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
```

Both the clt- based apprach and the simulation approach created similar bounds of around .49 for the lower bound and .81 for the upper bounds. The bounds for the simulation approach are slightly smaller as the lower bound is slightly higher as well as the upper bound is as well. In both cases the 95% constructed region creates bounds of the pain level of .487 to .809. If we would randomly sample the patient data for throat pain, following the same approach we would be 95% confident that 95% of the new constructed confidence intervals to encompass the true mean.

Exercise 2

significance level: $\alpha = 0.05$ $H_0 : \mu = 26$ $H_A : \mu \neq 26$

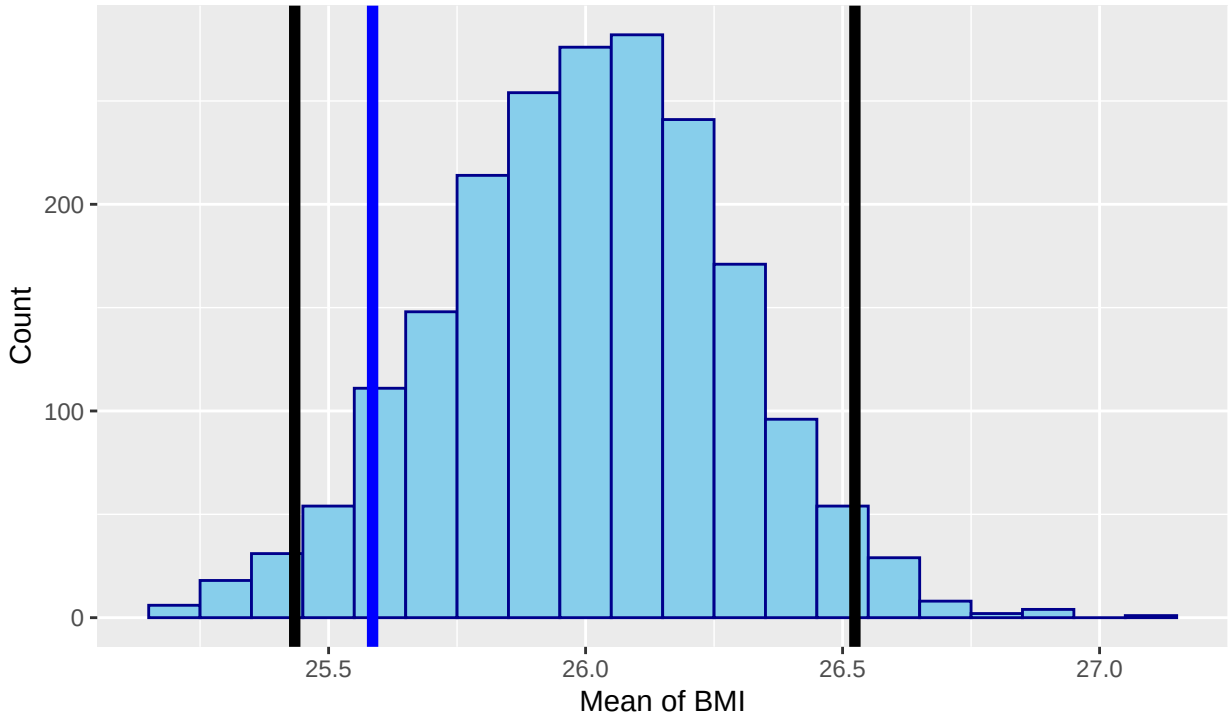
```
b_licorice <- licorice %>% summarize(mean_b = mean(preOp_calcBMI))

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 asumption 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") +
  geom_vline(xintercept = quantile(mean(boot_dist), label = "Sample BMI", 0.975),
             lwd= 2,
             color = "blue")
```

The graph displays the null distribution for
the mean BMI of the Preoperation Patients with the asumption 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

The blue line is the sample BMI of the patient's. Based on the simulated data. It seems as if that the probability that the mean of sampled patients' BMI have a 25.59 or more extreme BMI is 14.4%. This is above the general alpha level of .05 which means that we have insufficient evidence to reject the null hypothesis (that the BMI among patients is the same as the rest of the German adult population.)

We find that the p value of our test is 0.144. Since our p-value is greater than the alpha level of 0.05, we decide to fail to reject the null hypothesis that the mean patient BMI is different to the mean of the general German population.

In the visualization of the graph we can see that the blue line is within the two black lines which are the boundaries of a 95% significance level. Thus our data is not extreme enough to suggest that the difference between the German BMI of the entire population of 26 and the sample's mean is significant enough to reject the null hypothesis.

We ultimately conclude that there is not enough evidence to suggest the mean patients have a different than the German population.

Exercise 3

significance level:$\alpha = 0.05$ H_0: Throat pain 30 minutes after and ASA classification are independent; there is no association between them H_1: Throat pain 30 minutes after and ASA classification are not independent; there is association between them

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


chisq.test(table(B_pain$asapain, B_pain$preOp_asa))
```

```
##
##  Pearson's Chi-squared test
##
## data:  table(B_pain$asapain, B_pain$preOp_asa)
## X-squared = 2.4178, df = 2, p-value = 0.2985
```

The distribution of the test statistic under the null distribution is a chi-square distribution that has a 2 degree of freedoms.

