Inference for numerical data Solution

MATH224 - Intro to Stat

Exercise 1 (5 Points)

1 Point H0: p = 0.66, HA: p not = 0.66

OR

$H_0 : p = 0.66, H_A : p \neq 0.66$

```
yrbss <- yrbss %>%
  mutate(physical_3plus = if_else(physically_active_7d > 2, "yes", "no")) #1 Point

yrbss%>%
  filter(!is.na(physical_3plus))%>%
  prop_test(response = physical_3plus,
            success = "yes",
            p = 0.66,
            alternative = "two-sided",
            z = T) #1 Point for hypothesis test
```

```
## # A tibble: 1 x 3
##   statistic p_value alternative
##       <dbl>   <dbl> <chr>
## 1      2.22  0.0263 two.sided
```

```
yrbss%>%
  filter(!is.na(physical_3plus))%>%
  prop_test(response = physical_3plus,
            success = "yes",
            alternative = "two-sided",
            z = T,
            conf_int = T,
            conf_level = 0.98) #1 Point for confidence interval
```

```
## # A tibble: 1 x 5
##   statistic p_value alternative lower_ci upper_ci
##       <dbl>   <dbl> <chr>          <dbl>    <dbl>
## 1      39.0       0 two.sided      0.660    0.679
```

0.5 Points Since the p-value is greater than 0.02, we fail to reject the null hypothesis.

0.5 Points CI: (0.6596, 0.6785)

Exercise 2 (3 Points)

1 Point H0: pmale - pfemale = 0, HA: pmale - pfemale not = 0

OR

$H_0 : p_{male} - p_{female} = 0, H_A : p_{male} - p_{female} \neq 0$

```
yrbss%>%
  filter(!is.na(physical_3plus), !is.na(gender))%>%
  prop_test(response = physical_3plus,
            success = "yes",
            explanatory = gender,
            order = c("male","female"),
            z = TRUE,
            conf_int = TRUE,
            conf_level = 0.95)#1 Point
```

```
## # A tibble: 1 x 5
##   statistic   p_value alternative lower_ci upper_ci
##       <dbl>     <dbl> <chr>          <dbl>    <dbl>
## 1      23.2 4.50e-119 two.sided      0.196    0.231
```

0.5 Points Since p is less than 0.05, we reject the null hypothesis.

0.5 Points CI: (0.1962, 0.2313)

Exercise 3 (3 Points)

1 Point H0: mu = 66.82, HA: mu not = 66.82

OR

$H_0 : \mu = 66.82, H_A : \mu \neq 66.82$

```
yrbss%>%
  filter(!is.na(weight))%>%
  t_test(response = weight,
         mu = 66.82,
         conf_int = TRUE,
         conf_level = 0.95) #1 Point
```

```
## # A tibble: 1 x 7
##   statistic  t_df  p_value alternative estimate lower_ci upper_ci
##       <dbl> <dbl>    <dbl> <chr>          <dbl>    <dbl>    <dbl>
## 1      7.21 12578 5.86e-13 two.sided       67.9     67.6     68.2
```

