

Randomized Inference

Samantha Tian

2024-08-03

```
heartRates<-read.csv('HeartRatePetsFriends.csv')
HR_pets<-heartRates[heartRates$treatment=="Pets",]
HR_friend<-heartRates[heartRates$treatment=="Friend",]
HR_alone<-heartRates[heartRates$treatment=="Neither",]
```

The data set **heartRates** contains the resting heart rates of subjects in a study under 3 potential treatment conditions - with their pets, with a friend, or with neither. The individual subsets of the data are saved in **HR_pets**, **HR_friend**, and **HR_alone**. (Data from Presence of Human Friends and Pet Dogs as Moderators of Autonomic Responses to Stress in Women)

Question 1 (3 points)

Construct a bootstrap sample 90% confidence interval for the average resting heart rate of individuals in this study, irrespective of treatment. You should repeat the bootstrap process 10,000 times. You should save your individual bootstrap averages in the vector **q1.averages** provided in the code below, and your confidence interval in the variable **q1.CI**. The provided code will plot your responses and interval.

To ensure future questions function correctly, you should avoid modifying the original data in your response.

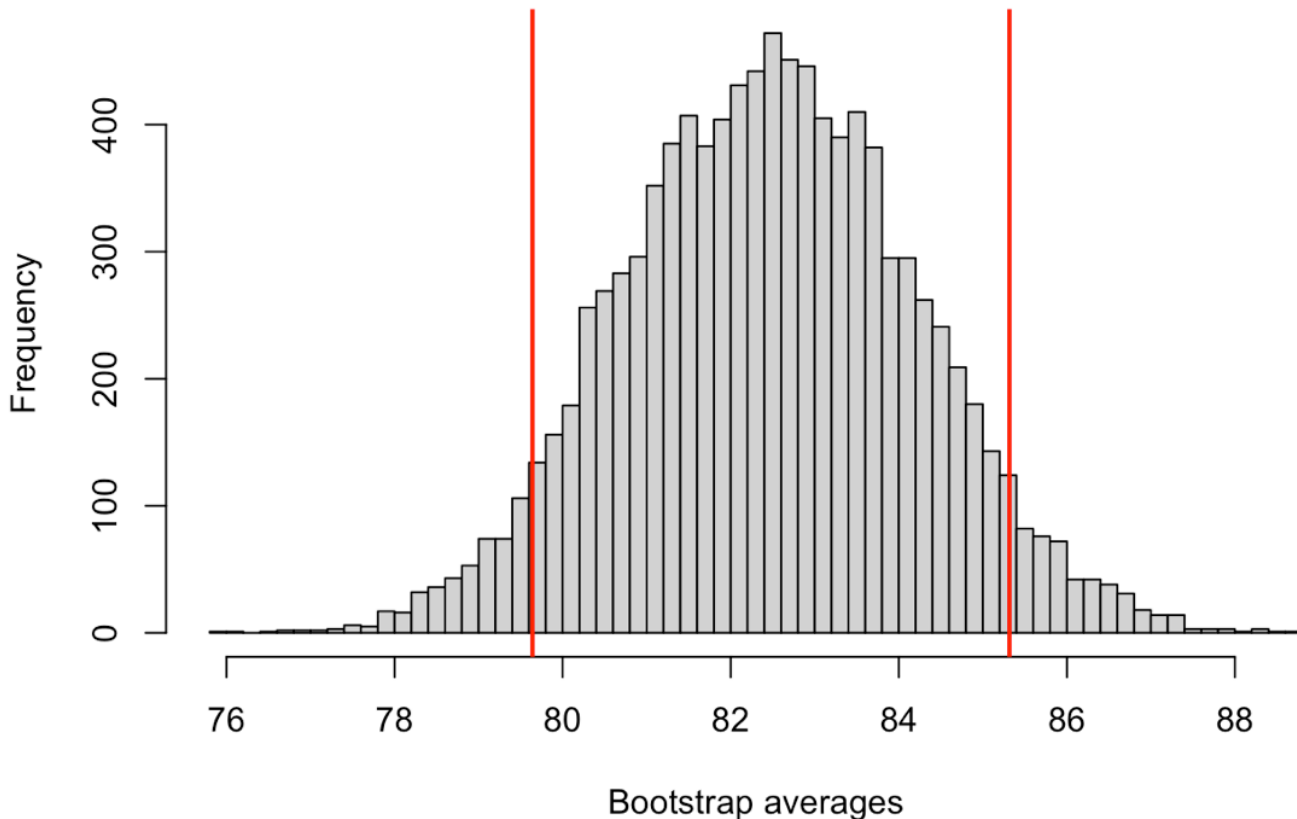
```
N=10000
q1.averages<-rep(0, N)
for(i in 1:N){
  # Calculate an average from a bootstrap sample here
  sample.bootstrap <- sample(heartRates$heartrate, replace = TRUE)
  # Save the average in the variable q1.averages
  q1.averages[i] <- mean(sample.bootstrap)
}

# From your sample, calculate the 90% confidence interval
q1.CI<- quantile(q1.averages, c(0.05, 0.95))
print(q1.CI)
```

```
##           5%           95%
## 79.63991 85.31603
```

```
# The following code will plot your bootstrap simulated means and your confidence interval
hist(q1.averages, breaks=50, xlab='Bootstrap averages', main='90% Bootstrap Confidence Interval')
abline(v=q1.CI[1], lwd=2, col='red')
abline(v=q1.CI[2], lwd=2, col='red')
```

