BACS HW5

109062710

April 9, 2021

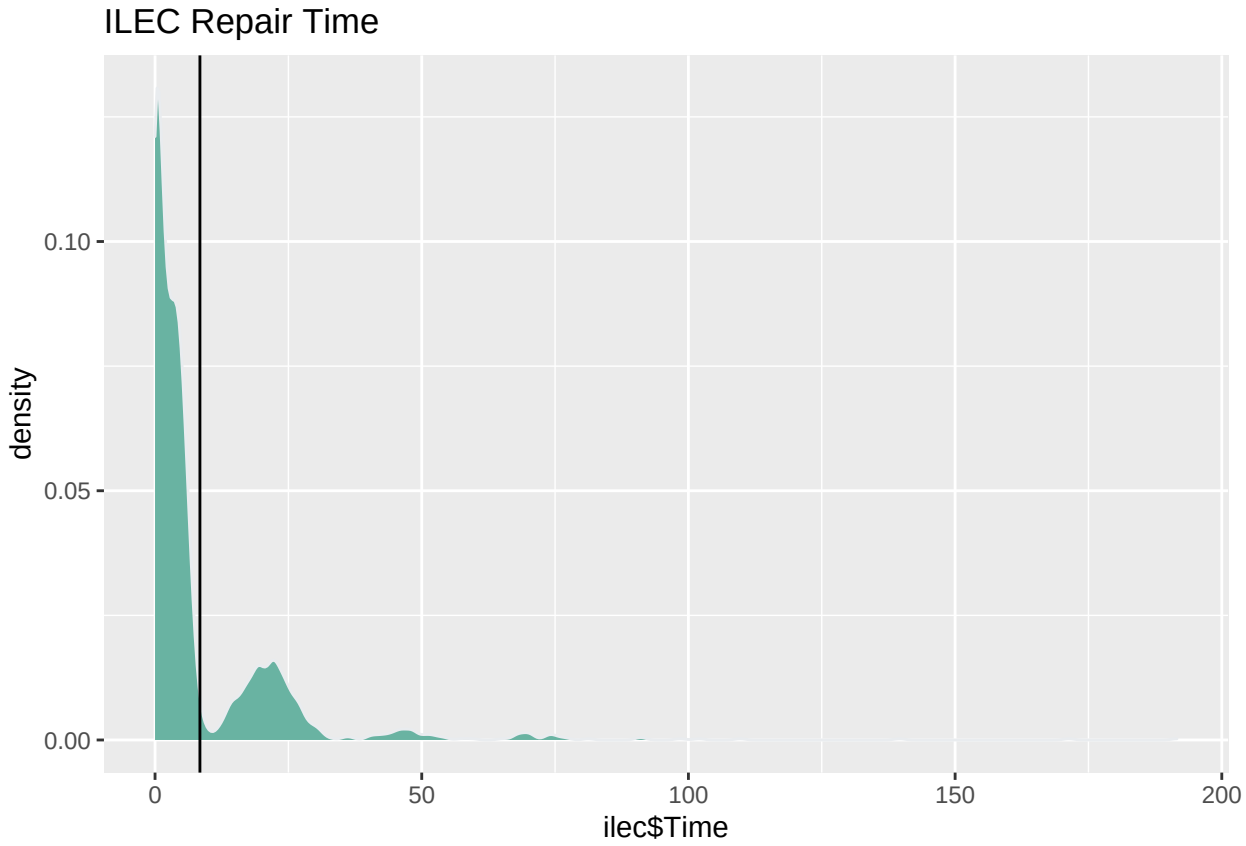
Helper functions

```
compute_tvalue <- function(sample) {
  sample_sderr <- sd(sample) / sqrt(length(sample))
  return (
    (mean(sample) - hypothesized_mean) / sample_sderr
  )
}

compute_mean <- function(sample) {
  return(mean(sample))
}
```

