Bootstrap Confidence Intervals

Yair Wexler

Based on:

An Introduction to the Bootstrap

Bradley Efron and Robert J. Tibshirani

Chapters 12-13

Introduction

- Chapters 12 and 13 discuss **approximate** confidence intervals to some parameter $\theta = t(F)$.
  - **Chapter 12** - Confidence intervals based on bootstrap “tables”
    - **Bootstrap-t intervals**
  - **Chapter 13** - Confidence intervals based on bootstrap percentiles
    - **Percentile intervals**
- Both chapters discuss **one sample non-parametric bootstrap**.

Bootstrap-t

- Normal theory approximate confidence intervals are based on the distribution of **approximate pivots**:

$$Z = \frac{\hat{\theta} - \theta}{\widehat{se}} \dot{\sim} N(0,1) \qquad\qquad Z = \frac{\hat{\theta} - \theta}{\widehat{se}} \dot{\sim} t_{n-1}$$

- The **bootstrap-t** method uses bootstrap sampling to estimate the distribution of the approximate pivot Z:

$$Z_b^* = \frac{\hat{\theta}_b^* - \hat{\theta}}{\widehat{se}_b^*}$$

  – For $\hat{\theta} = \bar{X}$ and $\widehat{se} = n^{-0.5}s$, a bootstrap-t interval is a Student-t interval.

Bootstrap-t

- Suggested by Efron (1979) and revived by Hall (1988).
- Creates an **empirical distribution table** from which we calculate the desired percentiles.
- Doesn’t rely on normal theory assumptions.
- **Asymmetric** interval (in general).
- “*bootstrap*” R package – **boott()**

Bootstrap-t algorithm

1. Calculate $\hat{\theta}$ from the sample **x**.
2. For each bootstrap replication b=1,…,B:
   1. Generate bootstrap sample **x*b**.
   2. Using *some measure* $\widehat{se}_b^*$ of the standard error of **x*b**, calculate:

$$Z_b^* = \frac{\hat{\theta}_b^* - \hat{\theta}}{\widehat{se}_b^*}$$

3. The bootstrap-t "table" qth quantile is:

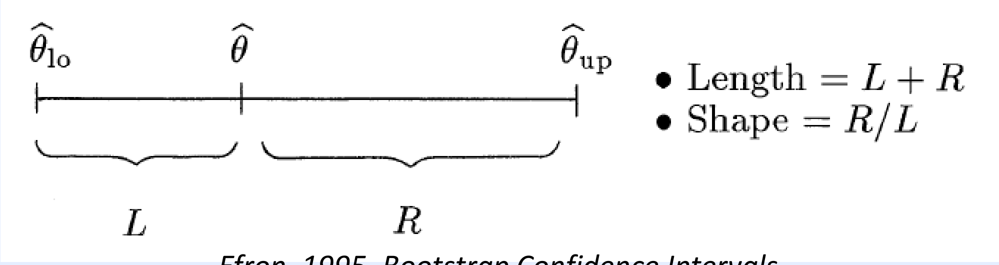
$$\hat{t}_q = Z^*_{(q\cdot B)}$$

4. 100(1-α)% Bootstrap-t confidence interval for θ:

$$\left[\hat{\theta} - \hat{t}_{1-\frac{\alpha}{2}} \cdot \widehat{se}, \hat{\theta} - \hat{t}_{\frac{\alpha}{2}} \cdot \widehat{se}\right]$$

Bootstrap-t vs Normal theory

- **Improved accuracy**:
  - Coverage tend to be closer to 100(1-α)% than in normal or t intervals.
  - Better captures the shape of the original distribution.



Efron, 1995. Bootstrap Confidence Intervals.

- **Loss of generality**:
  - Z table applies to all samples.
  - Student-t table applies to all samples of a fixed size n.
  - Bootstrap-t table is sample specific.