90% Bootstrap Confidence Interval



Question 2 (3 points)

Construct a bootstrap sample 95% confidence interval for the difference in average resting heart rate of individuals with pets compared to individuals with friends. You should repeat the bootstrap process 10,000 times. You should save your individual bootstrap average differences in the vector **q2.avgdiffs** provided in the code below, and your confidence interval in the variable **q2.CI**. The provided code will plot your responses and interval.

To ensure future questions function correctly, you should avoid modifying the original data in your response.

```
N=10000
q2.avgdiffs<-rep(0, N)
for(i in 1:N){
  # Calculate a difference of averages from two bootstrap sample here
  pet.heartrates <- sample(heartRates$heartrate[heartRates$treatment == "Pets"], size
= sum(heartRates$treatment == "Pets"), replace = TRUE)

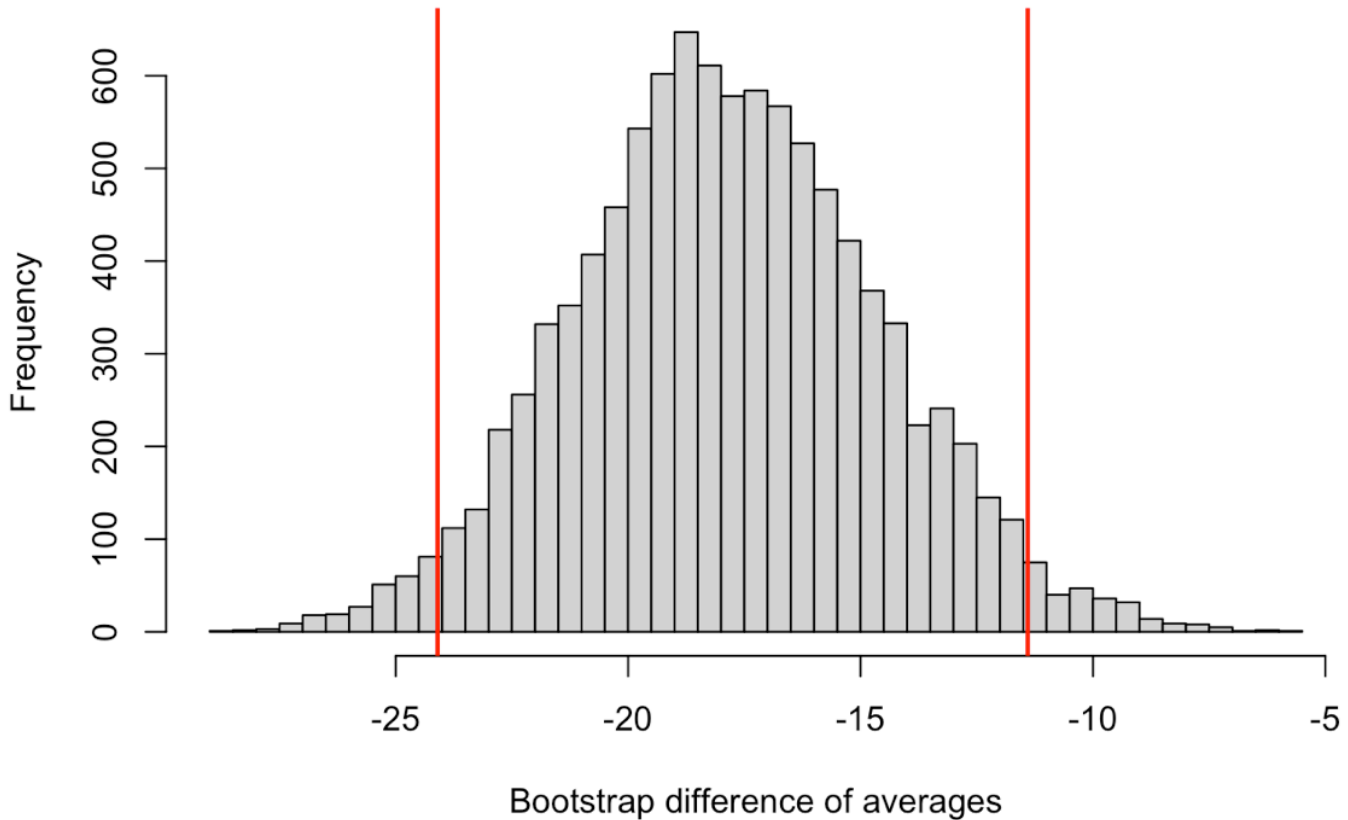
  friends.heartrates <- sample(heartRates$heartrate[heartRates$treatment == "Friend"], size = sum(heartRates$treatment == "Friend"), replace = TRUE)
  # Save the average difference in the variable q2.avgdiffs
  q2.avgdiffs[i]<-mean(pet.heartrates) - mean(friends.heartrates)
}

# From your sample, calculate the 95% confidence interval
q2.CI<- quantile(q2.avgdiffs, c(0.025, 0.975), na.rm = TRUE)
print(q2.CI)
```

```
##      2.5%      97.5%
## -24.09754 -11.40344
```

```
# The following code will plot your bootstrap simulated means and your confidence interval
hist(q2.avgdiffs, breaks=50, xlab='Bootstrap difference of averages', main='95% Bootstrap Confidence Interval')
abline(v=q2.CI[1], lwd=2, col='red')
abline(v=q2.CI[2], lwd=2, col='red')
```