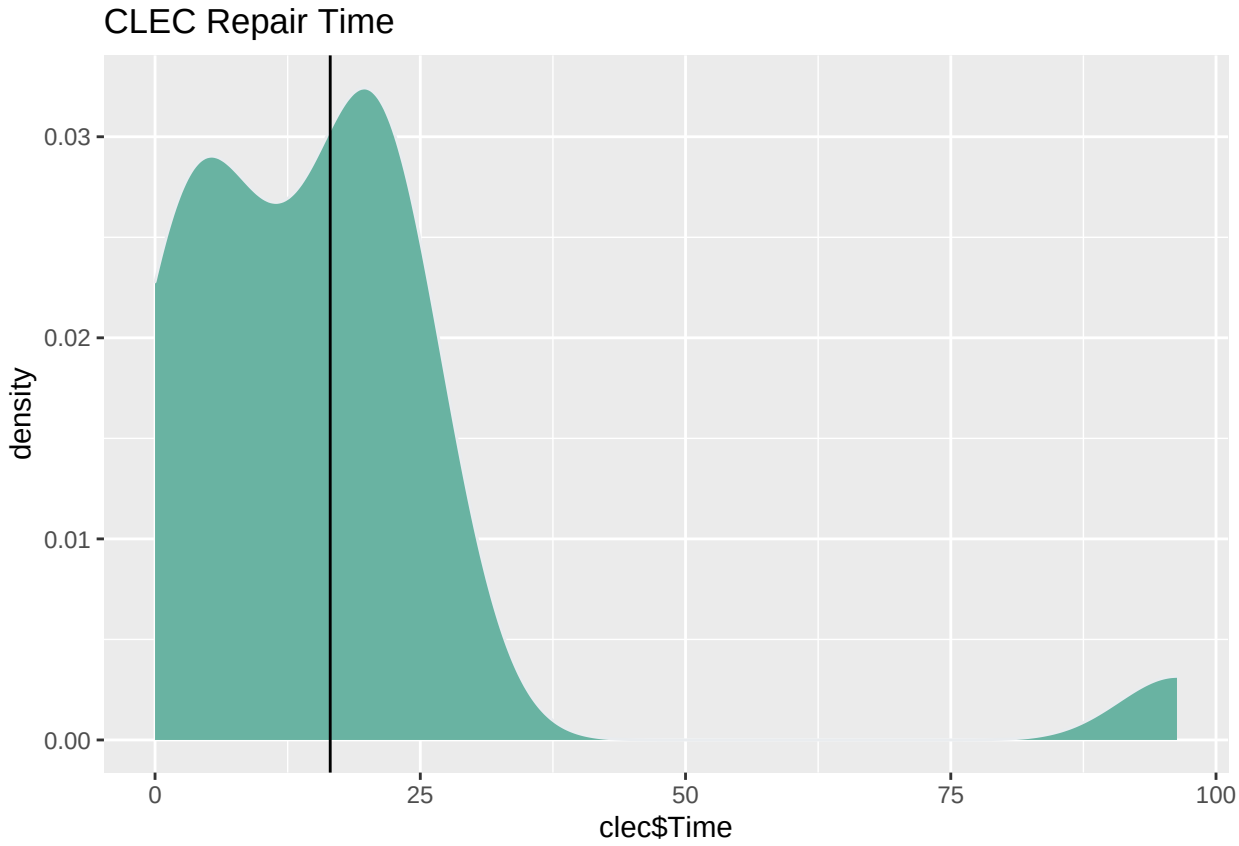
Question 1

a. Visualize Verizon's response times for ILEC vs. CLEC customers

```
ilec <- as.data.frame(verizon[verizon$Group == "ILEC", ])

ilec_mean <- mean(ilec$Time)

ggplot(mapping = aes(ilec$Time)) +
  geom_density(
    fill="#69b3a2",
    color="#e9ecef"
  ) +
  ggtitle("ILEC Repair Time") +
  coord_cartesian(xlim = c(0, max(ilec$Time))) +
  geom_vline(xintercept = ilec_mean)
```




```
clec <- as.data.frame(verizon[verizon$Group == "CLEC", ])

clec_mean <- mean(clec$Time)

ggplot(mapping = aes(clec$Time)) +
  geom_density(
    fill="#69b3a2",
    color="#e9ecef"
  ) +
  ggtitle("CLEC Repair Time") +
  coord_cartesian(xlim = c(0, max(clec$Time))) +
  geom_vline(xintercept = clec_mean)
```



b. Use the appropriate form of the t.test() function to test the difference between the mean of ILEC sample response times versus the mean of CLEC sample response times. From the output of t.test():

i. What are the appropriate null and alternative hypotheses in this case?