Bootstrap-t vs Normal theory

- **Example**:
  - Confidence intervals to the expected value θ of $X_i \sim exp\left(\frac{1}{\theta}\right)$.
  - Plug-in estimator for θ: $\hat{\theta} = \bar{X}$
  - Plug-in estimator for standard error:

$$\widehat{se} = n^{-1}\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{0.5}$$

  - n=100:

	CI	lo	up	length	shape
1	Normal	0.1045725	0.1529141	0.04834156	1.000000
2	Student-t	0.1042734	0.1532132	0.04893974	1.000000
3	Bootstrap-t	0.1070222	0.1565122	0.04949002	1.278430
4	Exact	0.1068152	0.1582313	0.05141604	1.344762

Bootstrap-t vs Normal theory

- Comparison of coverage:
  - n=15,100,5000

```
> coverage.15
           CI median.width  IQR.width  miss.left  miss.right miss.total coverage
1      Normal     0.1261544 0.05285748 0.10666667 0.003333333 0.11000000 89.00000
2   Student-t     0.1380507 0.05784189 0.09666667 0.003333333 0.10000000 90.00000
3 Bootstrap-t     0.1552218 0.09186576 0.06000000 0.003333333 0.06333333 93.66667
> coverage.100
           CI median.width   IQR.width  miss.left miss.right miss.total coverage
1      Normal    0.05572083 0.009626614 0.04333333 0.01666667 0.06000000 94.00000
2   Student-t    0.05641033 0.009745735 0.04000000 0.01666667 0.05666667 94.33333
3 Bootstrap-t    0.05897387 0.011755449 0.02333333 0.03333333 0.05666667 94.33333
> coverage.5000
           CI median.width    IQR.width  miss.left miss.right miss.total coverage
1      Normal   0.007923375 0.0002099254 0.02000000 0.02333333 0.04333333 95.66667
2   Student-t   0.007925294 0.0002099762 0.02000000 0.02333333 0.04333333 95.66667
3 Bootstrap-t   0.007931130 0.0003914792 0.01666667 0.02333333 0.04000000 96.00000
```

Issues regarding Bootstrap-t

- Bootstrap estimation of $\widehat{se}_b^*$ where there is no formula:
  - B_2 replications for each original replication b=1,…,B.
  - Total number of bootstrap replications: $B*B_2$.
  - Efron and Tibshirani suggest B=1000, B_2=25
    => total of **25,000** bootstrap replications.
- **Not** invariant to transformations.
  - Change of scale can have drastic effects.
  - Some scales are better than others.
- Applicable mostly to **location statistics**.

Bootstrap-t and transformations

- Example: **Fisher-z transformation** *Fisher 1921*
  - If (**X,Y**) has a *bivariate normal distribution* with correlation ρ.

$$\phi = \frac{1}{2} ln\left(\frac{1+\rho}{1-\rho}\right) = \operatorname{arctanh}(\rho)$$

$$\hat{\phi} - \phi \dot{\sim} N\left(0, \frac{1}{n-3}\right)$$

  - An approximate normal CI for ϕ:

