bm\_hw\_3

library(tidyverse)

## ── Attaching packages ────────────────────────────────────────────────────────────── tidyverse 1.2.1 ──

## ✔ ggplot2 3.0.0 ✔ purrr 0.2.5  
## ✔ tibble 1.4.2 ✔ dplyr 0.7.6  
## ✔ tidyr 0.8.1 ✔ stringr 1.3.1  
## ✔ readr 1.1.1 ✔ forcats 0.3.0

## ── Conflicts ───────────────────────────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

library(knitr)

Cigarette smoking continues to be a public health problem with major consequences on heart and lung diseases. Less is actually known about the consequences of quitting smoking. A recent study selected a group of 10 women working at a small medical practice, ages 50-64, that had smoked at least 1 pack/day and quit for at least 6 years (data “HeavySmoke.csv”).

1. The first question is to assess if their body mass index (BMI) has changed 6 years after quitting smoking. Perform an appropriate hypothesis test and interpret your findings. (5p)

smoke\_data = read\_csv("./HeavySmoke.csv") %>%  
 janitor::clean\_names() %>%  
 mutate(diff = bmi\_base - bmi\_6yrs)

## Parsed with column specification:  
## cols(  
## ID = col\_integer(),  
## BMI\_base = col\_double(),  
## BMI\_6yrs = col\_double()  
## )

diff\_mean = mean(smoke\_data$diff)  
diff\_sd = sd(smoke\_data$diff)  
n = 10  
t = (diff\_mean - 0)/(diff\_sd/sqrt(n))

$$
\bar{d}= \sum\_{i=1}^{n}\frac{d\_{i}}{n}=-3.36\\
s\_{d}=\sqrt{\frac{\sum\_{i=1}^{n}(d\_{i}-\bar{d})^{2}}{n-1}}=2.4627\\
n = 10\\
t=\frac{\bar{d}-0}{s\_{d}/\sqrt{n}}=-4.3145
$$

qt(0.975, n-1)

## [1] 2.262157

$$
t\_{n-1, 1-\alpha /2}=2.262157\\
\left | t \right |=4.3145\\
For\ \left | t \right |> t\_{n-1, 1-\alpha /2},\
reject\ H\_{0}
$$

t.test(smoke\_data$bmi\_base, smoke\_data$bmi\_6yrs, paired = TRUE)

##   
## Paired t-test  
##   
## data: smoke\_data$bmi\_base and smoke\_data$bmi\_6yrs  
## t = -4.3145, df = 9, p-value = 0.001949  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -5.121709 -1.598291  
## sample estimates:  
## mean of the differences   
## -3.36