Home exam

Tim Blauberger

PS2303 V23 Vetenskaplig metod III

André Hansla / Valgeir Thorvaldsson

March 23, 2023

Task 1 (a)

2p

The two-sided independent samples t-test for effect of stress on numeric performance is non-significant t(98)= -.016, p = .987. This means that stress level has no statistically significant effect on numeric performance. A strength of the t-test in this context is its simplicity. It isolates this effect and makes it easy to interpret. At the same time, this simplicity can also be considered a weakness. Other variables such as task difficulty and interactions with these variables are not taken into consideration.

Task 1 (b)

2p

I performed a one-way ANOVA to assess the effect of education on wordtest performance. An ANOVA was performed instead of a t-test, as education has three levels. The assumptions for an ANOVA were partially met, as group sizes differed, whereas Levene's test was nonsignificant with p = .684. The F-test was nonsignificant with F(2, 97) = 1.516 p = .225, η^2 = .030. As the F-test was nonsignificant, no post-hoc tests were performed. These results mean that education level has no statistically significant effect on performance in the word test. A strength of this analysis (ANOVA) is that it avoids problems of multiple comparisons. A weakness in this context is again the isolation of education level as the only included independent variable. Therefore, we cannot address interactions and influences of other variables.

Task 2

3,5p

I performed a two-way ANOVA (2 x 2 with stress level and task difficulty as between-subject factors) to assess the influence of these factors on motor test performance. Levene's test was insignificant with p = .237. The F-test for the main effect of task difficulty is significant with F(1, 96) = 13.339, p <.001, η_p^2 = .122. Participants in the easy task condition performed significantly better than those in the difficult task condition. The main effect of stress was slightly significant with F(1, 96) = 4.328, p <.04, η_p^2 = .043. Participants

in the stress condition performed significantly better than those in without stress induction. The interaction between difficulty and stress was nonsignificant with F(1, 96) = 2.372, p <.127, η_p^2 = .024. A strength of this analysis is its easy and clear interpretability (even the interaction would be easily interpretable in a 2 x 2 ANOVA) as well as decreased error rate to multiple t-tests. However, the isolation of two independent variables cannot provide a thorough understanding of the effects of other variables and possible confounders.

Task 3 a)

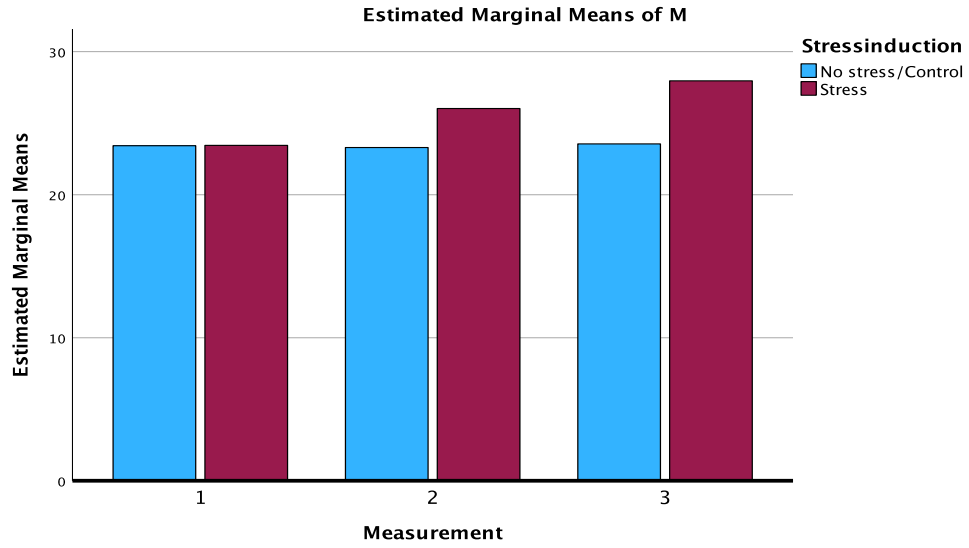
I am using a one-way repeated measures ANOVA for this task (3 repeated measurements of numeric performance; stress induction as a factor with two levels).

2p

Task 3 b)

Mauchly's test of sphericity is nonsignificant (p = .981). The main effects of both factors are significant, for stress induction with F(1, 98) = 6.992, p = .10, η_p^2 = .067 and for measurement with F(1, 98) = 10.429, p = .002, η_p^2 = .096. Also, the interaction between both factors is significant with F(1, 98) = 9.298, p = .003, η_p^2 = .087. Looking at the plot in figure 1 we can see that the point of measurement does not really affect numerical performance in the condition without stress induction. However, participants in the stress induction condition performed better with each measurement.

Figure 1



Task 4 a)

For this task I chose to perform a one-way ANOVA and a Kruskal-Wallis test as a nonparametric alternative. Both allow for comparing more than two groups, which is necessary as we have three categories of gender. The results differ, as the F-test of the ANOVA is significant with $F(2, 201) = 3.701$, $p = .036$, while the Kruskal-Wallis test is not significant with a test statistic of 1.210, $df = 2$, $p = .546$. However, Bonferroni post-hoc tests for the ANOVA revealed no statistically significant group differences either. The results might differ because the data violates the important assumptions of the parametric ANOVA, that groups should have equal sample sizes. As the group sizes differ strongly with only 8 participants identifying as nonbinary, it would be better to perform the nonparametric test (if any test at all, given the low group size of nonbinary individuals). 3p

Task 4 b)