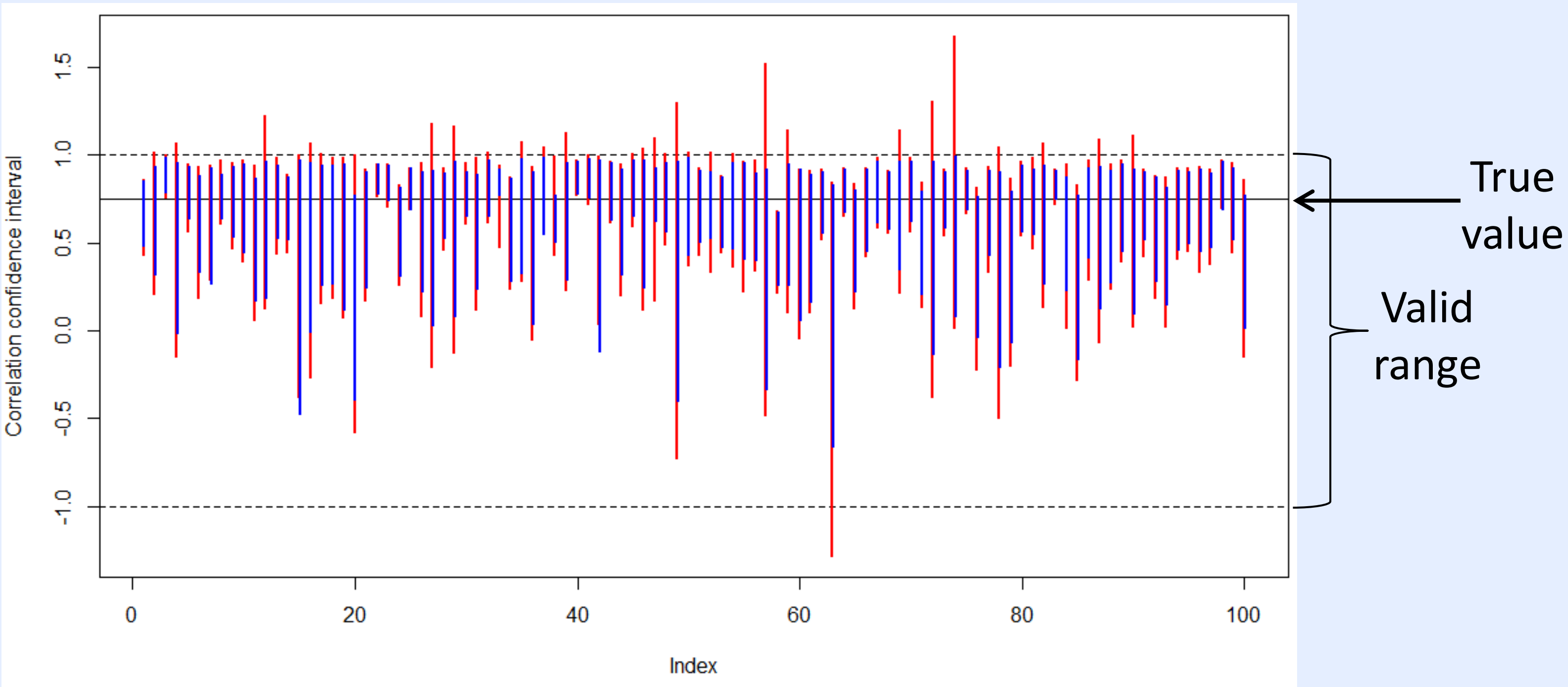
$$\left[\hat{\phi} - Z_{1-\frac{\alpha}{2}} \cdot \frac{1}{\sqrt{n-3}}, \hat{\phi} + Z_{1-\frac{\alpha}{2}} \cdot \frac{1}{\sqrt{n-3}}\right]$$

  - Apply the reverse transformation for an approximate CI for ρ.

