

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))
```

```
## # A tibble: 1 x 2
##   lower upper
##   <dbl> <dbl>
## 1 0.489 0.811
```

Exercise 2

```
licorice

## # A tibble: 235 x 19
##   preOp_gender preOp_asa preOp_calcBMI preOp_age preOp_mallampati preOp_smoking
##   <dbl>      <dbl>      <dbl>      <dbl>      <dbl>      <dbl>
## 1         0         3        33.0        67         2         1
## 2         0         2        23.7        76         2         2
## 3         0         2        26.8        58         2         1
## 4         0         2        28.4        59         2         1
```

```
## 5          0          1          30.4          73          1          2
## 6          0          2          35.5          61          3          1
## 7          0          3          25.5          66          1          1
## 8          0          2          31.1          61          2          1
## 9          0          3          21.2          83          1          1
## 10         0          3          27.2          69          2          3
## # ... with 225 more rows, and 13 more variables: preOp_pain <dbl>, treat <dbl>,
## #   intraOp_surgerySize <dbl>, extubation_cough <dbl>, pacu30min_cough <dbl>,
## #   pacu30min_throatPain <dbl>, pacu30min_swallowPain <dbl>,
## #   pacu90min_cough <dbl>, pacu90min_throatPain <dbl>, postOp4hour_cough <dbl>,
## #   postOp4hour_throatPain <dbl>, pod1am_cough <dbl>, pod1am_throatPain <dbl>
```

```
set.seed(2)
n_sims <- 900

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)

boot_means %>%
  summarize(unshifted_mean_HNR = mean(boot_dist))
```

```
## # A tibble: 1 x 1
##   unshifted_mean_HNR
##             <dbl>
## 1             25.6
```

```
offset <- boot_means %>%
  summarize(mu_0 - mean(boot_dist)) %>%
  pull()
```

```
boot_means <- boot_means %>%
  mutate(shifted_means = boot_dist + offset)
```

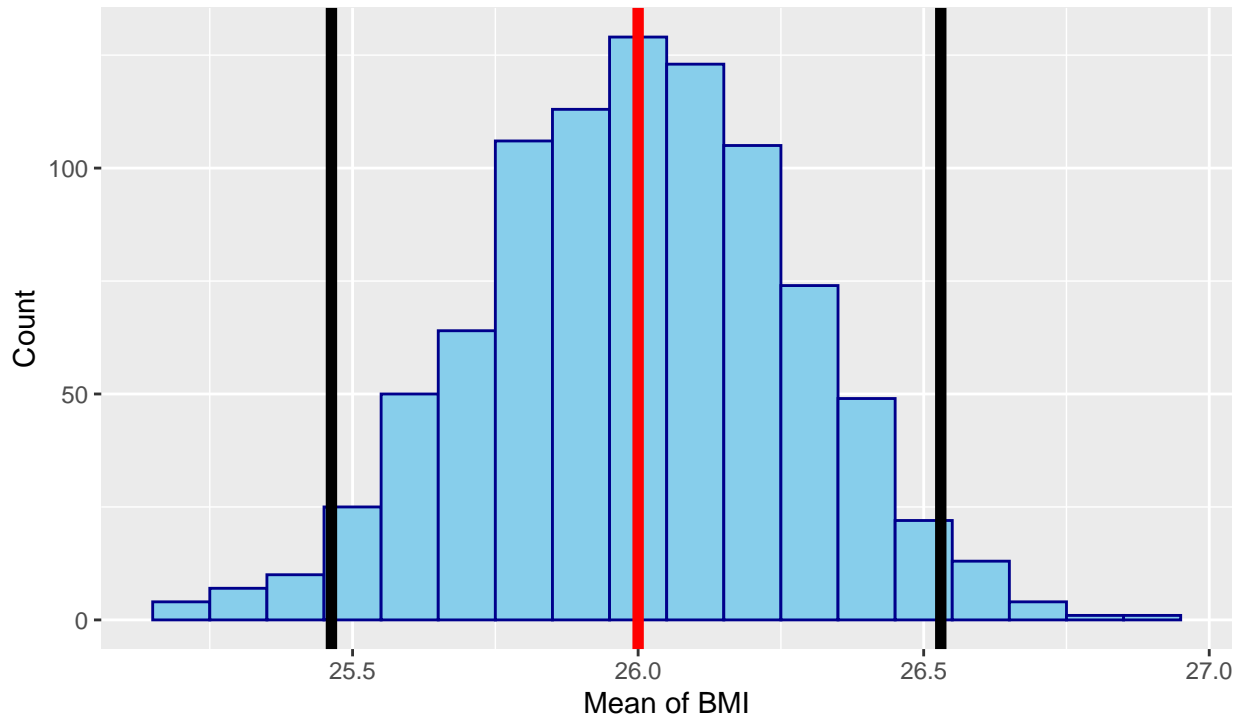
```
ggplot(data = boot_means, aes(x = shifted_means)) +
  geom_histogram(binwidth = 0.1, color = "darkblue", fill = "skyblue") +
  labs(x = "Mean of BMI",
       y = "Count",
       title = "The graph displays the null distribution for
the mean BMI of the Preoperation Patients with the assumption that the
mean BMI is 26.") +
  geom_vline(xintercept = 26, lwd = 2, color = "red") +
  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.



```

boot_means %>%
  summarize(lower = quantile(shifted_means, 0.025),
    upper = quantile(shifted_means, 0.975))

```

```

## # A tibble: 1 x 2
##   lower upper
##   <dbl> <dbl>
## 1  25.5  26.5

```

Exercise 3

```

B_pain <- 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
```

and so on.