Assessment 1

PS923 - Methods and Analysis in Behavioural Science

Henrik Singmann - Autumn Term 2019 (updated: 2019-11-13)

- **This assessment counts for 33% of your overall grade.**
- **Submission Instructions:** Submit your solution as **one** `html` or `pdf` document containing both `R` code, `R` output, figures, and written out text (i.e., full sentences) to Tabula as "Assessment 1" until **Wed, 27. November, 12:00 pm (noon)**.
- Please use `RMarkdown` to create the document.
- **Important:** Your document should be called `YOUR-STUDENT-ID_a21` (followed by the correct file extension). Please also add your student ID to the top of the document. To ensure anonymous marking, please refrain from using your name in either the document script or the file name.
- Your text does not need to contain references (i.e., references to scienitifc papers).

General Guidelines

Please complete the following questions. Your answer to each question should have two separate sections for each question, one immediately after the other.

In the first section, write out your answers using complete sentences, as you might for the results section of a paper. Include descriptive statistics in the text, or in tables or figures as appropriate. Tables and figures should be of publication quality (i.e., fully labelled, etc.). Integrate inferential statistics into your description of the results. Your answers might be short. **Given the correctness/appropriatness of the statistical analysis, the first section will play the main role for your mark.**

The second section should include the complete `R` code that you used and its output. Add comments (after a `#`) to explain what the code does. The code should show all of the commands that you used, enough for me to replicate exactly what you did (I will be copying and pasting code to run it, so make sure that works). You can include figures here that you used to explore the data that you do not wish to include in the first section. I will use the second section to help identify the source of any mistakes. For practical reports and papers you would only submit the first section, and thus the first section should stand alone without the second section.

Example Question

Does mere exposure to a stimulus improve its attractiveness? In an initial stage, participants were exposed to a series of pseudowords. Words were exposed at very short durations with a mask. (Pilot work established that participants were unable to report whether or not a word was presented before the mask in these conditions.) In a second phase, a mixture of the old, exposed pseudowords and new, previously unseen pseudowords were presented. Participants could view each word for as long as they liked before rating their liking for the word on a 1-10 scale. Using the data set `mere_exposure.csv`, test the hypothesis that mere exposure increases the attractiveness of pseudowords.

Example Answer

Section 1

Does exposure to a word improve attractiveness? To investigate this question, 32 participants took part in an experiment in which their main task was to rate the attractiveness of pseudowords on a 1-10 scale. Before the main task, participant were shown half of the pseudowords for a very short duration so that they could not perceive them consciously. Figure 1 shows the distribution and means of the attractiveness ratings and suggests that preexposed pseudowords (i.e., those shown briefly before the main task) were rated as more attractive than new, previously unseen pseudowords. We analysed the attractiveness ratings using an ANOVA with single repeated-measures factor exposure (old versus new). The difference in ratings (difference = 1.82, SE = 0.12) was significant, $F(1, 31) = 228.33$, $p < .0001$.
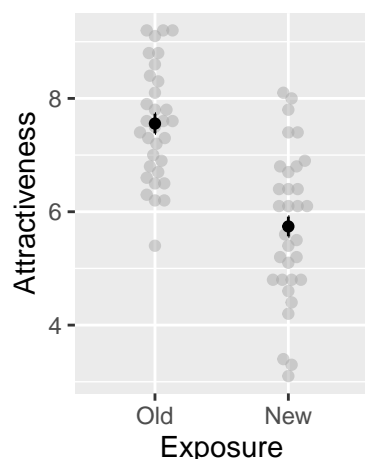

Figure 1. Attractiveness ratings of pseudowords as a function of prior exposure. Points in the background show the raw data (overlapping points are offset on the x-axis), black points in the foreground show the mean, error bars show 95% within-subjects confidence intervals.