$$r = \frac{e^{2\phi} - 1}{e^{2\phi} + 1} = \tanh(\phi)$$

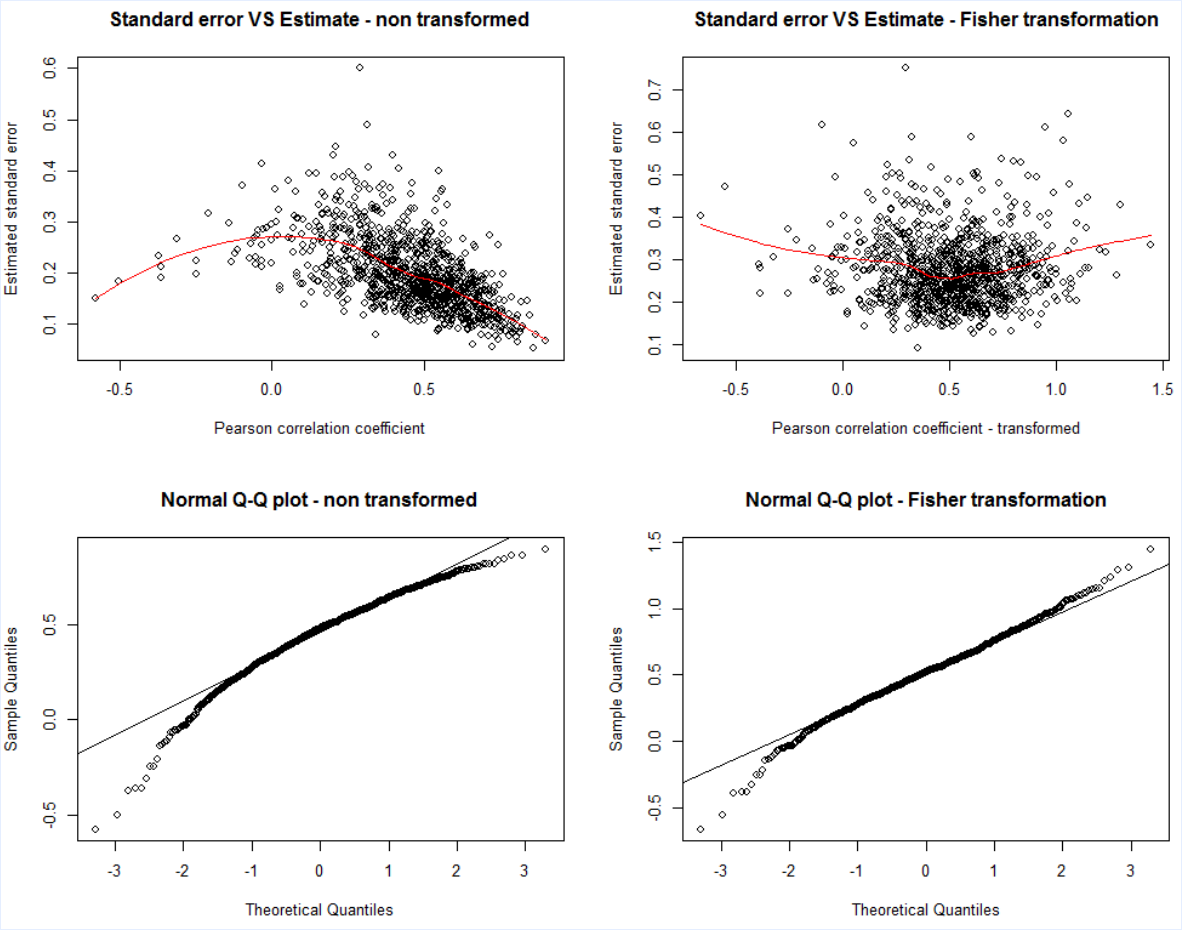
Bootstrap-t and transformations

- Simulation results for bootstrap-t with n=15:
  - **Red**: 95% CI bootstrap-t interval for r directly. (**96% coverage, 33% outside valid range**)
  - **Blue**: 95% CI bootstrap-t interval using Fisher transformation. (93% coverage, 0% outside valid range)



Bootstrap-t and transformations

- Variance stabilization and normalization of the estimate:



Variance stabilization

- In general, it is impossible to achieve both variance stabilization and normalization.
  - Bootstrap-t works better for variance stabilized parameters.
  - Normality is less important.
- In general, the variance stabilizing transformation is unknown.
  - Requires estimation.

Variance stabilization

- Tibshirani (1988) suggests a method to estimate the variance-stabilizing transformation using bootstrap:
  - Transformation is estimated using $\mathbf{B_1}$ replications.
    - Each replication requires $\mathbf{B_2}$ replications to estimate the **standard error**.
  - Bootstrap-t interval is calculated using new $\mathbf{B_3}$ replications.
- Efron and Tibshirani suggest $\mathbf{B_1=100}$**,** $\mathbf{B_2=25}$ and $\mathbf{B_3=1000}$ (total $B_1B_2+B_3=3500$).
- *“bootstrap”* package:
  - boott(…,**VS = TRUE**,…)

Bootstrap-t with variance stabilization

1. Generate $\mathbf{B_1}$ bootstrap samples $\mathbf{x^{*1},\ldots,x^{*B1}}$. For each bootstrap replication $b=1,\ldots,B_1$:
    1. Calculate $\hat{\theta}_b^*$.
    2. Generate $\mathbf{B_2}$ bootstrap samples $\mathbf{x^{**b}}$ to estimate $\widehat{se}_b^*$.
2. Smooth $\widehat{se}_b^*$ as a function of $\hat{\theta}_b^*$.
3. Estimate the variance stabilizing transformation $g(\hat{\theta})$.