The X^2 value of .2985.

We find that the p value of our test is .2985 Since our p-value is greater than the alpha level of 0.05, we decide to fail to reject the null hypothesis that Throat pain 30 minutes after and ASA classification are independent.

We ultimately conclude that there is not enough evidence to suggest that there is a correlation between between throat pain 30 minutes after and the ASA physical classification

Exercise 4

$\alpha = 0.05$ μ_s : mean throat pain 30 minutes level after the operation for those with the sugar placebo μ_l : mean throat pain level 30 minutes after the operation for those with the licorice H_0: Gargling licorice does not affect the mean throat pain of patients $H_0 : \mu_s = \mu_l$

H_A: Gargling licorice decreases the the mean throat pain of patients $H_A : \mu_s > \mu_l$

```
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
```

Those that were treated with sugar (treat 0) had a higher incidence of having throat pain 30 after. We received a pvalue of 1.8e-06 which is much smaller than 0.05 thus we can reject the null hypothesis and state that we have evidence that licorice decreases the incidence of having throat pain 30 min after the operation.

The distribution of the test statistic under the null distribution is a t distribution that has a 157.3 degrees of freedoms.

The test statistic is 4.805

We find that the p value of our test is 1.804e-06. Since our p-value is less than the alpha level of 0.05, we decide to reject the null hypothesis that the mean throat pain after 30 after a licorice gargle is greater than or equal to the
mean throat pain after 30 with the sugar gargle, in favor of the alternative hypothesis that the mean throat pain after 30 after a licorice gargle is less than to the mean throat pain after 30 with the sugar gargle.

We ultimately conclude that there is enough evidence to suggest a licorice gargle decreases the mean throat pain after 30 after a licorice gargle.

Exercise 5

significance level:$\alpha = 0.05$ μ_s : mean pain 30 minutes probability after for those with the sugar placebo μ_l : mean pain 30 minutes probability after for those with the licorice H_0: Gargling licorice does not affect the mean throat pain of patients $H_0 : \mu_s = \mu_l$

H_A: Gargling licorice decreases the the mean throat pain of patients $H_A : \mu_s > \mu_l$

```
set.seed(5)
cc_licorice <- c_licorice %>%
  mutate(pain = ifelse(pacu30min_throatPain >0 | pacu30min_swallowPain >0, 1, 0))
sugar <- cc_licorice %>%
  filter(treat == 0 )
licorice_l <- 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(sugar), replace = T)
  indices_p <- sample(1:nrow(licorice_l), replace = T)
  # bootstrap est. group means
  temp_h <- sugar %>%
    slice(indices_h) %>%
    summarize(mean1 = mean(pain)) %>%
    select(mean1) %>%
```

```
    pull()
  temp_p <- licorice_l %>%
    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
```

Since our p-value of (0.0055) is less than the alpha level of 0.05, we decide to reject the null hypothesis that the mean throat pain after 30 after a licorice solution is greater than or equal to the mean pain after 30 with the sugar placebo solution, in favor of the alternative hypothesis that the mean throat pain after 30 after a licorice gargle is less than to the mean pain after 30 with the sugar gargle.

We ultimately conclude that there is enough evidence to suggest a licorice gargle decreases the mean throat pain after 30 after a licorice gargle

Exercise 6

Since this is an experimental study if we assume that those who have a sore throat will have sore throat pain, we can conclude that a licorice solution will decrease sore throat pain 30 minutes after. This is because in exercise 4 we tested and found a p value of 1.804e-06 which suggests that there is significance to indicate that a licorice solution will decrease the sore throat pain 30 minutes after. Thus we can say the licorice solution prior to surgery is effective in r educing post-intubation sore throat.

However if we say that this conclusion holds true for longer periods of time (90 minutes or 4 hours or at 1 am), we can not say that is it effective. That is because we did not test for that duration of time and thus using the 30 minute data would be an extrapolation.

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. In addition we can not determine the strength our data has in supporting the null hypothesis.

Exercise 9

False - we do not know the power of the test thus we can not find the (1 - power) in order to determine the probability of having a type two error.

Exercise 10

FALSE- we are only 95% confident in that if we were to repeatedly construct 95% confidence intervals for the difference in mean BMI in the same way from the original population, then we know that 95% of those intervals would truly contain the true population difference in means.

Exercise 11

True - the variance is already in the confidence interval

Exercise 12

TRUE