**Name: Marwin Carmo**

Provide your answers and the code used for each question in the text boxes below. The boxes are approximately sized to match the amount of space you will need for each answer, but you are welcome to expand or shrink the text boxes as needed. You may submit answers as text or screenshots. Note: The text has been set to blue inside the text boxes. This is intentional and will make it easier for the TA’s to see your answers.

1. Why do we say that in repeated measures ANOVA time is treated as a categorical variable? Explain your answer. [1 pt.]

Answer:

Time is treated as a categorical variable because the focus is on comparing means across different levels of repeated measurement, where each level refers to a different point in time.

**Computational Problems**

Use the “hw7\_2023.csv” file for this assignment. The data was part of a study aimed to increase the physical fitness level of pregnant women. There is one between-subjects grouping variable, category. Half the women in the study received standard care and half were put into a training regimen. The researchers were interested in whether these different kinds of interventions impacted physical activity. We’ll use step count (a proxy of physical activity) as our outcome variable. Each woman was measured on their step count at weeks 14 (t1), 28 (t2) and 36 (t3) of their pregnancy.

2. Read in the dataset and convert it from wide format data to long format. Make sure that the independent variables and the ID variable are coded as factors. [1 pt.]

Code/Syntax:

hw7\_2023 <- readr::read\_csv("23-Fall/psc204-fq23/lab7/hw7\_2023.csv")

hw\_long <- hw7\_2023 |>

tidyr::pivot\_longer(

cols = t1:t3,

names\_to = "time",

values\_to = "value"

) |>

dplyr::mutate(

dplyr::across(id:time, factor)

)

3. Create a plot that shows the group means across time points. [2 pt.]

**Extra Credit:** You get 1 extra point if you include an error bar in your plot.

Answer:

A graph with red and blue lines

Description automatically generated

Code/Syntax:

hw\_long |>

dplyr::with\_groups(c(time, category),

summarise,

mean\_steps = mean(step\_count, na.rm = TRUE),

n = n(),

se = sd(step\_count, na.rm=TRUE)/sqrt(n)) |>

ggplot(aes(x = time, y = mean\_steps, color = category)) +

geom\_point() +

geom\_errorbar(aes(ymin = mean\_steps - se,

ymax = mean\_steps + se),

width = 0.1) +

scale\_x\_discrete(labels = paste("Time", 1:3)) +

labs(x = "Time", y = "Mean Steps") +

theme\_bw()

4. Conduct a repeated measures ANOVA to test if there is a difference in physical activity over time and category. Is there a significant effect for category? Time? Their interaction? Don’t forget to consider whether each variable is a between- or within-subjects variable. [2 pts.]

Answer:

There is no significant main effect of category, but the time variable and its interaction with category had statistically significant effects on the outcome.

Code/Syntax:

rma <- aov(step\_count ~ time + category + time \* category + Error(id/ (time)), data = hw\_long)

summary(rma)

5. Use polynomial contrasts to test if there are linear or quadratic changes across time within each level of the category variable. [2 pt.]

Answer:

There is a linear trend across the standard and training groups. This is shown by the statistically significant linear contrast for time for these two groups.

Code/Syntax:

poly\_contrasts <- contr.poly(3)

contrasts(hw\_long$time) <- poly\_contrasts

split\_list <- list(

time = list("linear" = 1, "quadratic" = 2)

)

std\_trends <- aov(step\_count ~ time + Error(id/ (time)), data = hw\_long,

subset = (category == "standard"))

summary(std\_trends, split = split\_list)

trn\_trends <- aov(step\_count ~ time + Error(id/ (time)), data = hw\_long,

subset = (category == "training"))

summary(trn\_trends, split = split\_list)

6. Write a paragraph reporting your results from the repeated measures ANOVA and the polynomial contrasts as you would in a research paper, including all relevant detail from the tests. [2 pts.]

Answer:

We used repeated measures ANOVA to test whether the intervention would increase the step count in pregnant women across three weeks. There was a significant main effect of time *F*(2, 54) = 15.03, *p* < .001, but no significant main effect of category *F*(1, 27) = 0.05, *p* = 0.825. Further, there was a significant interaction between these two factors *F*(2, 54) = 4.63, *p* = .014.

We also found evidence of a linear change in step count across time points for those in the standard, *F*(1, 24) = 30.0, *p* < .001, and training, *F*(1, 30) = 4.58, *p* = .04, groups.