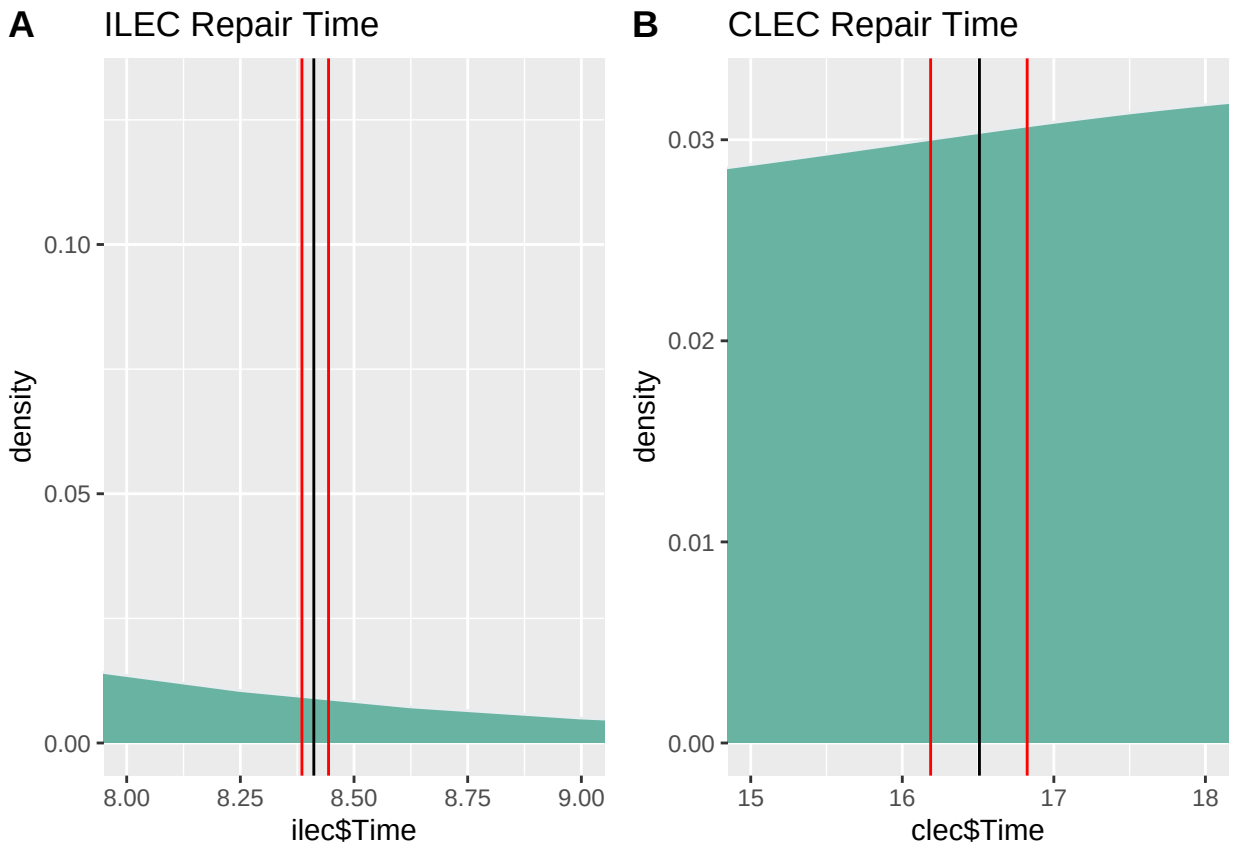
The null hypothesis is CLEC's respond and repair service time is as quick as ILEC's repair time.

The alternative hypothesis is CLEC's respond and repair service time is not as quick as ILEC's repair time.

```
ilec_samples <- replicate(1000, sample(ilec$Time, nrow(ilec), replace=TRUE))

ilec_means <- apply(ilec_samples, 2, FUN = compute_mean)
average_ilec_means <- mean(ilec_means)
ilec_means_sderr <- sd(ilec_means) / sqrt(length(ilec_means))
mean_ilec_ci99_low <- average_ilec_means - 2.58 * ilec_means_sderr
mean_ilec_ci99_high <- average_ilec_means + 2.58 * ilec_means_sderr
```

```
clec_samples <- replicate(1000, sample(clec$Time, nrow(clec), replace=TRUE))

clec_means <- apply(clec_samples, 2, FUN = compute_mean)
average_clec_means <- mean(clec_means)
clec_means_sderr <- sd(clec_means) / sqrt(length(clec_means))
mean_clec_ci99_low <- average_clec_means - 2.58 * clec_means_sderr
mean_clec_ci99_high <- average_clec_means + 2.58 * clec_means_sderr
```

```
ggarrange(ilec_plot, clec_plot,
          labels = c("A", "B"),
          ncol = 2, nrow = 1)
```



We can see from the graph above, we can reject our null hypothesis, saying **CLEC's respond and repair service time is as quick as ILEC's repair time**.

As the matter of fact, we should accept our alternative hypothesis, saying **CLEC's respond and repair service time is not as quick as ILEC's repair time**, instead