Bootstrapping Regression Example

TABLE 9.2 Copper–nickel alloy data for illustrating methods of obtaining a bootstrap confidence interval for β_1/β_0.

x_i	0.01	0.48	0.71	0.95	1.19	0.01	0.48
y_i	127.6	124.0	110.8	103.9	101.5	130.1	122.0
x_i	1.44	0.71	1.96	0.01	1.44	1.96	
y_i	92.3	113.1	83.7	128.0	91.4	86.2	

Bootstrapping Regression Example

y_i : Corrosion loss in copper-nickel alloy
x_i : iron content

- 13 measurements of corrosion loss (y_i) in copper-nickel alloys, each with a specific iron content (x_i).
- Of interest is the change in corrosion loss in the alloys as the iron content increases, relative to the corrosion loss when there is no iron. Thus, consider the estimation of $\theta = \beta_1/\beta_0$ in a simple linear regression

$$\frac{\text{Change Corrosion loss as the iron increase}}{\text{Corrosion loss when no iron (intercept)}} = \frac{\beta_1}{\beta_0}$$

Bootstrapping Regression Example

Observed Study => Bootstrap with pairs

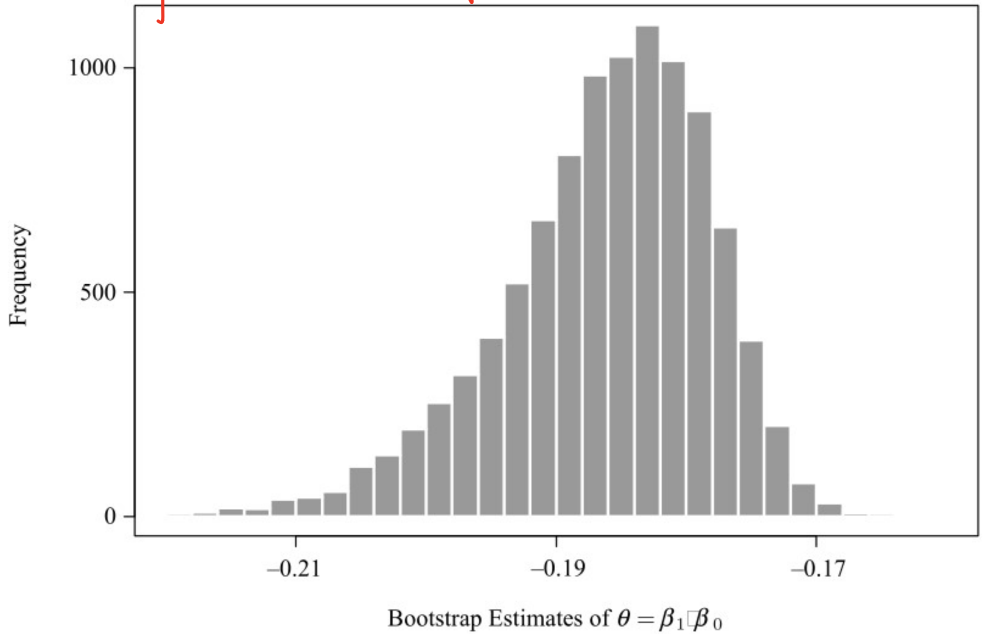


FIGURE 9.1 Histogram of 10,000 bootstrap estimates of β_1/β_0 from the nonparametric paired bootstrap analysis with the copper–nickel alloy data.

1. 2. 3. 4. 5
→ 1 2. 3. 4. 5
→ 1. 2. 3. 4 5

Bootstrap Bias Correction

- A particular interesting choice for bootstrap analysis when $T(F) = \theta$ is the quantity $R(\mathcal{X}, F) = T(\hat{F}) - T(F)$. This represents the bias of $T(\hat{F}) = \hat{\theta}$, and it has mean equal to $E\{\hat{\theta}\} - \theta$. The bootstrap estimate of the bias is

$T(F) = \theta$

$R(\mathcal{X}, F) = T(\hat{F}) - T(F)$ ↳ bias $= E\{\hat{\theta}\} - \theta$

$$\sum_{i=1}^{B} \frac{(\hat{\theta}_i^* - \hat{\theta})}{B} = \bar{\theta}^* - \hat{\theta}.$$

↳ $\frac{1}{B}\sum_{i=1}^{B} \hat{\theta}_i^*$

Bootstrap Sample $\mathcal{X}_1^* \rightarrow \hat{\theta}_1^*$
⋮
$\mathcal{X}_B^* \rightarrow \hat{\theta}_B^*$
only use statistics (parameters of interests)

- An improved bias estimate requires only a little additional effort.
- Let $\hat{F}_j^*$ denote the empirical distribution of the j-th bootstrap pseudo-dataset, and define

$$\bar{F}^*(\mathbf{x}) = \sum_{j=1}^{B} \hat{F}_j^*(\mathbf{x})/B.$$

$\bar{F}^*(x)$ ↳ Statistics (parameter of interest)

→ the mean of empirical distribution

Bootstrap Sample $\mathcal{X}_1^* \rightarrow \hat{F}_1^*$
$\mathcal{X}_2^* \rightarrow \hat{F}_2^*$
⋮
$\mathcal{X}_B^* \rightarrow \hat{F}_B^*$
use the entire empirical distⁿ

Then, $\bar{\theta}^* - T(\bar{F}^*)$ is a better estimate of bias.