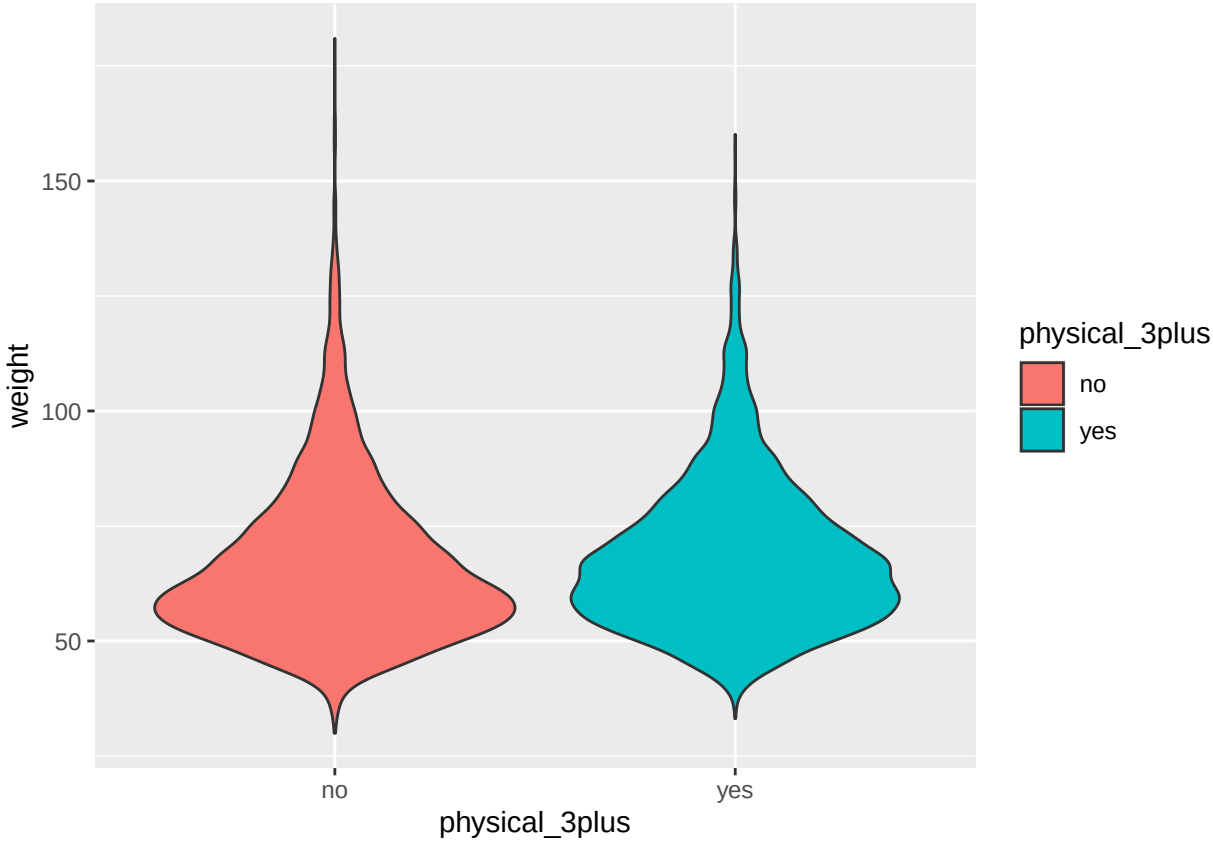
0.5 Points Since p-value is less than 0.05, we reject the null hypothesis.

0.5 Points CI: (67.61, 68.20)

Exercise 4 (3 Points)

2 Points Explanation: There's seems to be a slight difference in the violin plots. Physically inactive students have a higher standard deviation compared to the physically active students. Physically inactive students have a slightly lower mean weight.

```
yrbss %>%
  filter(!is.na(physical_3plus), !is.na(weight))%>%
  ggplot(aes(x = physical_3plus, y = weight, fill = physical_3plus))+
  geom_violin() #1 Point
```



Exercise 5 (3 Points)

2 Points There is an observable difference. But the difference isn't big enough for us to deem it statistically significant without an hypothesis test.

```
yrbss %>%
  filter(!is.na(physical_3plus), !is.na(weight))%>%
  group_by(physical_3plus) %>%
  summarise(mean_weight = mean(weight)) #1 Point
```

```
## # A tibble: 2 x 2
##   physical_3plus mean_weight
##   <chr>                <dbl>
```

```
## 1 no                    66.7
## 2 yes                   68.4
```

Exercise 6 (3 Points)

1 Point H0: mu_yes - mu_no = 0, HA: mu_yes - mu_no not = 0

OR

$H_0 : \mu_{yes} - \mu_{no} = 0, H_A : \mu_{yes} - \mu_{no} \neq 0$

```
yrbss%>%
  filter(!is.na(weight))%>%
  t_test(response = weight,
         explanatory = physical_3plus,
         order = c("yes","no"),
         mu = 0,
         conf_int = TRUE,
         conf_level = 0.95) #1 Point
```

```
## # A tibble: 1 x 7
##   statistic  t_df      p_value alternative estimate lower_ci upper_ci
##       <dbl> <dbl>        <dbl> <chr>          <dbl>    <dbl>    <dbl>
## 1      5.35 7479. 0.0000000891 two.sided       1.77     1.12     2.42
```

0.5 Points Since p-value is less than 0.05, we reject the null hypothesis.

0.5 Points CI: (1.12, 2.42)