Section 2

```
library("tidyverse")
library("afex")
library("emmeans")
mere_exposure <- read_csv("mere_exposure.csv")
glimpse(mere_exposure)
#> Observations: 32
#> Variables: 3
```

```
#> $ id         <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, ...
#> $ old_liking <dbl> 7.4, 6.5, 6.8, 7.8, 6.9, 9.2, 6.2, 7.9, 6.5, 9.1, 7...
#> $ new_liking <dbl> 5.6, 5.4, 3.4, 5.2, 6.1, 8.0, 4.4, 6.1, 3.1, 6.4, 6...

me_tidy <- mere_exposure %>%
  pivot_longer(cols = -id, names_to = "Exposure", values_to = "Attractiveness")
str(me_tidy)
#> Classes 'tbl_df', 'tbl' and 'data.frame':    64 obs. of  3 variables:
#>  $ id           : num  1 1 2 2 3 3 4 4 5 5 ...
#>  $ Exposure      : chr  "old_liking" "new_liking" "old_liking" "new_liking" ...
#>  $ Attractiveness: num  7.4 5.6 6.5 5.4 6.8 3.4 7.8 5.2 6.9 6.1 ...

me_tidy %>%
  group_by(Exposure) %>%
  summarise(mean = mean(Attractiveness),
            sd = sd(Attractiveness))
#> # A tibble: 2 x 3
#>   Exposure    mean    sd
#>   <chr>      <dbl> <dbl>
#> 1 new_liking  5.74  1.34
#> 2 old_liking  7.56  1.01

ggplot(me_tidy, aes(Attractiveness)) +
  geom_histogram() +
  facet_wrap(~Exposure)
```
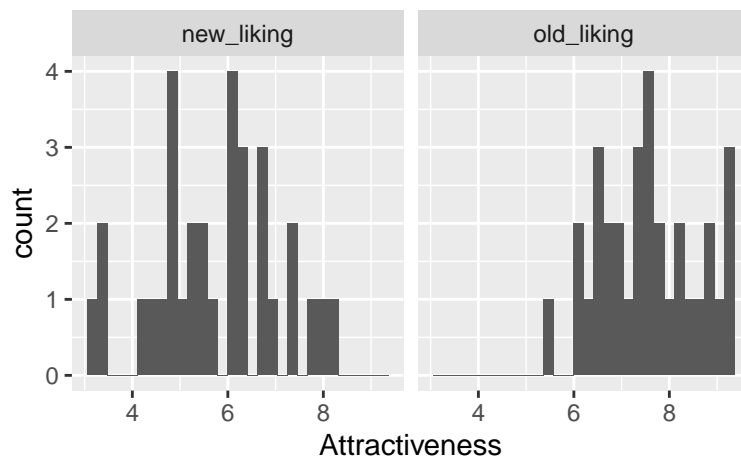


```
(a1 <- aov_ez(id = "id", dv = "Attractiveness", me_tidy, within = "Exposure"))
#> Anova Table (Type 3 tests)
#>
#> Response: Attractiveness
#>     Effect    df  MSE         F ges p.value
#> 1 Exposure 1, 31 0.23 228.33 *** .38  <.0001
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1 ' ' 1

pairs(emmeans(a1, "Exposure"))
#>  contrast              estimate   SE df t.ratio p.value
```

```
#>  old_liking - new_liking    1.82 0.12 31 15.110  <.0001
```

```
afex_plot(a1, "Exposure", error = "within",
          factor_levels = list(Exposure = c("Old", "New")),
          data_arg = list(cex = 3.5, color = "darkgrey"))
```

Task 1

One important research question in many applied fields, such as marketing, is which factors predict the subjective valuation of objects. For example, which value do we associate with a specific box of chocolates? Intuitively it makes sense to assume that the judgment made by a person of how much someone else would value an object should be a good predictor for how much the person themselves values an object. The goal of this task is to investigate this relationship: How good of a predictor is a value judgment participants do for other people in predicting their own value judgment for the same object?

`self_other_judgments.csv` contains parts of the data from a large study in which participants were asked to judge the value of a box of chocolates for themselves (variable `self`) as well as for others (variable `other`). Participants were randomly assigned to one of two groups (variable `task`). In the `WTP` group, participants had to judge how much they would be willing to pay for the box of chocoloates and how much they believed someone else would be willing to pay for the box of chocolate (in US$). In the `Enjoy` group, participants had to judge how much they would enjoy the box of chocoloates and how much they believe someone else would enjoy the box of chocoloates on a scale from 0 to 100. The data also contains a participant identifier (variable `pid`) as well as the information which judgment had to be performed first (`Order`).

Your task is to analyse the data using a linear regression model. The main research questions are (a) if other-judgments predict self-judgments and (b) whether they do so equally for both groups (i.e., level of `task` factor). Present the results as you would do in a paper in APA format. That means, describe your statistical model and results and describe which conclusions the results allow with regards to the research questions. Also, include only one figure (which may contain multiple panels) with appropriate figure caption in the first section for this task.

Task 2

An important determinant for individuals' preferences for certain products is a products' value or rank on a relevant dimension compared to similar products. For example, when choosing between different laptops, you might be comparing them in terms of their disk space; one laptop may have 0.5 TB disk space and another may have 1 TB. One research question in this literature is whether the **absolute** rank (e.g., the absolute amount of disk space) or the **relative** rank (e.g., how much disk space compared to other laptops) of the product is a better predictor of individuals' preferences. The former ignores the distribution of the relevant dimension whereas the latter ignores the absolute value on the relevant dimension.

For example, imagine a set of 10 laptops 8 of which have a disk space below 0.5 TB, one laptop has a disk space of 0.5 TB, and one laptop has a disk space of 1 TB. In this case, the laptops with 0.5 TB disk space and 1 TB disk space both have a high relative rank (relative ranks of 2 and 1, respectively). However, the laptop with 0.5 TB has a medium absolute rank whereas the laptop with 1 TB has a high absolute rank. By making the distribution of attribute values in a set less extreme, relative and absolute rank can be more aligned with each other. By making the distribution of attribute values in a set even more extreme, relative and absolute rank can be even more strongly pitted against each other.

`wtp_factorial.csv` contains data from a (simulated) factorial experiment which investigated this research question. Participants were presented with objects which either had a `low`, `medium`, or `high rank` and we manipulated the `type` of rank, either `relative` or `absolute`. In addition, we sampled participants from two different `population`s, either `general` population (excluding economists) and `economists`. Participants preferences were obtained via willingness-to-pay (`wtp`) elicitations (i.e., `wtp` was the dependent variable).

Your task is to analyse the data using an ANOVA (preferably using `afex`) and present the results as you would do in a paper in APA format. That means, describe your statistical model and results and describe which conclusions the results allow with regards to the research question. Be careful to not draw unwarranted causal conclusions. Also, include only one figure (which may contain multiple panels) with appropriate figure caption in the first section for this task.