Permutation Tests Examples

the distribution of the test statistic under null hypothesis is obtained by calculating all possible values of the test statistic under all possible rearrangement of the observed data points

Consider a medical experiment where rats are randomly assigned to treatment and control groups.

rat — treatment
rat — control group

- The outcome X_i is then measured for the i-th rat.
- Under the null hypothesis, the outcome does not depend on whether a rat was labeled as treatment or control.
- Under the alternative hypothesis, outcomes tend to be larger for rats labeled as treatment.
- A test statistic T measures the difference in outcomes observed for the two groups. For example, T might be the difference between group mean outcomes, having value t_1 for the observed dataset.

Permutation Tests.

1. Calculate test statistics of interests in observed data set.
2. Calculate same test statistics in each permuted data set and get the distribution of test statistics from permuted data set
3. Compare observed test statistics with distribution of permuted test statistics

Empirical p-value

$P(|\text{Permuted test statistic}| > |\text{Observed Test Stat}|)$

If empirical p-value < significance level,
we reject H_0

Observed Dataset

	1.	2.	3.	4,	- - - .	n.
Treatment	1	0	1	0		
Control	0	1	0	1		

For each sample, we have two assign possibilities : T or C.

Total Possible Allocation $2^n - 1$.

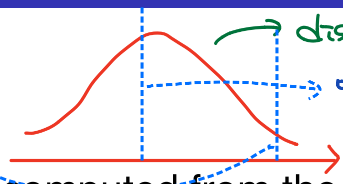
↳ But if we assume the balanced experiment, it will be much less.

Permutation Tests Examples

- Under the null hypothesis, the individual labels “treatment” and “control” are meaningless because they have no influence on the outcome.
- Since they are meaningless, the labels could be randomly shuffled among rats without changing the joint null distribution of the data.
- Shuffling the labels creates a new dataset: Although one instance of each original outcome is still seen, the outcomes appear to have arisen from a different assignment of treatment and control.
- Each of these permuted datasets is as likely to have been observed as the actual dataset, since the experiment relied on random assignment.

Permutation Tests Examples

do not follow null hypothesis.



dist of test stat

obs : follow null hypothesis

- Let t_2 be the value of the test statistic computed from the dataset with this first permutation of labels.

→ we need to consider all possible combination

- Suppose all M possible permutations (or a large number of randomly chosen permutations) of the labels are examined, thereby obtaining t_2, $\cdots$, t_M.

$t_2, \cdots, t_M.$ ⇒ Under the null hypothesis, same distribution with t_1

- Under the null hypothesis, $t_2, \cdots, t_M$ were generated from the same distribution that yielded t_1.
- Therefore, t_1 can be compared to the empirical quantiles of $t_1, \cdots, t_M$ to test a hypothesis or construct confidence limits.

Permutation Tests

from empirical dist of stat constructed by permuted dataset.

- To pose this strategy more formally, suppose that we observe a value t for a test statistic T having density f under the null hypothesis.
- Suppose ~~large values of T indicate that the null hypothesis is false~~.
- Monte Carlo hypothesis testing proceeds by generating a random sample of $M - 1$ values of T drawn from f.
- If the observed value t is the k-th largest among all M values, then the null hypothesis is rejected at a significance level of k/M.
- If the distribution of the test statistic is highly discrete, then ties found when ranking t can be dealt with naturally by reporting a range of p-values.

a rank 1 – 5
b rank 6 – 8
⋮
z rand --. M

the range of p-value from permutation test for test stat a 1/M ~ 5/M

Permutation Tests

- There are a variety of approaches for sampling from the null distribution of the test statistic.
- The permutation approach works because "treatment" and "control" are meaningless labels assigned completely at random and independent of outcome, under the null hypothesis.
- This simple permutation approach can be broadened for application to a variety of more complicated situations.
- In all cases, the permutation test relies heavily on the condition of exchangeability. order doesnot matter $\left(\begin{array}{c} X_1, X_2, \ldots, X_n \\ X_2, X_3, \ldots, X_n, X_1 \end{array}\right) \Rightarrow$ same information
- The data are exchangeable if the probability of any particular joint outcome is the same regardless of the order in which the observations are considered.

Permutation Tests

There are two advantages to the permutation test over the bootstrap.

- If the basis for permuting the data is random assignment, then the resulting p-value is exact (if all possible permutations are considered).
  - For such experiments, the approach is usually called a randomization test.
  - In contrast, standard parametric approaches and the bootstrap are founded on asymptotic theory that is relevant for large sample sizes.

Permutation Tests

use statistics

There are two advantages to the permutation test over the bootstrap.

- Permutation tests are often more powerful than their bootstrap counterparts.
  - However, the permutation test is a specialized tool for making a comparison between distributions, whereas a bootstrap tests hypotheses about parameters, thereby requiring less stringent assumptions and providing greater flexibility.
  - The bootstrap can also provide a reliable confidence interval and standard error, beyond the mere p-value given by the permutation test.
  - The standard deviation observed in the permutation distribution is not a reliable standard error estimate.

Final Exam

Due : Dec. 20. 23:55

# of Question : 4.

Q1, Q2 ⇒ 5.

Q3, Q4 ⇒ 15.

Range : Numerical Optimization → Bootstrap

Q1 (a) Fixed point iteration (X)

Do not discuss with others.

Implement by RCpp (3-4)

R (1-2)

Next Week (12/9)

10 AM – 11:30 AM.

- Final Exam Question
- Check HW 3/4.