PSYC 5210 Exam 1

Due Sunday night

Instructions

This exam is focused on interpreting data using the methods we have learned so far in class. Each question should be answered as if you were writing a description for public presentation or a research article. Your main answers should not include example R code, but rather figures, tables, and text description of your findings. However, because it is easier for me to give partial credit for answers if I can see how you arrived at the solution, include an Appendix which includes the specific R code you used to solve each problem. *Do not turn in the direct output from a Markdown file. Turn in a report on your findings, and you may turn in the markdown or output from the markdown as an appendix.*

All questions concern a data set for human behavioral task called the 'match-to-sample' task.

Work by yourself on the exam. If you have clarification questions, you can ask via email, and responses will be forwarded to all students.

Overview

In this study, participants took part in a delayed match-to-sample task over multiple days and multiple times-of-day. This task involves looking at an pattern, and then a few seconds later being shown that pattern along wih another pattern not seen. They must decide which pattern was the one they actually saw. You might think about this as perhaps a user test that tracks a user over time to determine how they behave on some neurological testing, perhaps to understand how a drug or training intervention may help.

You can read the data in with the following command:

```
dat <- read.table("pooled.csv",header=T,sep=",",
        colClasses=c("factor","factor","numeric","numeric","numeric","numeric",
        "factor","factor","numeric","numeric","numeric","factor","numeric"))
```

```
subnum,group,session,tod,block,trial,targ,foil,studytime,rt,order,resp,corr
222,A,10,2,3,1,100011111110000010011010 1,110011011110000010011010 1,1006,654,0,<rshift>,1
222,A,10,2,3,2,011101100010011101100001 0,011101100010011101000001 1,1005,1635,0,<rshift>,1
222,A,10,2,3,3,001110001010011100111100 0,001110000010011110111100 0,1005,1333,0,<rshift>,1
222,A,10,2,3,4,100010100111110110000001 1,100010100111110010000011 1,1005,973,0,<rshift>,1
222,A,10,2,3,5,110110001101000000110111 0,101110001101000000110111 0,1006,1364,0,<rshift>,1
```

Here:

- targ and foil ar bitmaps indicating the patterns seen
- order refers to whether the correct stimulus was on the left or right (and thus whether the correct response was left-shift or right-shift).
- session refers to the day it was taken
- to refers to time of day; 1=morning, 2=late morning 3= afternoon 4=evening
- trial is the trial number within a block.
- ignore the 'block' variable, studytime, and group.
- Note that participants performed up to 12 sessions.
- the main dependent measures of interest are rt (response time, in milliseconds) and corr (accuracy, with correct responses coded as 1 and incorrect as 0.)

Problem 1

Use aggregate, tapply, or similar to determine the mean and sds for the following, and answer these questions. For each, show tables, and write text that both asks and answers the following questions. When you observe a difference, explain why you might have seen this and what it might mean.

(a) What was the effect of time of day on RT and accuracy?

(b) What was the effect of session on RT and accuracy?

(c) What was the effect of the correct response side (i.e., order) on RT and accuracy?

(d) Was there improvement in RT/accuracy within trial blocks? Answer this both overall, and for only block 1.

(e) Categorize the sessions into three session blocks: 1-4, 5-8, and 9-12. What was the effect of time of day and session block (i.e., a 4x3 table) on RT and accuracy?

Problem 2

For each part of 1, create a figure that shows the effect on response time. Across the different parts, use each of the following at least once: a barchart with error bars specifying either standard deviation or standard error, a matrix plot (matplot), a boxplot, and a plot where you overlay means on top of individual data points (similar to a bandplot). Each plot must use colors. For each plot, full credit will only be given for figures that are fully annotated and labeled. Write a figure caption for each, as it would appear in a paper in which you used each plot, indicating what is being shown (e.g., what are the error bars, the dv, etc.).

Problem 3

. Create a function that takes as its argument a session id and participant code, as well as the data, like this:

```
PlotCombo <- function(subnum,session,data) {...}
```

`PlotCombo` should:

(a) use the entire data set as an input for the data argument

(b) select the subset of data that match the given values of subnum and session.

(c) If an illegal combination is given (i.e, one such that the resulting data set is empty), it should exit with a warning, but create an empty plot that reads something like "no data available". Be sure to demonstrate and test that this works.

(d) Create a plot of mean response time by trial number for that session:

(e) Make titles or text that label the session, and subject code for the data set, as well as the time-of-day (tod).

Once the PlotCombo is complete, using either the `pdf()` command or an RMarkdown file, create a multi-page summary that shows each participant by iterating over at least the first 8 blocks of each participant, (e.g., plotting a 2x4 or 4x2 grid on each page.) If you call `pdf()`, it will automatically create new pages once each page is filled, and save the file when `dev.off()` is called. For example, something like:

```
pdf("filename.pdf")
PlotCombo(3,1,data)
PlotCombo(3,2,data)
PlotCombo(3,3,data)
dev.off()
```

Again, be sure describe what you have found–are there any patterns, outliers, or interesting results?

Some pointers

1. Be sure you do a good job at describing the results, labeling figures, interpreting what is going on and using the data to back it up. Show and Tell

2. Don't just turn in a dump of output. Treat this as a deliverable for a customer, or a prospective customer, or a boss. Think about things like significant digits, cleanliness of figures, use of irrelevant colors, etc. axis labels, and using common color themes. Don't use of default labels unless they are appropriate. Choose appropriate x and y limits, and make sure your figures are readable in the document you turn in.

3. Some of the challenges students have had previously on similar problems include:

    - using global data instead of data passed into function in part 3
    - doing a subset but not using the subset of data
    - trying to iterate when you wanted to just do a subset.
    - managing the syntax for subsetting.

4. Be sure to think about data in terms of information to communicate. Any table or figure you show should be part of your argument. Don't show a plot you can't understand, and won't explain to the reader.