

# HW 02

Due Thursday, October 15, 11:59 PM

Matthew Xiao

```
library(tidyverse)
licorice <- read_csv("data/licorice.csv")
```

## Exercise 1

```
c_licorice <- licorice %>% filter(!is.na(pacu30min_throatPain))

n_sims <- 1000
set.seed(1)
boot_dist = numeric(n_sims)

for (i in 1:n_sims){
  set.seed(i)
  indices <- sample(1:nrow(c_licorice), replace = TRUE)
  boot_mean <- c_licorice %>%
    slice(indices) %>%
    summarize(boot_meean = mean(pacu30min_throatPain)) %>% pull()
  boot_dist[i] = boot_mean
}

boot_means <- tibble(boot_dist)
boot_means %>% summarize(Lower = quantile(boot_dist, 0.025),
                        Upper = quantile(boot_dist, 0.975)) %>%
  knitr::kable()
```

Lower	Upper
0.4891631	0.8112661

```
t.test(c_licorice$pacu30min_throatPain, conf.level = .95)
```

```
##
## One Sample t-test
##
## data: c_licorice$pacu30min_throatPain
## t = 7.9287, df = 232, p-value = 9.291e-14
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
## 0.4870277 0.8091096
## sample estimates:
```

```
## mean of x
## 0.6480687
```

## Exercise 2

```
b_licorice <- licorice %>% summarize(mean_b = mean(preOp_calcBMI))

set.seed(2)
n_sims <- 2000
boot_dist = numeric(n_sims)
for(i in 1:n_sims){
  set.seed(i)
  indices <- sample(1:nrow(licorice), replace = T)
  boot_mean <- licorice %>%
    slice(indices) %>%
    summarize(boot_mean = mean(preOp_calcBMI)) %>%
    pull()

  boot_dist[i] <- boot_mean
}
mu_0 = 26
boot_means <- tibble(boot_dist)

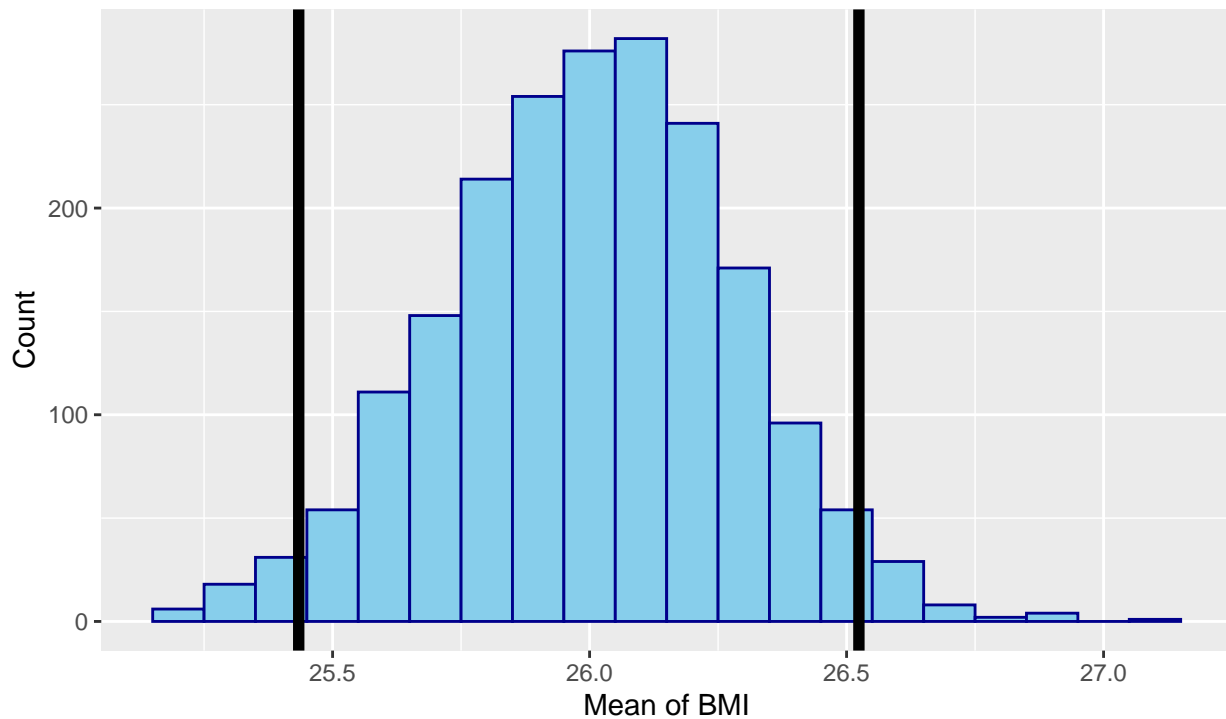
offset <- boot_means %>%
  summarize(mu_0 - mean(boot_dist) )%>% pull()
boot_means <- tibble(boot_dist)
boot_means %>% summarize(Lower = quantile(boot_dist, 0.025),
                        Upper = quantile(boot_dist, 0.975)) %>%

knitr::kable()
```

