Hypothesis Test with the bootstrap

Xinyi Lu

Data set from surveys of customers of 168 Italian restaurants in New York City.[1]

The variables are:

Price = the price (in $ US) of dinner (including 1 drink and tip)

Food = customer rating of the food (out of 30)

Decor = customer rating of the decor (out of 30)

Service = customer rating of the service (out of 30)

East = dummy variable, 1 (0) if the restaurant is east (west) of Fifth Avenue

Data set load from Prof.Fox's github data

[1]Zagat Survey 2001: New York City Restaurants

H_0 and H_1

Claim: Restaurants on East Fifth Avenue have higher food ratings than those on the west side.

H_0: $\mu_{East} = \mu_{West}$

$H_{1:}$ $\mu_{East} > \mu_{West}$

Achieved significance level (ASL): p-value

ASL = $Prob_{H0}\{ \mu_{boot} \geq \mu_{obv} \}$

ASL < .10 borderline evidence against H0.

ASL < .05 reasonably strong evidence against H0.

ASL < .025 strong evidence against H0.

ASL < .01 very strong evidence against H0.

1. Loading packages and read data to RStudio

```
```{r}
# loading package
library(tidyverse)
library(dplyr)
```

```{r}
# read cvs from Prof.Fox's github data set
nyc <- read.csv("https://ericwfox.github.io/data/nyc.csv")
str(nyc)
```
```

```
'data.frame':	168 obs. of  6 variables:
 $ Restaurant: chr  "Daniella Ristorante" "Tello's Ristorante" "Biricchino" "Bottino" ...
 $ Price     : int  43 32 34 41 54 52 34 34 39 44 ...
 $ Food      : int  22 20 21 20 24 22 22 20 22 21 ...
 $ Decor     : int  18 19 13 20 19 22 16 18 19 17 ...
 $ Service   : int  20 19 18 17 21 21 21 21 22 19 ...
 $ East      : int  0 0 0 0 0 0 0 1 1 1 ...
```

2-1. Data wrangling

````
```{r}
#data wrangling
East <- nyc %>% filter(East == 1) %>% dplyr::select(Food)
West <- nyc %>% filter(East == 0) %>% dplyr::select(Food)
```

```{r}
str(East)
```
````

```
'data.frame':	106 obs. of  1 variable:
 $ Food: int  20 22 21 19 21 21 19 20 21 22 ...
```

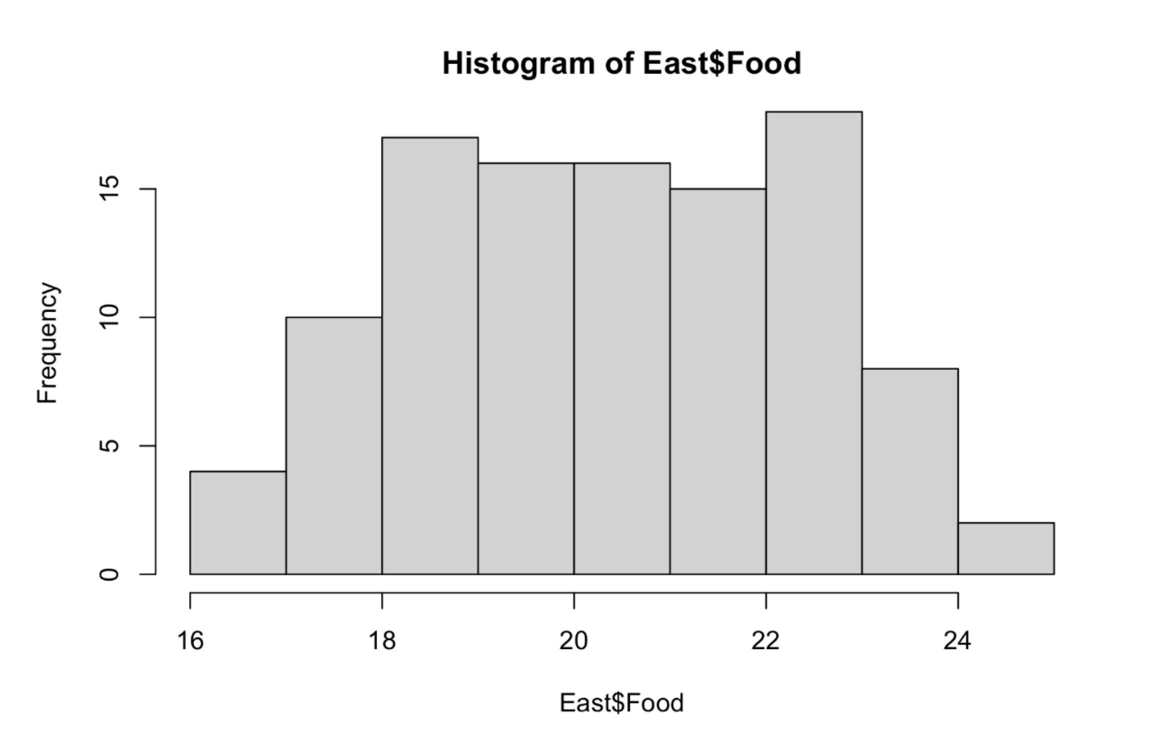
````
## Note: East and West are data frame

