Lesson 7: Do Diets Really Work? (continued #2)

LTC James K Starling

Last compiled on 24 January, 2022

Background

Recall that in Lesson 5 and 6 we investigated the impact diet (Vegetarian, Keto, and None) has on weight loss (kg lost).

```
library(tidyverse)
library(pwr) # see https://www.statmethods.net/stats/power.html for good tutorial.

## Load data
dietdat <- read_csv('https://raw.githubusercontent.com/jkstarling/MA376/main/WeightLoss.csv')

# glimpse(dietdat)
## change data types appropriately.
dietdat <- dietdat %>% mutate(diet = as.factor(diet)) %>%
  mutate(exercise = as.factor(exercise))
```

1. What was our null and alternative hypotheses (in words and symbols, assuming we are going to conduct ANOVA)?

The null hypothesis is there is no association between weight loss and diet type, the alternative hypothesis is that there is an association between weight loss and diet type.

$$H_0 : \mu_{Vegetarian} = \mu_{Keto} = \mu_{None}$$

$$H_A : \text{at least one of the } \mu \text{ is different from the others.}$$

2. For the diet study, what is a Type I error/Type I error rate? Type II error/Type II error rate?

Type I error (false positive)- rejecting H_0 when H_0 is TRUE. In the diet study, a type I error will occur when we say there is an association between diet and weight loss when there actually is no association. The type I error rate is the probabilty of saying the above (also known as the significance level (α).

Type II error (false negative)- failing to reject H_0 when H_0 is FALSE. In the diet study a Type II error will occur when say there is no association between diet and weight loss when in fact there is an association. The Type II error rate is the probability of saying the above (β).

3. What is statistical power?

Statistical power is the probability that the researchers find evidence against the null hypothesis and in favor of the alternative hypothesis when the alternative hypothesis is TRUE (p.103). $(1 - \beta)$

    4. What elements of a study can affect statistical power?

These elements are sample size (n) and amount of unexplained variation (# of groups, R^2. Additionally, the significance level will affect the power as this will affect the probability that we reject the null hypothesis (whether or not H_0 is TRUE)

    5. Focusing on the F-statistic, we can explore the impact of how the elements will affect statistical power.

$$F = \left[\frac{R^2}{1 - R^2} \right] \times \left[\frac{n - \text{\# of groups}}{\text{\# of groups - 1}} \right]$$

    Using the code from Lesson 6:

```
multi.lm <- lm(wloss~0+diet, data=dietdat)
SST <- sum((dietdat$wloss - mean(dietdat$wloss))^2)
SSE <- sum((dietdat$wloss -multi.lm$fitted.values)^2)
SSM <- SST-SSE
n <- nrow(dietdat)
dfmod <- 2
dfred <- n - dfmod - 1
Fstat <- (SSM/dfmod) / (SSE/dfred); Fstat
```

```
## [1] 10.68527
```

```
pmm <- pf(Fstat, dfmod, dfred, lower.tail = FALSE); pmm
```

```
## [1] 4.158189e-05
```

```
Rsq <- SSM/SST; Rsq
```

```
## [1] 0.1077304
```

We saw that we strong evidence to reject the null hypothesis ('rather large' F-statistic and 'rather small' p-value).

Let's see what happens when we reduce the number of observations.

```
dietdat2 <- dietdat %>% group_by(diet, exercise)  %>%  slice_head(n=10)
```

6. Create a multiple-means model using 'dietdat2' data. Calculate the F statistic and p-value. Use the names 'multi.lm2', 'Fstat2', 'pmm2', and 'Rsq2'. Explain the results.

```
multi.lm2 <- lm(wloss~0+diet, data=dietdat2)
SST <- sum((dietdat2$wloss - mean(dietdat2$wloss))^2)
SSE <- sum((dietdat2$wloss -multi.lm2$fitted.values)^2)
SSM <- SST-SSE
n <- nrow(dietdat2)
dfmod <- 2
dfred <- n - dfmod - 1
Fstat2 <- (SSM/dfmod) / (SSE/dfred); Fstat2
```

```
## [1] 7.43918
```

```
pmm2 <- pf(Fstat2, dfmod, dfred, lower.tail = FALSE); pmm2
```

```
## [1] 0.001041173
```

```
Rsq2 <- SSM/SST; Rsq2
```

```
## [1] 0.1460404
```

We see that the F statistic decreased to 7.44 and the corresponding p-value to 0.001. We see that although we would still reject the null hypothesis at the $\alpha = 0.05$ significance level, we have reduced the evidence towards doing so. We also see that the R^2 value increased slightly from 0.1077 to 0.01460.

7. Let's see what happens when we reduce the number of groups by combining Vegetarian and Keto diets.

Create a multiple-means model using 'dietdat3'. Calculate the F statistic and p-value. Use the names 'multi.lm3', 'Fstat3', 'pmm3', and 'Rsq3'. Explain the results.

```
dietdat3 <- dietdat %>% mutate(kv=fct_collapse(diet, KV = c("Keto", "Vegetarian")))
# add updated code here
multi.lm3 <- lm(wloss~0+kv, data=dietdat3)
SST <- sum((dietdat3$wloss - mean(dietdat3$wloss))^2)
SSE <- sum((dietdat3$wloss -multi.lm3$fitted.values)^2)
SSM <- SST-SSE
n <- nrow(dietdat3)
dfmod <- 1    ######
dfred <- n - dfmod - 1
Fstat3 <- (SSM/dfmod) / (SSE/dfred); Fstat3
```

```
## [1] 13.85284
```

```
pmm3 <- pf(Fstat3, dfmod, dfred, lower.tail = FALSE); pmm3
```

```
## [1] 0.0002649767
```

```
Rsq3 <- SSM/SST; Rsq3
```

```
## [1] 0.07220555
```

We see that the F statistic increased to 13.85 but the corresponding p-value to increased slightly to 2.6e-4. We also see that the R^2 value decreases slightly from 0.1077 to 0.0722.

