

1 Item 1

1.1 Algorithm

Step 1: Let B be the number of bootstrap samples taken. With $n = 15$, do SRSWR from schools 1 to 15.

$$\begin{aligned} B_1^* &= \{(X_{11}^*, Y_{11}^*), (X_{21}^*, Y_{21}^*), \dots, (X_{n1}^*, Y_{n1}^*)\} \\ B_2^* &= \{(X_{12}^*, Y_{12}^*), (X_{22}^*, Y_{22}^*), \dots, (X_{n2}^*, Y_{n2}^*)\} \\ &\vdots \\ B_B^* &= \{(X_{1B}^*, Y_{1B}^*), (X_{2B}^*, Y_{2B}^*), \dots, (X_{nB}^*, Y_{nB}^*)\} \end{aligned} \quad (1.1)$$

Step 2: Let $S_{x_b}^*$ and $S_{y_b}^*$ be the standard deviations of the variables, X_b^* and Y_b^* , respectively, where $b = \{1, 2, \dots, B\}$. Calculate the pearson product coefficient of correlation, r_b^*

$$r_b^* = \frac{\frac{1}{n-1} \sum (X_{ib}^* - \bar{X}_b^*) (Y_{ib}^* - \bar{Y}_b^*)}{S_{x_b}^* S_{y_b}^*} \quad (1.2)$$

to yield

$$r = \{r_1^*, r_2^*, \dots, r_B^*\} \quad (1.3)$$

Step 3: Calculate $\widehat{se}(r)$ using

$$\widehat{se}(r) = \sqrt{\frac{\sum_{b=1}^B (r_b^* - \bar{r}^*)^2}{B - 1}} \quad (1.4)$$

1.2 Algorithm implementation

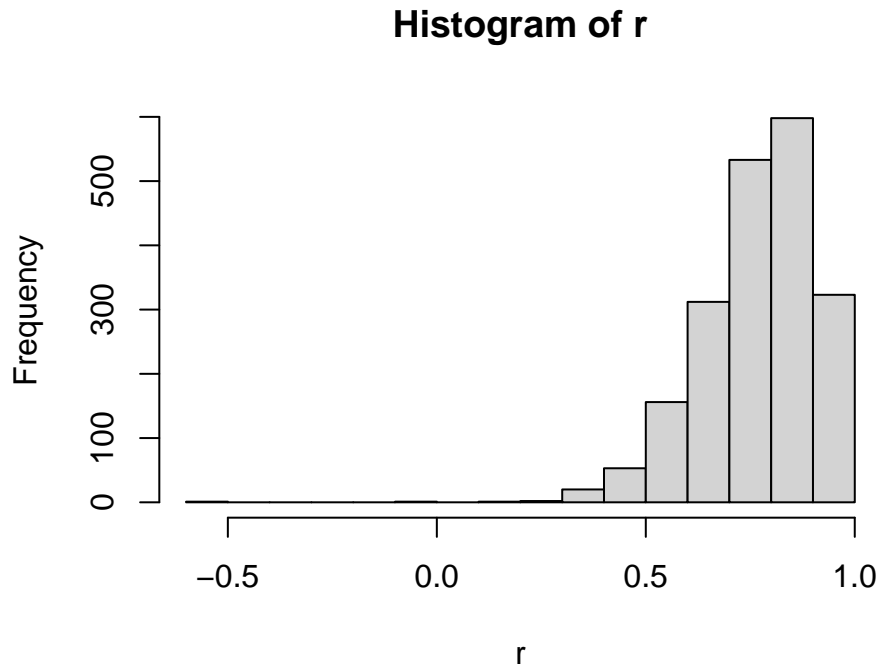
```
# set seed for reproducibility
set.seed(7)

r <-
# Step 2: take the pearson correlation for each bootstrap sample
sapply(
# Step 1: 2000 bootstrap samples (with replacement) of size 15
lapply(
1:2000,
function(x){ law_school_data[sample(law_school_data$School,
15,
replace = TRUE),
c(2,3)] }
),
function(x){cor(x$LSAT,x$GPA, method = "pearson")})

# Step 3: take the bootstrap estimate of the standard error
se_r_boot <- sd(r)

# take the percentile 95% CI for r
ci_r_boot <- c(quantile(r,.025),quantile(r,.975))
```

- i. bootstrap estimate of the standard error of r : 0.1368617
- ii. 95% confidence interval for ρ (the true population correlation): (0.4510449, 0.9625053)
- iii. a histogram showing the bootstrap distribution of the correlation r



2 Item 2

3 Item 3

References

Fox, J.-P., & Marianti, S. (2017). Person-fit statistics for joint models for accuracy and speed. *Journal of Educational Measurement*.