Assignment 6: Who busts the Mythbusters?

Minh Tran

2021-05-29

Exercise 1

The explanatory variable is the 'group' and the response variable is 'yawn'. The value in the response variable that is considered as a success is 'yes'.

Exercise 2

The quantity that should be used to build the null distribution is iv.The average difference in fraction_yawned between the treatment and control groups

Exercise 3

```
specify(yawn ~ group, success = "yes")
```

Exercise 4

```
hypothesize(null = "independence")
```

Exercise 5

```
generate(reps = 10000,  type = "permute")
```

Exercise 6

```
calculate(stat = "diff in props",  order = combine("Treatment", "Control"))
```

Exercise 7

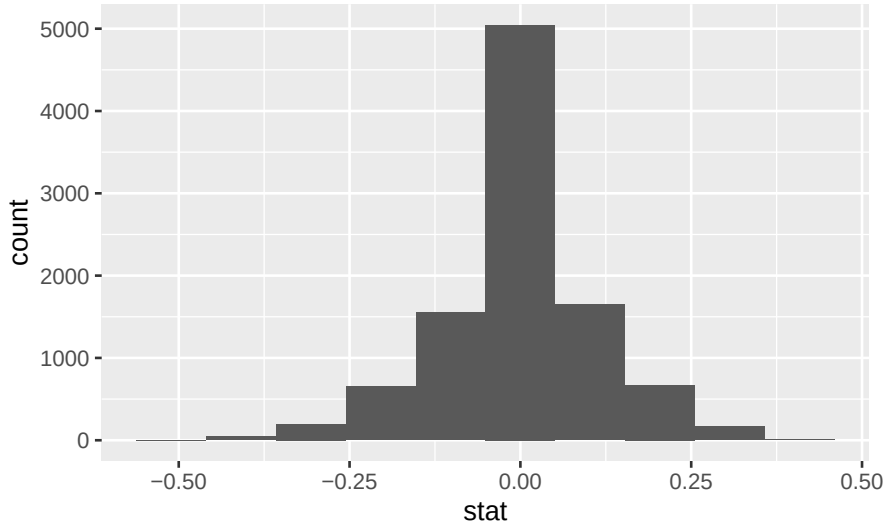
```
set.seed(42)
yawn_null <- experimental_data %>%
  specify(yawn ~ group, success = "yes") %>%
  hypothesize(null = "independence") %>%
  generate(reps = 10000,  type = "permute") %>%
  calculate(stat = "diff in props",  order = combine("Treatment", "Control"))
```

```
## Warning: `combine()` is deprecated as of dplyr 1.0.0.
## Please use `vctrs::vec_c()` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_warnings()` to see where this warning was generated.
```

Exercise 8

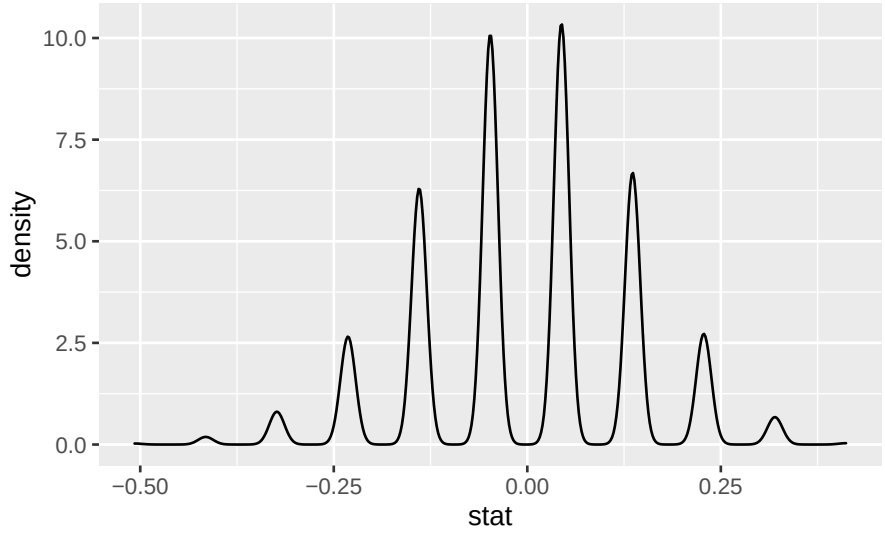
i.

```
yawn_null %>%
  ggplot() +
  geom_histogram(aes(x = stat), bins = 10)
```