    8. Let's see what happens when we reduce the number of groups by combining Vegetarian and None diets (not recommended… why?).

Create a multiple-means model using 'dietdat4'. Calculate the F statistic and p-value. Use the names 'multi.lm4', 'Fstat4', 'pmm4', and 'Rsq4'. Explain the results.

```
dietdat4 <- dietdat %>% mutate(nv=fct_collapse(diet, NV = c("None", "Vegetarian")))
# add updated code here
multi.lm4 <- lm(wloss~0+nv, data=dietdat4)
SST <- sum((dietdat4$wloss - mean(dietdat4$wloss))^2)
SSE <- sum((dietdat4$wloss -multi.lm4$fitted.values)^2)
SSM <- SST-SSE
n <- nrow(dietdat4)
dfmod <- 1    ######
dfred <- n - dfmod - 1
Fstat4 <- (SSM/dfmod) / (SSE/dfred); Fstat4
```

```
## [1] 0.1485173
```

```
pmm4 <- pf(Fstat4, dfmod, dfred, lower.tail = FALSE); pmm4
```

```
## [1] 0.7004165
```

```
Rsq4 <- SSM/SST; Rsq4
```

```
## [1] 0.0008336711
```

We see that the F statistic decreases substantially with an increase in the p-value, while the R^2 value decreases to 0.00083. We should expect this since in the context of the problem we would not recommend combining None with any other diet factor level.

9. Now we will go over HW #1.6.11 to show how calculate power in R:

Suppose you are testing to see whether, on average, heart rates measured in beats per minute (bpm), after jumping for 30 seconds are higher than after sitting for 30 seconds. There are 20 students from your statistics class who have volunteered to participate. You randomly assign 10 to the jumping treatment and 10 to the sitting treatment. You may assume that the SD of heart rates within each condition is 8 beats per minute.

a. *Use the Comparing Two Populations (we will use R) applet with both means set to 80. Use a significance level of Which of the following is the best rejection region for detecting that the population mean heart rate after jumping is higher than after sitting?*

How do we account for this without the applet? ANS: Use the standardized t-statistic:

$$t^*_{df} = \frac{\overline{\Delta} - \mu_0}{s/\sqrt{n}}$$

or rewriting it (with $_0 = 0$):

$$\overline{\Delta} = t_{df}^* \cdot \frac{s}{\sqrt{n}}$$

```
qt(0.025, df = 9, lower.tail = FALSE) * 8 / sqrt(10)
```

```
## [1] 5.722855
```

The rejection region for detecting the jumping heart rate being higher than the sitting rate is a difference of 5.7 bpm.

    b. If you find a difference in the mean heart rate of 7 bpm, what will you conclude about the two treatments?

7 bpm > 5.7 bpm, so we reject the null hypothesis (that the heart rates are the same).

    c. If you find a difference in the mean heart rate of 4 bpm, what will you conclude about the two treatments?

4 bpm < 5.7 bpm, so we fail to reject the null hypothesis (that the heart rates are the same).

    d. What is the probability of making a Type I error?

(Trick question?) This is simply the significance level ($\alpha = 0.05$).

    e. What is the probability of rejecting the null hypothesis if the difference in mean heartbeats is 5 bpm?

```
library(pwr)
n_students <- 10
sd <- 8
del <- 5
pwr.t.test(n=n_students, d=del/sd, sig.level = 0.05, alternative = "greater")
```

```
##
##      Two-sample t test power calculation
##
##              n = 10
##              d = 0.625
##      sig.level = 0.05
##          power = 0.3821113
##    alternative = greater
##
## NOTE: n is number in *each* group
```

    f. What is the probability of rejecting the null hypothesis if the difference in mean heartbeats is 10 bpm?

```
del <- 10
pwr.t.test(n=n_students, d=del/sd, sig.level = 0.05, alternative = "greater")
```

```
##
##      Two-sample t test power calculation
##
##             n = 10
##             d = 1.25
##       sig.level = 0.05
##         power = 0.8514775
##     alternative = greater
##
## NOTE: n is number in *each* group
```

g. (Fill out the table… find the power for bpm = 5, 10, 15, 20, 25)

```
del <- 15
pwr.t.test(n=n_students, d=del/sd, sig.level = 0.05, alternative = "greater")
```

```
##
##      Two-sample t test power calculation
##
##             n = 10
##             d = 1.875
##       sig.level = 0.05
##         power = 0.9914222
##     alternative = greater
##
## NOTE: n is number in *each* group
```

```
del <- 20
pwr.t.test(n=n_students, d=del/sd, sig.level = 0.05, alternative = "greater")
```

```
##
##      Two-sample t test power calculation
##
##              n = 10
##              d = 2.5
##      sig.level = 0.05
##          power = 0.9999011
##    alternative = greater
##
## NOTE: n is number in *each* group
```

```
del <- 25
pwr.t.test(n=n_students, d=del/sd, sig.level = 0.05, alternative = "greater")
```

```
##
##      Two-sample t test power calculation
##
##              n = 10
##              d = 3.125
##      sig.level = 0.05
##          power = 0.9999998
##    alternative = greater
##
## NOTE: n is number in *each* group
```