4. Generate $\mathbf{B_3}$ bootstrap samples.
    1. Compute a bootstrap-t interval for $\phi = g(\theta)$.
    2. Standard error is (roughly) constant => $Z_b^* = \hat{\phi}_b^* - \hat{\phi}$

5. Perform reverse transformation.

Confidence intervals based on bootstrap percentiles

The percentile interval

- The **bootstrap-t method** estimates the distribution of an approximate pivot Z_b^*:

$$Z_b^* = \frac{\hat{\theta}_b^* - \hat{\theta}}{\widehat{se}_b^*}$$

- The **percentile interval** (Efron 1982) is based on calculating the CDF of the bootstrap replications $\hat{\theta}_b^*$.
  - A 100(1-α)% percentile interval is:

$$\left[\hat{\theta}^*_{\left(\frac{\alpha}{2}\right)}, \hat{\theta}^*_{\left(1-\frac{\alpha}{2}\right)}\right]$$

The percentile interval

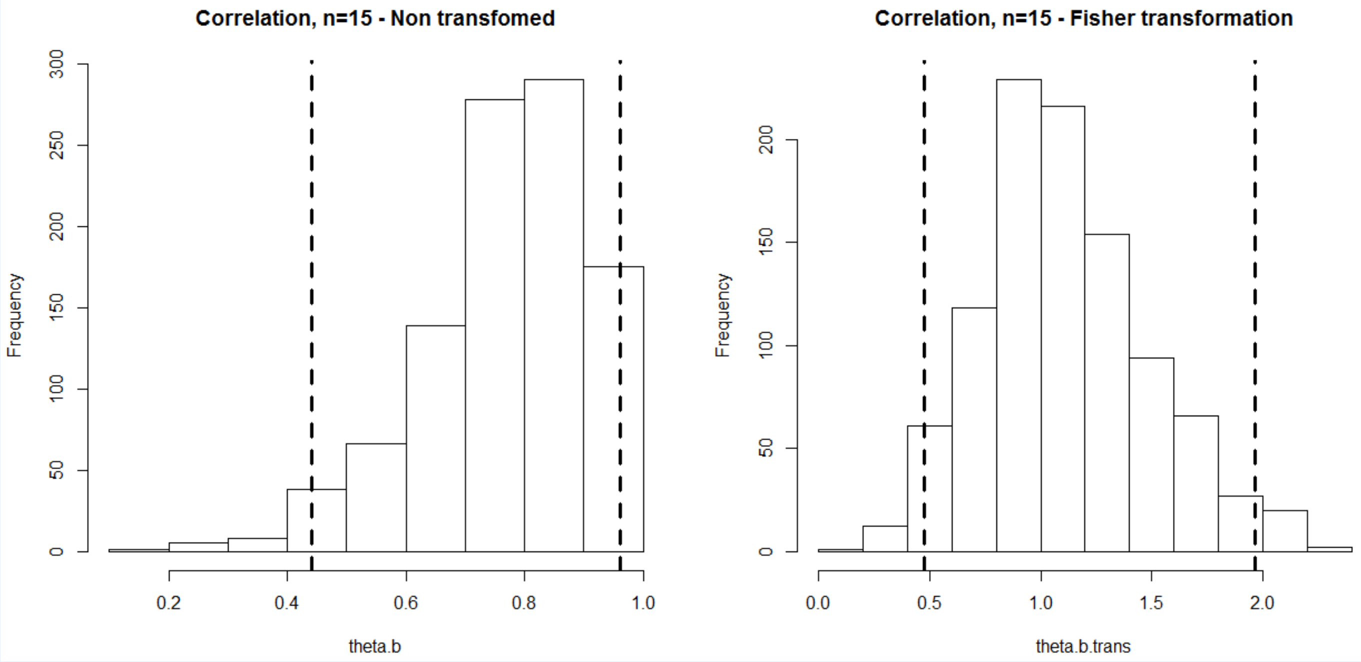
- The percentile interval has 2 major assets:
  - **Invariance to monotone transformation**.
    - For any monotone transformation $\phi = g(\theta)$

$$\left[\hat{\theta}^*_{\left(\frac{\alpha}{2}\right)}, \hat{\theta}^*_{\left(1-\frac{\alpha}{2}\right)}\right] = \left[g^{-1}\left(\hat{\phi}^*_{\left(\frac{\alpha}{2}\right)}\right), g^{-1}\left(\hat{\phi}^*_{\left(1-\frac{\alpha}{2}\right)}\right)\right]$$

    - No knowledge of an appropriate transformation is required.
  - **Range preservation**.
    - $\hat{\theta}$ and $\hat{\theta}^*$ obey the same restrictions on the values of θ.
    - The percentile interval will always fall in the allowable range.

