Worksheet 6

Peyton Hall

02/21/2025

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(nlme)
```

```
##
## Attaching package: 'nlme'
##
## The following object is masked from 'package:dplyr':
##
##     collapse
```

```
library(multcomp)
```

```
## Loading required package: mvtnorm
## Loading required package: survival
## Loading required package: TH.data
## Loading required package: MASS
##
## Attaching package: 'MASS'
##
## The following object is masked from 'package:dplyr':
##
##     select
##
##
## Attaching package: 'TH.data'
##
## The following object is masked from 'package:MASS':
##
##     geyser
```

Activity 1 We would like to study the effect of pizza consumption and the timing of eating pizza on mood. The study recruited 16 volunteers, and randomly assigned 8 of them to eat Pop Johns' Pizza, and the other 8 eat Domino's Pizza. All 16 volunteers had Pizza for both breakfast and dinner. The researchers recorded their mood scores 3 hours after they had Pizza. The data is on the worksheet. Perform an appropriate analysis to test the main effect of the type of Pizza, the main effect of timing, and the interaction between type of Pizza and timing

Individual, Type of Pizza, After Breakfast, After dinner 1 Pop John's, 52, 77 2 59, 69 3 66, 68 4 43, 71 5 Dominos, 55, 82 6 61, 69 7 62, 80 8 39, 59

```
individuals <- 1:8
pizza <- rep(c("Pop Johns", "Dominos"), each = 4)
meals <- rep(c("After breakfast", "After dinner"), each = 8)
moodscore <- c(52,59,66,43, 55,61,62,39, 77,69,68,71, 82,69,80,59)
mydf <- data.frame(individuals, pizza, meals, moodscore)
model2 <- lme(moodscore~pizza+meals+pizza*meals, random = ~1+meals|individuals, data=mydf)
anova(model2)
```

```
##             numDF denDF  F-value p-value
## (Intercept)     1     6 649.1883  <.0001
## pizza           1     6   0.0142  0.9089
## meals           1     6  22.2650  0.0033
## pizza:meals     1     6   0.0748  0.7936
```

a) What are the f statistic and p-value for testing the main effect of type of Pizza? f = 0.0142; p-value = 0.9089

b) What conclusion can you make about the main effect of Pizza? $H_0 : \mu_{PJ} = \mu_D$ vs $H_a : \mu_{PJ} \neq \mu_D$ Fail to reject H0; there is no significant difference in mood score between Pop John's and Dominos pizza.

c) What are the F statistic and p-value for testing the main effect of timing/meal? f = 22.2650; p-value = 0.0033

d) What conclusion can you make about the main effect of timing/meal? $H_0 : \mu_{AB} = \mu_{AD}$ vs $H_a : \mu_{AB} \neq \mu_{AD}$ Reject H0; there is significant mood score difference between breakfast and dinner

e) What is the p-value for the interaction term? p-value = 0.7936 Fail to reject H0; there is no significant interaction between pizza and meal.

2. Does physical exercise alleviate depression? We find some depressed people and check that they are all equivalently depressed to begin with. Then we allocate each person randomly to one of the three groups: no exercise; 20 minutes of jogging per day; or 60 minutes of jogging per day. At the end of a month, we ask each participant to rate how depressed they feel on a Likert scale from 1 (totally miserable) through 100 (ecstatically happy). Here is the data. Use 0.05 significance level to test the claim that the depression level (score) are significantly different between the three exercise groups. No exercise, 23, 26, 32, 37, 29 20-min exercise, 22, 27, 39, 35, 46 60-min exercise, 59, 66, 38, 49, 56

```
exercise <- rep(c("No exercise", "20-min exercise", "60-min exercise"), each = 5)
depscore <- c(23, 26, 32, 37, 29, 22, 27, 39, 35, 46, 59, 66, 38, 49, 56)
depdf <- data.frame(exercise, depscore)
kruskal.test(depscore~exercise, data = depdf)
```

```
##
##  Kruskal-Wallis rank sum test
```

```
##
## data:  depscore by exercise
## Kruskal-Wallis chi-squared = 8.34, df = 2, p-value = 0.01545
```

```
pairwise.wilcox.test(depscore, exercise, p.adj="none")
```

```
##
##  Pairwise comparisons using Wilcoxon rank sum exact test
##
## data:  depscore and exercise
##
##                 20-min exercise 60-min exercise
## 60-min exercise 0.0317          -
## No exercise     0.5476          0.0079
##
## P value adjustment method: none
```

Step 1: State the null and alternative hypotheses H_0 : Depression no exercise = Depression twenty min = Depression sixt vs H_a : At least two are different

Step 2: What test would you choose? Kruskal-Wallis rank sum test

Step 3: What is the test statistic and p-value? t = 8.34; p-value = 0.01545

Step 4: What is your decision? Reject H0; there is a sig. diff. bt. at least 2 types of exercises.

Step 5: What can you conclude from post-hoc (follow up) test? The significant difference is bt. 20 min and 60 min bt. none and 60 min