultant's claim (that the reduction in complications is due to her work) using the data?
>
> No. The claim is that there is a causal connection, but the data are observational, so we must be on the lookout for confounding variables. For example, maybe patients who can afford a medical consultant can afford better medical care, which can also lead to a lower complication rate. While it is not possible to assess the causal claim, it is still possible to understand the consultant's true rate of complications.

Figure 12.1 shows how the unknown original population can be estimated by using the sample to approximate the proportion of successes and failures (in our case, the proportion of complications and no complications for the medical consultant).



By taking repeated samples from the estimated population, the variability from sample to sample can be observed. In Figure 12.3 the repeated bootstrap samples are obviously different both from each other and from the original population. Recall that the bootstrap samples were taken from the same (estimated) population, and so the differences are due entirely to natural variability in the sampling procedure.

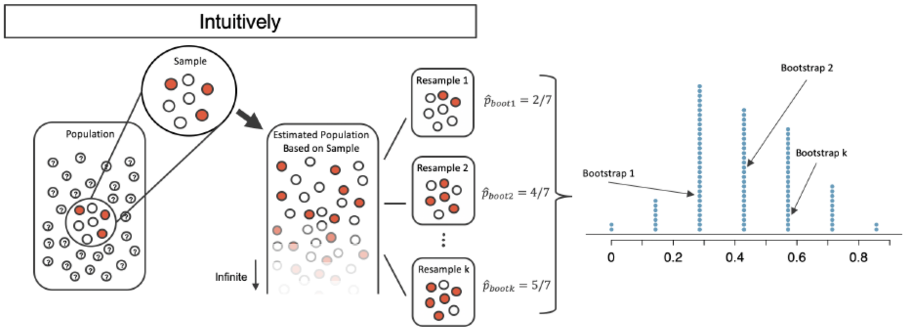
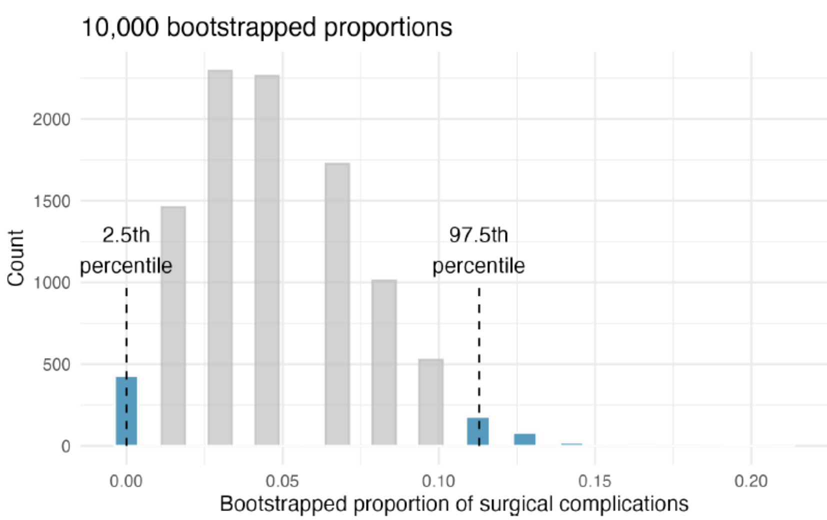


Figure 12.3: The bootstrapped proportion is estimated for each bootstrap sample. The resulting bootstrap distribution (dotplot) provides a measure for how the proportions vary from sample to sample

One simulation isn't enough to get a sense of the variability from one bootstrap proportion to another bootstrap proportion, so we repeat the simulation 10,000 times using a computer. See figure 12.6 below.



We are 95% confident that, in the population, the true probability of a complication is between 0% and 11.3%.

Figure 12.6: The original medical consultant data is bootstrapped 10,000 times. Each simulation creates a sample from the original data where the probability of a complication is $\hat{p} = 3/62$. The bootstrap 2.5 percentile proportion is 0 and the 97.5 percentile is 0.113. The result is: we are confident that, in the population, the true probability of a complication is between 0% and 11.3%.

To solve the example:

```
Given:  x=3 n=62 and the national rate of complications is 10%
```

Calculate a 95% confidence interval.

```
prop.test(x=3, n=62, p=.10, conf.level = .95)
```

Results

```
95 percent confidence interval:
  0.0125801 0.1437904
```

Conclusion

```
We are 95% confident that this consultant's true proportion of complications is
between 1.3% and 14.4%. That is, because the national rate of 10% is included in
this interval, there is no compelling evidence that her rates are lower than the
national rates.
```

12.3 Confidence intervals

A point estimate provides a single plausible value for a parameter. However, a point estimate is rarely perfect; usually there is some error in the estimate. In addition to supplying a point estimate of a parameter, a next logical step would be to provide a *plausible range of values for the parameter.* This plausible range of values is called the **confidence interval.**

Bootstrap sampling is often called sampling with replacement.