Invariance to transformation

- **Example:** a percentile interval for the ρ=corr(X,Y), using the distribution of $\hat{\theta}^*$ directly (left), and the distribution of $\hat{\phi}^* = g\left(\hat{\theta}^*\right)$ (Fisher transformation, right)



$$\left[\hat{\theta}^*_{\left(\frac{\alpha}{2}\right)}\ ,\hat{\theta}^*_{\left(1-\frac{\alpha}{2}\right)}\right] = [\,0.442{,}0.961] = [g^{-1}\,(0.475)\ ,g^{-1}\,(1.966)\,] = \left[g^{-1}\left(\hat{\phi}^*_{\left(\frac{\alpha}{2}\right)}\right)\ ,g^{-1}\left(\hat{\phi}^*_{\left(1-\frac{\alpha}{2}\right)}\right)\right]$$

Issues with percentile intervals

- Doesn't cope with biased estimators.
- Tendency for under-coverage in small samples.

Table 13.3. *Results of 300 confidence interval realizations for $\theta = \exp(\mu)$ from a standard normal sample of size 10. The table shows the percentage of trials that the indicated interval missed the true value 1.0 on the left or right side. For example, "Miss left" means that the left endpoint was > 1.0. The desired coverage is 95%, so the ideal values of Miss left and Miss right are both 2.5%.*

Method	% Miss left	% Miss right
Standard normal $\hat{\theta} \pm 1.96\widehat{\mathrm{se}}$	1.2	8.8
Percentile (Nonparametric)	4.8	5.2

- Both issues are present in bootstrap-t and normal theory intervals.

Comparison of bootstrap confidence intervals

- Comparison of coverage for the correlation example, with **n=15,100,5000**.

```
> coverage.15
                CI median.width IQR.width   outside  miss.left miss.right miss.total coverage
1         Fisher.z    0.4907985 0.2389673 0.0000000 0.01666667 0.03333333 0.05000000 95.00000
2      Bootstrap-t    0.6962289 0.5743575 0.3233333 0.01000000 0.02333333 0.03333333 96.66667
3 Bootstrap-t_trans    0.5480309 0.4184388 0.0000000 0.03666667 0.06666667 0.10333333 89.66667
4       Percentile    0.4407891 0.2979692 0.0000000 0.01333333 0.07666667 0.09000000 91.00000
5   Bootstrap-t_VS    0.5116025 0.2799836 0.0100000 0.05333333 0.04333333 0.09666667 90.33333
> coverage.100
                CI median.width  IQR.width outside  miss.left miss.right miss.total coverage
1         Fisher.z    0.1726172 0.03065182       0 0.01333333 0.03000000 0.04333333 95.66667
2      Bootstrap-t    0.1898071 0.04457929       0 0.01333333 0.02333333 0.03666667 96.33333
3 Bootstrap-t_trans    0.1856566 0.04238529       0 0.02000000 0.02333333 0.04333333 95.66667
4       Percentile    0.1701845 0.03753713       0 0.02000000 0.04000000 0.06000000 94.00000
5   Bootstrap-t_VS    0.1758487 0.04019904       0 0.03666667 0.04333333 0.08000000 92.00000
> coverage.5000
                CI median.width    IQR.width outside  miss.left miss.right miss.total coverage
1         Fisher.z   0.02419992 0.0006988088       0 0.01000000 0.03000000 0.04000000 96.00000
2      Bootstrap-t   0.02531704 0.0019345037       0 0.02000000 0.02333333 0.04333333 95.66667
3 Bootstrap-t_trans   0.02532924 0.0019570371       0 0.02000000 0.02333333 0.04333333 95.66667
4       Percentile   0.02419462 0.0012691209       0 0.02666667 0.02333333 0.05000000 95.00000
5   Bootstrap-t_VS   0.02433918 0.0012317460       0 0.03666667 0.04333333 0.08000000 92.00000
```

Fin