95% Bootstrap Confidence Interval



Question 3 (4 points)

We wish to test the hypothesis that, on average, having pets around reduces average resting heart rate compared to having no pets or friends around. Specifically:

H_0 : Average resting heart does not differ when having pets around or not having pets or friends around

$$\mu_{\text{pets}} = \mu_{\text{neither}}$$

H_a : Average resting heart is lower when having pets around compared to not having pets or friends around

$$\mu_{\text{pets}} < \mu_{\text{neither}}$$

You should simulate a total of 10,000 permutations of treatment, and save each difference of averages in the variable **q3.avgdiffs** provided in the code below. The data frame **HR_petsAndAlone** has been created and can be freely modified during your permutation tests. Plots will assume that you look at the difference between the average heartrate with pets ('Pets') minus the average heart rate alone ('Neither').

Calculate the simulated p-value and save it in the variable **q3.pvalue**

```
# combine data frames
HR_petsAndAlone=rbind(HR_pets, HR_alone)

observed_diff<-mean(HR_pets$heartrate)-mean(HR_alone$heartrate)

N=10000
q3.avgdiffs<-rep(0, N)
for(i in 1:N){
  # Permute the treatments and calculate the averages for each treatment type
  perm.treatment <- sample(HR_petsAndAlone$treatment)

  pets.avg <- mean(HR_petsAndAlone$heartrate[perm.treatment == "Pets"])
  neither.avg <- mean(HR_petsAndAlone$heartrate[perm.treatment == "Neither"])

  # Save the average difference in the variable q3.avgdiffs
  q3.avgdiffs[i]<- pets.avg - neither.avg
}

# From your permutation samples, calculate the estimated p-value for the hypotheses of
# interest
q3.pvalue<- sum(q3.avgdiffs <= observed_diff) / N
print(q3.pvalue)
```

```
## [1] 0.0086
```

```
# The following code will plot your permutation simulated mean differences and your observed
# difference
hist(q3.avgdiffs, breaks=50, xlab='Permutation test difference of averages', main='Permutation
# tests of heartrate with pets vs alone')
abline(v=observed_diff, lwd=2, col='red')
```

Permutation tests of heartrate with pets vs alone

