Florian Wickelmaier
Forschungsmethoden und Mathematische Psychologie
Fachbereich Psychologie
Universität Tübingen

Statistik III/Advanced Statistics
SS 2026

Exercises

Exercise 21. Hedeker and Gibbons (2006) report on a study designed to determine the efficacy of a school-based smoking prevention curriculum in conjunction with a television-based prevention program. The study was based on a (2×2)-factorial design with the factors prevention curriculum (yes = 1, no = 0) and TV prevention program (yes = 1, no = 0). Students were randomly assigned to one of the four groups at the school level. Students' knowledge of tobacco-related health risks was assessed by administering the tobacco and health knowledge scale (THKS) before and after the intervention.

a) Set up the general model equation of an analysis of covariance (i. e., the adjusted follow-up values) for this study, using the labels y_{i1} and y_{i2} for individual i's THKS scores before and after the intervention. In addition, let $c_i = 1$ if person i completes the intervention with prevention curriculum ($c_i = 0$ otherwise), and let $tv_i = 1$ if intervention is completed with TV prevention program ($tv_i = 0$ otherwise).

b) Interpret the parameters included in the model equation.

c) Read the data contained in the file `television.txt` into R. (Note: Skip the data description at the beginning of the file in lines 1-42 by specifying the argument `skip = 42` in the function `read.table()`).

d) Calculate means and standard deviations for baseline, follow-up, and change scores for each of the four groups separately. How many students are in each group?

e) Plot the follow-up means as a function of the presence of the curriculum separately for the two TV conditions (`lattice` package allowed). What effects do you suppose?

f) Perform an analysis of the follow-up scores (ANOVA), change scores, adjusted follow-up scores (ANCOVA), and adjusted change scores.

g) Do the factors of the latter two models have identical effects on the change in knowledge?

h) Answer the following questions based on the parameter estimates in the ANCOVA model:

   (a) Do the two factors interact?

(b) What is the mean THKS score expected as a follow-up value for the reference group ($c_i = 0$, $tv_i = 0$) at a baseline value of $y_{i1} = 0$?

(c) How large is the predicted change in knowledge in the remaining groups relative to the reference group?

(d) Determine the follow-up value predicted for the reference group and the group with $c_i = 1$ and $tv_i = 1$ at a baseline value of $y_{i1} = 2$ (median)?

i) Check the ANCOVA model's assumption of identical slopes in the regression of the follow-up value on the baseline value in all groups with an appropriate statistical test ($\alpha = 0.05$).