

Advanced Statistics Practice Test

Model solution

Bernhard Angele

Part 1 – Data analysis (50 %)

Scenario

The university wants to find out whether therapy animals are useful in lowering student stress ahead of exams. In order to test this, they asked 24 students to have four different exam revision sessions: one with no therapy animals present, one with a therapy cat, and one with a therapy dog. The students were then asked to report their stress levels on a scale from 0 to 100, where 0 is no stress at all and 100 is extreme stress.

Assignment

Conduct the appropriate ANOVA and the corresponding post-hoc t -tests comparing all three factor levels to each other (assume all assumptions are met). In order to help you with this, a number of intermediate values have already been computed for you below. Report the results of the ANOVA and t -tests as one would in the results section of an academic manuscript (using APA style). Finally, in layperson (non-academic) language describe the results and summarise what you can conclude about the effect of therapy dogs and cats. There is no word limit, but the whole report should be rather brief. You may use a calculator, Microsoft Excel, or any other computer programme, as well as your pet notes and any other offline resources you wish to bring to the test. Your test must be your own work and you may not collaborate with anyone else. Show your work so you can get partial credit in case you make small calculation mistakes.

Data

In order to save you time, part of the ANOVA table is provided below, as well as the group means and standard deviations and the means and standard deviations of the differences between factor levels for each student [Note: nope, not in the practice assignment].

```
## Warning: Converting "subject" to factor for ANOVA.
```

Means table

Condition	Mean	SD	N
Cat	47.0	4.42	24
Dog	46.0	5.10	24
No animal	57.6	7.08	24

Differences between groups

$Condition_1$	$Condition_2$	$Mean(Condition_1 - Condition_2)$	$SD(Condition_1 - Condition_2)$
Cat	No animal	-10.6	9.3
Dog	No animal	-11.5	8.82
Cat	Dog	0.958	7.15

ANOVA Table

Source	SS	df	MS	F	p
Between Subjects	549.875	23	23.908		
Within Subjects	3620	48	75.417		
–Animal condition	1969.083	2	984.542	27.433	$p < .01$
–Residual	1650.917	46	35.889		
Total	4169.875	71			

Part 2: Questions (50 %)

Please answer the following questions. Your answers should not be longer than a few sentences or a paragraph and should be comprehensible to a layperson who does not know anything about hypothesis testing but is familiar with general concepts such as sample, population, probability, probability distribution, degrees of freedom, mean, standard deviation, effect, predictor, variance, continuous and discrete variables, correlation, etc. If it helps, feel free to use concrete examples to aid your explanation.

Example question (*from last year's take-home test – questions in the in-class test will be a little shorter and less detailed*): In a non-technical manner, explain the concept of “Type I”, “Type II”, and “Type III” sums of squares in multiway ANOVA and multiple regression. Under which circumstances will the different sums of squares tests yield different results?

Answer: All sums of squares and the resulting F-values are model comparisons. They differ with regard to which models are compared. Type I sums of squares for a factor compare a model with all previously entered factors with a model containing all previously entered factors and the factor in question. They are order-sensitive. Type II sums of squares for a factor compare a model with all other factors except the factor in question *and its interactions* to a model with all other factors plus the factor in question, but none of its interactions. Type III sums of squares for a factor compare a model with all other factors except the factor in question (but including its interactions) to a model with all other factors (including interactions with the factor in question) plus the factor in question itself. The different sums of squares will yield different results if there is multicollinearity, i.e. if any of the factors are correlated. In an ANOVA design, multicollinearity can occur if the design is unbalanced (i.e. group sizes are not even).

Questions you should be able to answer:

1. Why do we do hypothesis tests?
2. What is the difference between a sample and a population?
3. What is a random variable?
4. Why do we care about theoretical distributions (e.g. the normal distribution)?
5. When can we assume that a random variable is normally distributed?
6. What is a z -value, and how is it related to the normal distribution?
7. What is a χ^2 -value, and how is it related to the normal distribution?
8. What is a t -value, and how is it related to the normal distribution?
9. What is an F -value, and how is it related to the normal distribution?

10. What is null hypothesis significance testing?
11. What is a p -value?
12. What does a p -value tell us about the sample?
13. What does a p -value tell us about the population?
14. What does it mean if an effect is “significant”?
15. What is a Type I error (and what is the α -rate)?
16. What is a Type II error (and what is the β -rate)?
17. What is the relationship between α -rate and β -rate?
18. What is statistical power?
19. How can we increase statistical power?
20. What is an unbiased estimator?
21. What are the assumptions we are making for the z -test?
22. What are the assumptions we are making for the t -test?
23. What are the assumptions we are making for the *ANOVA* (between subjects)?
24. What are the assumptions we are making for the repeated measured *ANOVA*?
25. Why is it a problem when assumptions aren’t met? What can happen?
26. What is a “robust” statistical test?
27. Which practices can lead to an inflated α -rate?
28. Why is an inflated α -rate a problem?
29. What is a one-tailed/two-tailed test?
30. How can the experimenter’s intentions determine the outcome of a null-hypothesis significance test?
31. Why is this dependence on experimenters’ intentions a problem?*
32. What is a stopping rule?
33. Why shouldn’t you use stopping rules?
34. Why is everybody using them anyway?*
35. What are “experimenter degrees of freedom”?*

- These questions are intended to make you think and the answers may be highly subjective depending on who you ask. I will obviously not ask questions phrased like that on the test!