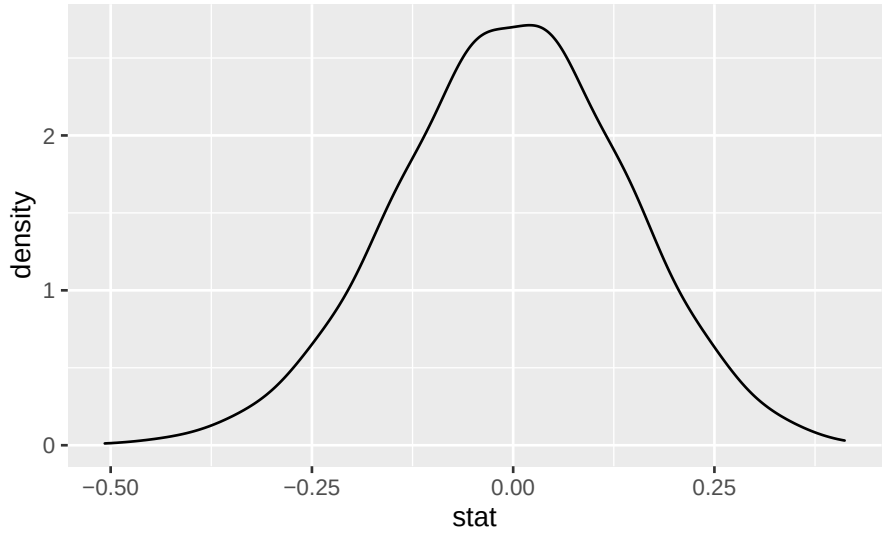
ii.

```
yawn_null %>%
  ggplot() +
  geom_density(aes(x = stat))
```



iii.

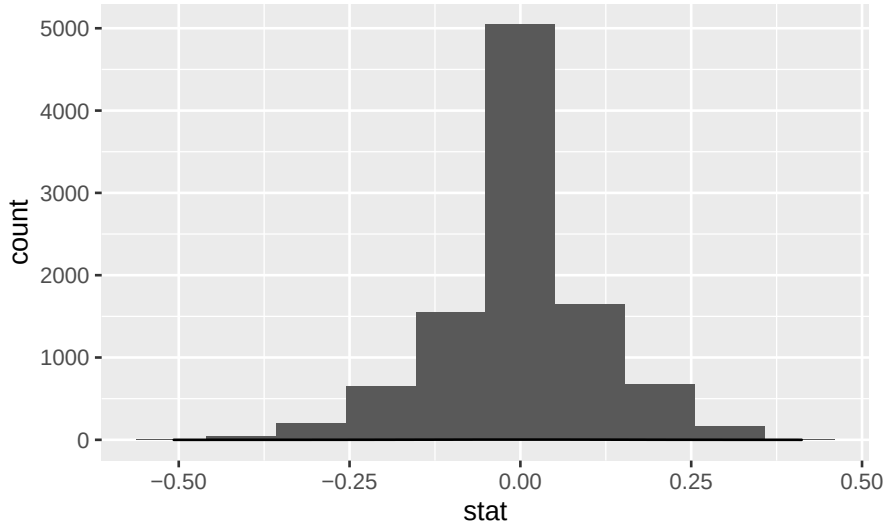
```
yawn_null %>%
  ggplot() +
  geom_density(aes(x = stat), adjust = 5)
```



iv. The center of distribution is around 0 which makes sense since the proportion of average difference in fraction_yawned between the treatment and control group are due to random chance, so it is a normal distribution.