Here are the "steps" to creating bootstrap sampling:

- Randomly sample one observation from the original sample of size *n* – if marbles in a bag represent the observations, replace the marble back into the bag so as to keep the population constant.
- Randomly sample a second observation from the original sample of size *n*. Because we sample with replacement (i.e., we don't actually remove the marbles from the bag), there is a 1-in-*n* chance that the second observation will be the same one sampled in the first step!
- Keep going one sampled observation at a time.
- Randomly sample the n^{th} observation from the *n* samples.

A bootstrap sample behaves similarly to how an actual sample from a population would behave, and we compute the point estimate of interest (here, compute $\hat{p}_{boot}$).

95% Bootstrap percentile confidence interval for a parameter *p.*

The 95% bootstrap confidence interval for the parameter p can be obtained directly using the ordered $\hat{p}_{boot}$ values.

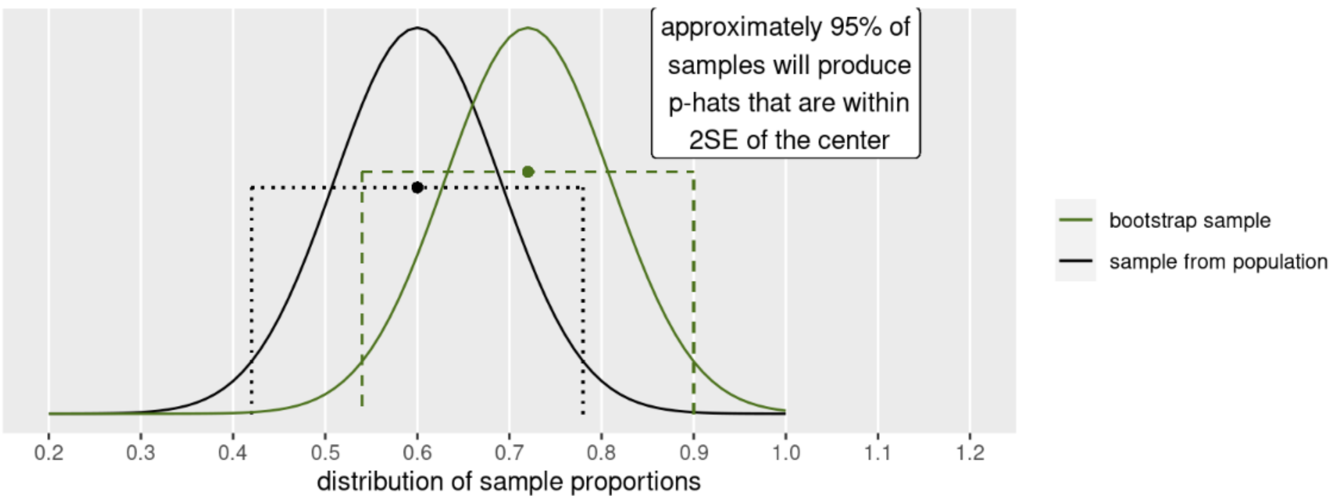
Consider the sorted $\hat{p}_{boot}$ values. Call the **2.5% bootstrapped proportion value "lower"**, and call the **97.5% bootstrapped proportion value "upper".**

The 95% confidence interval is given by: (lower, upper)

Standard Error of the Sample Proportion

The amount of discrepancy between $\bar{y}$ and μ is described using probability terms with the SAMPLING distribution of $\hat{\boldsymbol{p}}$, and the standard deviation of $\hat{\boldsymbol{p}}$ is given as the **standard error of the sample proportion**, which is (this is sometimes called the ***estimated standard error***.

- SD describes the **dispersion** of the data
- SE describes the **unreliability** (due to sampling error) in the mean of the sample as an estimate of the mean of the population. In other words, the SE measures how far $\hat{\boldsymbol{p}}$ is likely to be from the population mean p.



Technical conditions

Methods for creating bootstrap confidence intervals work for **any** statistic and parameter, as long as the following technical conditions hold:

1. the distribution of the statistic is reasonably symmetric and bell-shaped
2. the sample size is reasonably large
3. the sample was representative of the population.

A plot of the bootstrap $\hat{\boldsymbol{p}}$ values will give a good indication for whether the technical conditions are valid.

Homework Chapter 12

1. Review section 12.4 (the chapter review)
2. **Suggested**: from textbook section 12.5 exercises: 2,4,6,8
3. **Suggested**: Unit 4 Tutorial on Foundations of Inference: Sampling Variability

    4 - [Parameters and confidence intervals]