To test this hypothesis, I performed a Chi-squared test of gender (excluding nonbinary) and the "follow-up" variable. The results are nonsignificant $X^2 = .753$, $df = 1$, $p = .391$, indicating that there is no difference between genders. 0,5p

Task 5

For this task I performed a Bayesian Bonferroni proportion test. I used the data from the prior experiment to form a prior, with a = 5 and b = 15. Using the data of the "new" experiment as the likelihood, I computed the posterior distribution and Bayes Factors in Jamovi. The results clearly prefer the null-hypothesis with a BF = 13.136, which can be considered as strong evidence. One could imagine that the significant result in the old experiment was caused by the extremely small sample size, leading to a high standard error. The combined evidence of the Bayesian approach can be considered more meaningful in this context and should be preferred. This means that the test supports the hypothesis of our animal expert, that there is no sex-difference of owners in dog attachment. 3p

PS2303 Methods III
Individual Home Exam
(Individual assignment 2)
Spring 23

Welcome to the individual home exam!

The individual home exam consists of 5 tasks (4 points per task) covering the different parts of the course.

For this exam there are no restrictions concerning which aid to use. However, you are not allowed to copy any portion of any text or in any other way plagiarize. If we judge that a text or portion of a text has been plagiarized this may have disciplinary consequences.

We expect you to do the exam on your own.

Good Luck!

Maximal score: 20 points.
Pass with distinction (75%): 15 points
Pass (60%): 12 points

Tasks 1, 2 & 3

Download the datafile "Methods III Fictive Data for Home Exam".

It contains simulated/fictive data from an imagined experiment on stress and performance. Participants were one hundred citizens from Gothenburg who signed up to participate in a performance experiment, which they were informed about via a newspaper announcement. They were compensated with one cinema ticket.

Participants were randomly allocated to four independent groups, each group representing a combination of the following two factors: Factor A: Stress induction (1) versus No stress induction (0); Factor B: Difficult (1) versus Easy Task (0).

Each participant then repeated a numeric performance test ("huvudräkning" ("mental arithmetic")) three times with one-minute paus in between where right before the onset of the next test round the stress or control procedure was repeated. After the numeric test, participants completed two other tests measuring word and motoric performance but these they took only once. For each test, the possible range of performance is 0 (No solution correct) to 50 (All solutions correct).

Participants background in terms of gender, age, and education were also recorded in the experiment and are defined in the data file.

Task 1 *(4 points)*
(a) T-test if there is an effect of stress on numeric performance (for the first trial). Report and interpret the main results. What conclusions do you draw? Also reflect upon one strength and one limitation with the analysis applicable to the context (2p). Word limit: 200 words
(b) Test potential mean differences in word performance related to education level. Report and interpret the main results. What conclusions do you draw? Also reflect upon one strength and one limitation with the analysis applicable to the context (2p). Word limit: 200 words

Task 2 *(4 points)*
Test possible main and interaction effects of stress and task difficulty on motoric performance. Report and interpret the main results. What conclusions do you draw? Also reflect upon one strength and one limitation with the analysis applicable to the context. Word limit: 200 words

Task 3 *(4 points)*
Test if there is an effect of stress induction on numeric performance (using all three measurements).
(a) What type of ANOVA (design) are you using for this analysis? (1p)
(b) Report and interpret the main results. Present a suitable graph illustrating the main results. What conclusions do you draw? (3p). Word limit: 200 words

Task 4 (4 points)

Download the datafile ”dating_data”.

204 people took part in a speed dating event. Each person briefly met 20 other people and engaged in a conversation with them. Afterward, for each person they met, they indicated whether they would be interested in meeting again. After being informed about the people with whom they “matched” (i.e., each person in the pair indicated they would be interested in meeting again), they had an opportunity to send a message to the people with whom they matched.

The data set contains the following variables:

gender: the participant’s self-reported gender
age: the participant’s self-reported age
likes_received: how many people indicated they would like to meet this participant again
matches: how many matches this participant received
followed_up: whether the participant sent a message to a match (0 = no, 1 = yes)

You have two tasks:

- Test whether there were gender differences in the number of “likes” a person received (i.e., how many people indicated they would like to meet the person again). Note that there are several ways to approach this research question. Briefly explain (in 1-3 sentences) why you have taken the approach you have selected. Use both parametric statistical tests and rank-based statistical tests to address this question, and briefly describe (in 2-4 sentences) the similarities/differences in the results. If the results are different, explain why they might be different. (3 points)

- Test the hypothesis that men are more likely than women to send a message to a match. Conduct a statistical test of this hypothesis and interpret the results (in 1-3 sentences). (1 point)

Task 5 (4 points)

Download the data file: "DogAttachmentExp2.sav"

Bayesian inferences
An animal expert reads about a study claiming that male dogs form a closer attachment with male owners (in comparison with female owners). The conclusion is derived from a study of 20 heterosexual couples (living together without children) that had in the beginning of the study just purchased a male dog. After a one year of dog ownership several valid tests of the dog attachment were conducted. The results section report that 15 of the dogs had formed a closer attachment with their male owner, in comparison to only 5 with the female owner ($p = .04$). Given our expert cannot really think of any obvious justification for this finding, and the small sample size, he is a bit skeptical and decides to replicate/redo the study. The data that he derives is on Canvas (DogAttachmentExp2.sav). Can you help our animal expert to combine, and interpret, the evidence from these two studies using Bayes´ rule (e.g., in JASP)? Do male dogs really prefer a specific sex? Provide a detailed evaluation of the combined evidence.