Lesson 14: Hopelessness and Exercise

CDT I.M. Smrt

A recent study (Dunn et al., J. Cardiopulm. Rehab. Prevent., 2017) gathered data on patients who had recently had a heart attack requiring hospitalization. Before patients were discharged from the hospital, patients took a series of surveys which included measures of chronic hopelessness. Chronic hopelessness was measured based on a series of Likert scale questions which were averaged together yielding a Chronic Hopelessness score of 1–4. Higher scores indicate a more hopeless feeling. Researchers were interested in learning about whether exercise after a heart attack was associated with hopelessness levels eight months later. Two-hundred and twenty-three (n = 223) patients have data measuring hopelessness eight months after hospitalization, as well as information about their exercise habits during the eight-month period following hospitalization. Exercise was classified as "Regular home exerciser" (exercised at home at least three times per week) or "Not regular home exerciser" (did not exercise at home at least three times per week).

```
### Loading packages
library(tidyverse)
library(ggResidpanel)
library(car)
## Load data
hope <- read.table('https://raw.githubusercontent.com/jkstarling/MA376/main/hopelessness8months_2group.t

# setwd('C:/Users/james.starling/OneDrive - West Point/Teaching/MA376/JimsLessons/Block II/LSN14/')
# hope <- read.table('hopelessness8months_2group.txt', header = TRUE, stringsAsFactors = TRUE)

# add effect coding for Exercise and BaselineGroup
contrasts(hope$Exercise) <- contr.sum
contrasts(hope$BaselineGroup) <- contr.sum
```

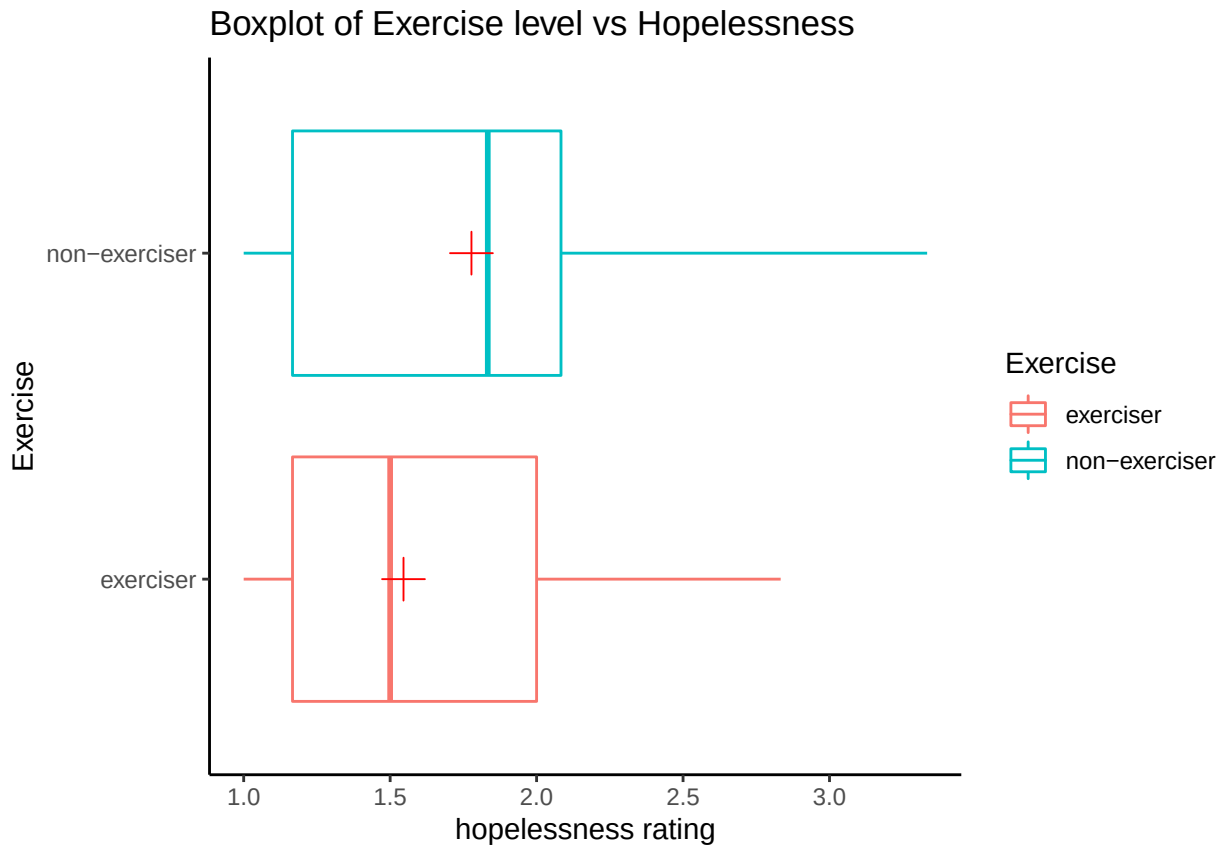
(1) State the null and alternative hypothesis for the research question to learn whether exercise after a heart attack was associated with hopelessness levels eight months later. Identify the explanatory and response variables, and whether they are categorical or quantitative.