```{r}
length(East$Food)
length(West$Food)
```
````

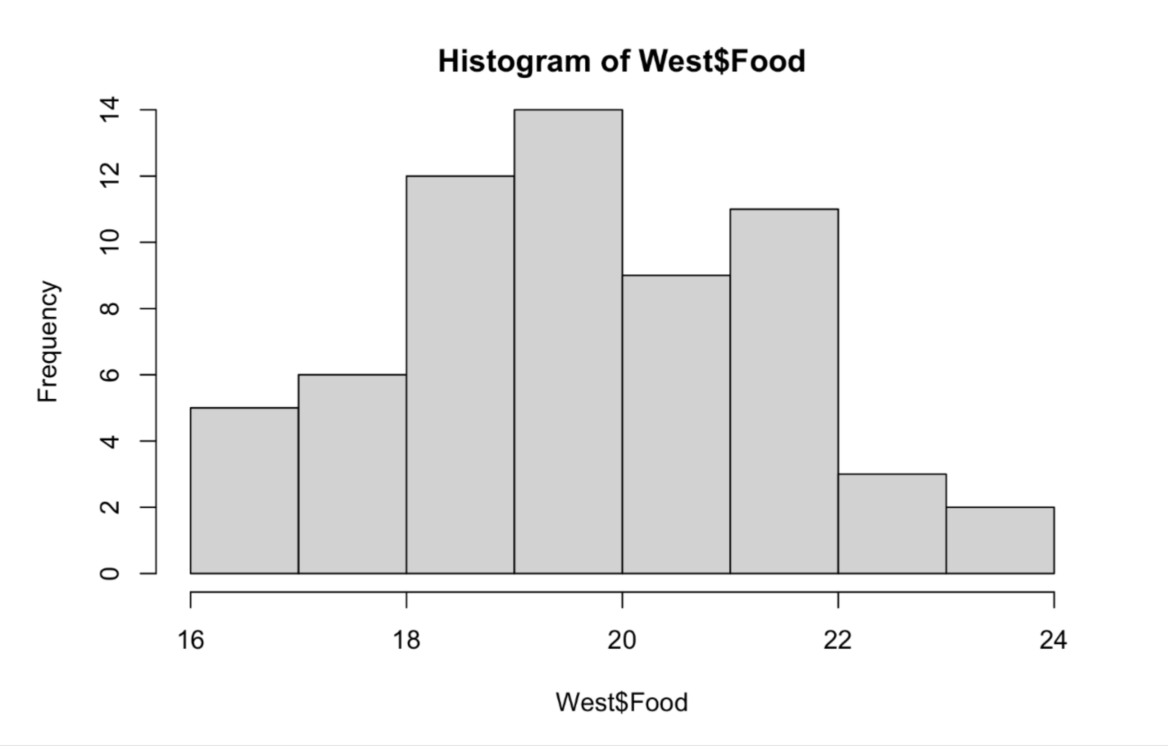
```
[1] 106
[1] 62
```

2-2 Data Visualization

```{r}
hist(East$Food)
```




```{r}
hist(West$Food)
```



3. Regular T Test for Mean difference

```{r}
t.test(East$Food, West$Food, alternative = "greater", var.equal = TRUE)
```

```
	Two Sample t-test

data:  East$Food and West$Food
t = 2.3627, df = 166, p-value = 0.009651
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.2215988       Inf
sample estimates:
mean of x mean of y 
 20.86792  20.12903 
```

Achieved significance level (ASL): p-value = 0.009651

P-value < α = 0.05 and 0.1, reject H_0.

4. Permutation Test for Mean difference

(without replacement)

````
```{r}
# Permutation Test
obv.mean <- mean(East$Food) - mean(West$Food)
obv.mean
```
````