Exercise 9

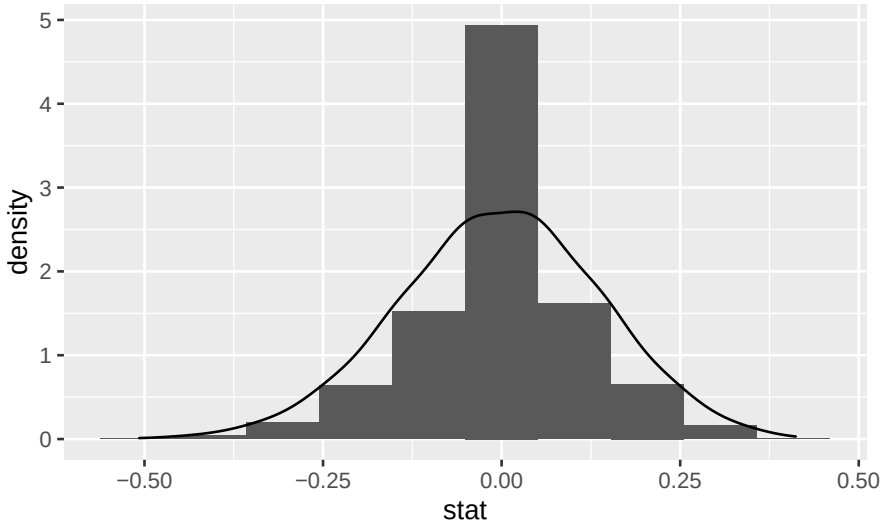
```
ggplot(data = yawn_null) +
  geom_histogram(aes(x = stat), bins = 10) +
  geom_density(aes(x = stat), adjust = 5)
```



The histogram is clearly present, but the density plot is not present.

Exercise 10

```
ggplot(data = yawn_null) +
  geom_histogram(aes(x = stat, y = ..density..), bins = 10) +
  geom_density(aes(x = stat), adjust = 5)
```



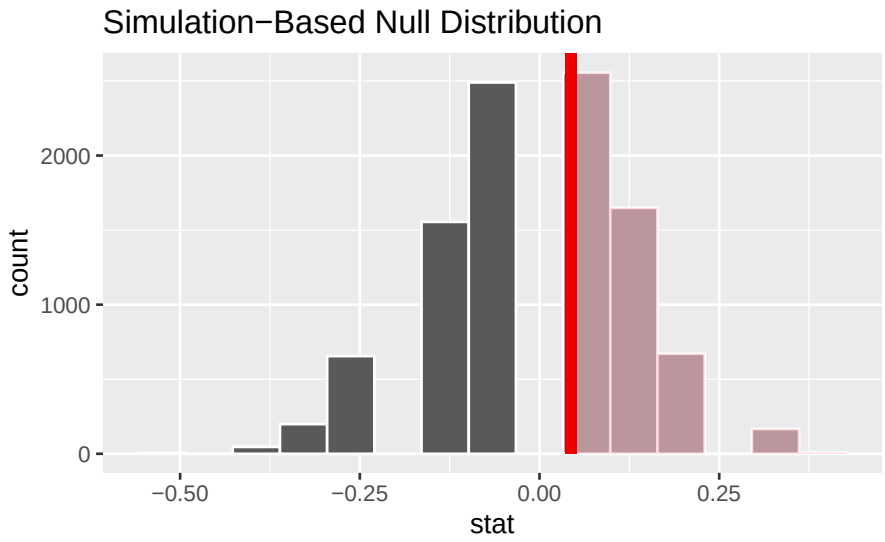
Exercise 11

```
yawn_obs_stat <- experimental_data %>%
  specify(yawn ~ group, success = "yes") %>%
  calculate(stat = "diff in props",  order = combine("Treatment", "Control"))
```

```
yawn_null %>%
  get_p_value(obs_stat = yawn_obs_stat, direction = "right")
```

p_value
0.5049

```
yawn_null %>%
  visualize() +
  shade_p_value(obs_stat = yawn_obs_stat, direction = "right")
```



Based on the results, the p-value of 0.50 is greater than $\alpha = 0.05$, so we fail to reject the null hypothesis. So, we conclude that there are no significant difference in the amount of yawns between the control and treated group.