Florian Wickelmaier
Forschungsmethoden und Mathematische Psychologie
Fachbereich Psychologie
Universität Tübingen

Statistik III/Advanced Statistics
SS 2026

Exercises

Exercise 19. In the context of a methadone program, Anglin, McGlothin & Speckart (1981) examined a sample of 647 heroine addicts. The following variables were collected:

C	County of residence: urban, suburban, rural
S	Socioeconomic status: low, not low
P	Parents available until age 16: yes, no

The response variable F indicates whether or not the first heroine use was before the 18th birthday. The following frequencies were observed:

		C	urban		suburb		rural	
		F	< 18	≥ 18	< 18	≥ 18	< 18	≥ 18
P	S							
yes	low		32	65	21	20	10	40
	not		34	97	14	28	8	23
no	low		41	42	18	22	6	13
	not		41	45	3	13	1	10

Wickens (1989) models these data using a logit model including the interaction of variables C and P in addition to all main effects. Parameter estimation yields the following results under dummy coding (last category as reference):

	Estimate	Std. Error	z value	$P(>\|z\|)$
(Intercept)	-1.416	0.448	-3.161	0.002
Slow	0.346	0.172	2.014	0.044
Curban	1.187	0.461	2.577	0.010
Csuburb	0.655	0.514	1.275	0.202
Pyes	-0.058	0.509	-0.114	0.909
Curban:Pyes	-0.764	0.552	-1.384	0.166
Csuburb:Pyes	0.330	0.622	0.531	0.595
Residual deviance: 6.2088 on 5 degrees of freedom				

Calculate by hand and using R.

a) Specify the underlying design matrix $\mathbf{X}$. Arrange columns according to the sequence of parameter estimates in the above table.

b) Determine the data vector $\mathbf{x}$ of a heroine addict who grew up without parents in the suburbs, and has low socioeconomic status.

c) What are the odds and the probability predicted by the model for this person?

d) Compute the odds for a person who, all other things being equal (ceteris paribus), does *not* have low socioeconomic status. What is the odds ratio of the previously considered person compared to this person?

e) How does the result of the preceding computation relate to the estimated value of the parameter associated with *S low* in the table?

f) Interpret the effect of the availability of the parents until age 16 on the time of first heroine use according to the model.

g) Test the lack of fit of the model ($\alpha = 0.05$, $\chi^2_{0.95}(5) = 11.07$).

Exercise 20. The file `KIMS_codebook.txt` lists the items of the *Kentucky Inventory of Mindful Skills* (KIMS; Baer, Smith, & Allen, 2004), and the file `KIMS_data.csv` contains the publicly available (`https://openpsychometrics.org/_rawdata/`) data that were collected using an online version of the KIMS (sample size $n = 601$). The questionnaire collects responses to 39 items (coded 1–5) in four scales labeled "Observing," "Describing," "Act With Awareness," and "Accept Without Judgment". Read the data into R:

```
dat <- read.table("KIMS_data.csv", header = TRUE, sep = ",", na.strings = 0)
dat <- na.omit(dat[, 1:39])
```

Then `dat` contains the data matrix for the following analyses.

a) Create a scree plot and add an illustration of the results of the analysis for random data.

b) Decide how many components to extract according to the Kaiser-Guttman criterion and according to parallel analysis. Let the further analyses be based on the components extracted according to parallel analysis.

c) Compute the factor loadings resulting from a varimax rotation and assign items to components based on the maximum absolute value.

d) Compare this assignment to the "Observing", "Describing", "Act With Awareness", and "Accept Without Judgment" scales from Baer et al. (2004) and try to interpret any discrepancies.

e) Compute the factor scores after varimax rotation for each individual. Note: Variable `scores` of function `princomp()` does not return the unrotated factor scores $\mathbf{F}$, but $\mathbf{Y}$ in the equation $\mathbf{F} = \mathbf{Y}\mathbf{\Lambda}^{-1/2}$. The rotation matrix $\mathbf{M}$ is returned by `varimax()` in variable `rotmat`. (For more details and examples see the help pages `?princomp` and `?varimax`.)

References: Baer, R.A., Smith, G.T., & Allen, K.B. (2004). Assessment of mindfulness by self-report: The Kentucky Inventory of Mindfulness Skills. Assessment, 11, 191-206. `https://doi.org/10.1177/1073191104268029`