(2) Was this an observational study or a randomized experiment? If the latter, are there any blocking variables to account for?

(3) Plot the distributions of hopelessness 8 months after a heart attack between the two exercise levels. Also provide the mean, standard deviation, and count for each level of exercise.

Is there preliminary evidence to suggest that there is a relationship between exercise status and hopelessness level? What are the estimated effects for exercisers and non-exercisers?

```
hope %>% ggplot(aes(x=hopelessness8, y = Exercise, color=Exercise)) +
  geom_boxplot() +
  labs(title="Boxplot of Exercise level vs Hopelessness", x='hopelessness rating') +
  stat_summary(fun=mean, geom="point", shape=3, size=5, color="red", fill="red") +
  theme_classic()
```




```
#
# hopeless.means <- xxx
# hopeless.means
#
# # estimated effects
#
```

Single-variable model

(4) Conduct a one-variable analysis evaluating the association between exercise and hopelessness 8 months after a heart attack and provide an ANOVA table.

How much of the variation in hopelessness scores is explained by whether or not the individual exercises? Save the sum of squared error total as `SST_ex`.

```
# hope.lm <- hope %>% lm(xxx ~ xxxx, data=.)
# summary(hope.lm)
#
# hope.anv <- anova(hope.lm)
# hope.anv
#
# SST_ex <- xxxx
# SST_ex
#
# r.sq <- xxxx
# r.sq
```

Two-variable model

(5) Conduct a two-variable analysis evaluating the association between exercise and hopelessness 8 months after a heart attack and provide an ANOVA table.

Give the prediction equation and residual standard error.

What variables are significant (at the $\alpha = 0.05$ level)? (Load the package `car` and use `Anova(hope.lm2, type="3")`)

```
# hope.lm2 <- hope %>% lm(xxx)
# summary(hope.lm2)
#
# hope.anv2 <-xxxx
# hope.anv2
#
# se2 <- xxxx
# se2
```

(6) Provide the model R^2. What is the value of the covariation? (Hint: You will need to use `SST_ex` calculated above.)

```
# SSexercise <- xxxx
# SSbasegroup <- xxx
# SSE <- xxx
#
# SScovariation2 <- xxxx
# SScovariation2
#
# # rsquared
# r.sq2 <- xxxx
# r.sq2
```

Interaction

(7) Create an interaction plot between Baseline hopelessness (x-axis) and Exercise (group & color). What story is being told with this graph?

```
# hope.int <- hope %>% group_by(xxx, xxx) %>% summarise(mn=mean(xxx))
#
# hope.int %>% ggplot(aes(x=xxx, y = xxx, group=xxx, color=xxx)) +
#   geom_line() + geom_point() +
#   labs(title="Interaction Plot of mean Hopelessness accouting for BaseLine Group and Exercise")
```

(8) Conduct a two-variable analysis evaluating the association between exercise and hopelessness 8 months after a heart attack WITH INTERACTION. Provide an ANOVA table.

Give the prediction equation and residual standard error. What variables are significant (at the $\alpha = 0.05$ level)?
What is the R^2 value?

```
# hope.lm3 <- hope %>% lm(xxxx)
# summary(hope.lm3)
#
# hope.anv3 <- Anova(hope.lm3, type="3")
# hope.anv3
#
# se3 <- xxxx
# se3
```

(9) Explain how to get the degrees of freedom for SSM (not shown), SSexercise, SSBaselineGroup, SSinteraction, and SSresiduals.

(10) What are the null and alternative hypotheses for the "model" F-test (with interaction)? What do you conclude about the model?

(11) What are the null alternative hypotheses for the interaction F-test? What do you conclude about the interaction between baseline hopelessness and exercise? What percentage of the variation in eight-month hopelessness is explained by this interaction?

(12) Calculate the SS for the Model (SSM). What is the covariation of this model? How does it compare to the previous model?

```
# SScovariation3 <- xxxx;
# SScovariation3
```

(13) How does the SSError change when comparing the two-variable model with and without the interaction? Why?

(14) Discuss whether or not the validity conditions are met for this analysis. Cite specific characteristics of graphs or numbers from the output to support your statements.

```
# resid_panel(xxxx)
# hope %>% group_by(xxxx, xxxx) %>%
#   summarise(xxxx)
```

(15) What are the null and alternative hypotheses for the "model" F-test? What do you conclude about the model?

(16) What are the null alternative hypotheses for the interaction F-test? What do you conclude about the interaction between baseline hopelessness and exercise? What percentage of the variation in eight-month hopelessness is explained by this interaction?

** Bonus: How to do Type III SS w/o car package

```
# from bottom of page: https://mcfromnz.wordpress.com/2011/03/02/anova-type-iiiiii-ss-explained/
#
# model <-
#
#
# # wrong way:
# aov(model)
# anova(model)
```