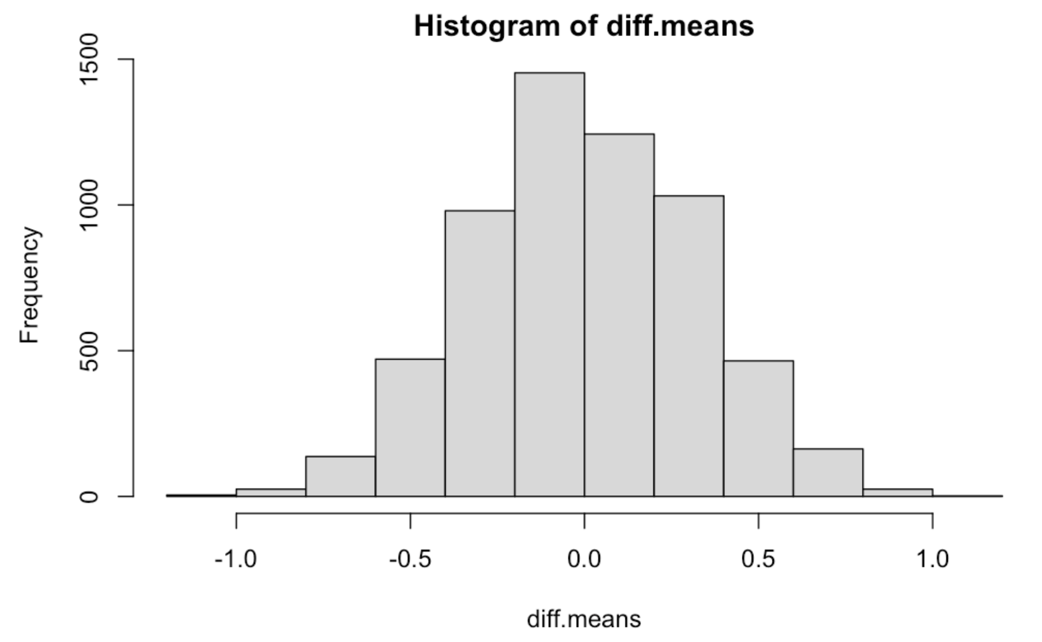
```
[1] 0.7388923
```

````

```{r}
comb <- c(East$Food,West$Food)
diff.means <- numeric()

set.seed(234)
for (i in 1:5000) {
  means <- sample(comb, 168, replace = FALSE)
  diff.means[i] <- mean(means[1:106]) - mean(means[107:168])
}
```
````

````
```{r}
hist(diff.means)
```
````




````

```{r}
length(diff.means[ diff.means >= obv.mean]) / 5000
```
````

```
[1] 0.0076
```

p-value = 0.0076, which is close to the t-test p-value: 0.0096, not under the assumptions of normality

P-value < α = 0.05 and 0.1, reject H_0

4-1: 1000 permutations vs 5000 permutations

```r
```{r}
comb <- c(East$Food,West$Food)
diff.means <- numeric()

set.seed(234)
for (i in 1:1000) {
  means <- sample(comb, 168, replace = FALSE)
  diff.means[i] <- mean(means[1:106]) - mean(means[107:168])
}
```

```{r}
hist(diff.means)
```

```{r}
length(diff.means[ diff.means >= obv.mean]) / 1000
```
```

```
[1] 0.013
```

When I used 1000 permutations, the p-value of test is 0.013. This p-value is small enough.

5. Hypothesis Test with The Bootstrap (with replacement)

The bootstrap test statistic for $H_0 : F = G$

```
```{r}
# HT
# comb <- c(East$Food,West$Food)
# obv.mean <- mean(East$Food) - mean(West$Food) = 0.7388923
set.seed(345)
boot_sample <- matrix(sample(comb, 168 *10000, replace = TRUE), nrow = 10000)
diff_means <- apply(boot_sample, 1,
                    function(x){mean(x[1:106]) - mean(x[107:168])})

length(diff_means[diff_means>= obv.mean])/10000
```
```

```
[1] 0.0096
```

p-value = 0.0096 , which is same as the t-test p-value: 0.0096
P-value < α = 0.05, reject H_0.

5-1. Studentized Statistics

$$t(x) = \frac{\bar{z} - \bar{y}}{\sqrt{\sigma^2 (1/n + 1/m)}}$$

but σ^2 is unknown.

we use $\bar{\sigma}^2$ (pool variance) to replace σ^2.

$$\bar{\sigma}^2 = \frac{\sum_{i=1}^{n} (z_i - \bar{z})^2 + \sum_{i=1}^{m} (y_i - \bar{y})^2}{n+m-2}.$$

```
```{r}
# HT with studentized statistics
stud.t <- function(x){
  num <- mean(x[1:106]) - mean(x[107:168])
  pool.var <- (var(x[1:106])*105 + var(x[107:168])*61)/166
  denominator <- sqrt(pool.var*(1/106 + 1/62))
  return(num/denominator)
}

obv.stud.t <- stud.t(comb)

new.t <- apply(boot_sample, 1, stud.t)

length(new.t[new.t>= obv.stud.t])/10000
```
```

```
[1] 0.0094
```

I used the same set of bootstrap samples. Studendization gives a different p-value 0.094.
P-value < α = 0.05 and 0.1, reject H_0.

6. Summary

T-test, Permutation test and hypothesis test are all have similar p-value, and they are all rejecting H0, when < α = 0.05 and 0.1. Therefore, restaurants on East Fifth Avenue have higher food ratings than those on the west side.

	T-test	Permutation test	Hypothesis Test	Hypothesis Test with studendization
ASL p-value	0.009651	0.0076	0.0096	0.0094

Thank you!
Question?

Xinyi Lu