Lower	Upper
25.01972	26.10958

```
boot_means <- boot_means %>%
  mutate(shifted_means = boot_dist+offset)
ggplot(data = boot_means, aes(x = shifted_means)) +
  geom_histogram(binwidth = 0.1, color = "darkblue", fill = "skyblue") +
  labs(x = "Mean of BMI",
       y = "Count",
       title = "The graph displays the null distribution for
the mean BMI of the Preoperation Patients with the assumption that the
mean BMI is 26.") +
  geom_vline(xintercept = quantile(boot_means$shifted_means, 0.025),
            lwd = 2,
            color = "black") +
  geom_vline(xintercept = quantile(boot_means$shifted_means, 0.975),
            lwd = 2,
            color = "black")
```

The graph displays the null distribution for the mean BMI of the Preoperation Patients with the assumption that the mean BMI is 26.



```
obs <- licorice %>%
  summarize(mean(preOp_calcBMI)) %>%
  pull()
dif <- mu_0 - obs
boot_means <- boot_means %>%
  mutate(shifted_means = boot_dist + offset)

boot_means1 <- boot_means %>%
  mutate(extreme = ifelse(shifted_means <= obs |
                           shifted_means >= mu_0 + dif, 1, 0)) %>%
  count(extreme) %>%
  mutate(prob = n / sum(n)) %>%
  select(prob) %>%
  slice(2)
boot_means1 %>% knitr::kable()
```

prob
0.144

### Exercise 3

```
B_pain <- c_licorice %>%
  mutate(asapain = ifelse(pacu30min_throatPain > 0, 1, 0))
```

```
t.test(B_pain$preOp_asa ~ asapain,
      data = B_pain,
      mu = 0,
      var.equal = FALSE,
      alternative = "two.sided",
      conf.level = 0.95)

##
## Welch Two Sample t-test
##
## data: B_pain$preOp_asa by asapain
## t = -1.2976, df = 124.15, p-value = 0.1968
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.29954887 0.06230774
## sample estimates:
## mean in group 0 mean in group 1
## 2.053254 2.171875
```

#### Exercise 4

```
t.test(pacu30min_throatPain ~ treat,
      data = c_licorice,
      mu = 0,
      var.equal = FALSE,
      alternative = "greater",
      conf.level = 0.95)

##
## Welch Two Sample t-test
##
## data: pacu30min_throatPain by treat
## t = 4.8035, df = 157.3, p-value = 1.804e-06
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
## 0.4932019 Inf
## sample estimates:
## mean in group 0 mean in group 1
## 1.0258621 0.2735043
```

#### Exercise 5

```
set.seed(5)
cc_licorice <- c_licorice %>%
  mutate(pain = ifelse(pacu30min_throatPain > 0 | pacu30min_swallowPain > 0, 1, 0))
licorice_l <- cc_licorice %>%
  filter(treat == 0)
sugar <- cc_licorice %>%
  filter(treat == 1)
n_sims <- 2000
boot_diffs <- numeric(n_sims)

for(i in 1:n_sims){
```

```

# create indices
indices_h <- sample(1:nrow(licorice_1), replace = T)
indices_p <- sample(1:nrow(sugar), replace = T)
# bootstrap est. group means
temp_h <- licorice_1 %>%
  slice(indices_h) %>%
  summarize(mean1 = mean(pain)) %>%
  select(mean1) %>%
  pull()
temp_p <- sugar %>%
  slice(indices_p) %>%
  summarize(mean_jitter = mean(pain)) %>%
  select(mean_jitter) %>%
  pull()

boot_diffs[i] <- temp_h - temp_p
}
boot_diffs <- tibble(diffs = boot_diffs)
boot_diffs %>%
  summarize(lower = quantile(diffs, 0.025),
            upper = quantile(diffs, 0.975))

## # A tibble: 1 x 2
##   lower upper
##   <dbl> <dbl>
## 1 0.0446 0.277

obs_diff <- boot_diffs %>%
  summarize(obs_diff = mean(diffs)) %>%
  pull()
offset <- boot_diffs %>%
  summarize(offset = 0 - mean(diffs)) %>%
  pull()
null_dist <- boot_diffs %>%
  mutate(centered_diffs = diffs + offset) %>%
  select(centered_diffs)
null_dist %>%
  mutate(extreme = ifelse(centered_diffs > abs(obs_diff), 1, 0)) %>%
  summarize(p_val = mean(extreme))

## # A tibble: 1 x 1
##   p_val
##   <dbl>
## 1 0.0055

```

## Exercise 6

## Exercise 7

TRUE

## Exercise 8

FALSE. Our p-value can not be evidence for, but rather it would only be evidence that does not reject it can not support.

**Exercise 9**

True

**Exercise 10**

TRUE

**Exercise 11**

FALSE we can not know for sure

**Exercise 12**

TRUE