Florian Wickelmaier
Forschungsmethoden und
Mathematische Psychologie

Fachbereich Psychologie
Universität Tübingen

Hausaufgaben
Generalized linear models

Erzeugen Sie ein lauffähiges und kommentiertes R-Skript (Dateiendung `.R`), in dem Sie die folgenden Aufgaben bearbeiten.

a) (5 Pkt.) Es soll überprüft werden, wie sich drei verschiedene Arten des Kreativitätstrainings (Faktor A) auf die Kreativität von Versuchspersonen auswirken. Drei Zufallsstichproben à 5 Versuchspersonen werden vor Beginn des Trainings (b_1), während des Trainings (b_2) und nach Abschluss des Trainings (b_3) hinsichtlich ihrer Kreativität getestet, wobei jede Stichprobe ein anderes Kreativitätstraining erhält. Die folgende Tabelle zeigt die Testwerte des Experiments (vgl. Bortz, 1999, S. 327). Niedrige Werte bedeuten mehr Kreativität:

id	A	B1	B2	B3
1	a1	54	52	50
2	a1	54	48	50
3	a1	52	44	43
4	a1	57	53	50
5	a1	55	49	47
6	a2	54	51	50
7	a2	59	50	50
8	a2	54	48	44
9	a2	50	45	45
10	a2	56	54	49
11	a3	51	45	42
12	a3	54	46	49
13	a3	55	49	46
14	a3	56	49	49
15	a3	52	46	48

(a) Erzeugen Sie einen "breiten" Dataframe, der die angegebenen Variablen enthält. Verwenden Sie dann die Funktion `reshape()`, um den Dataframe in das "lange" Format zu transformieren.

(b) Stellen Sie die Mittelwerte für die einzelnen Stichproben in einem Interaktionsplot graphisch dar. Zeichnen Sie die Standardfehler für jeden Mittelwert ein. Vergessen Sie nicht eine sinnvolle Beschriftung. Exportieren Sie die Abbildung als Pdf-Datei.

(c) Führen Sie eine Varianzanalyse mit der Funktion `aov()` durch ($\alpha = 0.01$). Interpretieren Sie die Ergebnisse.

b) (7 Pkt.) Erzeugen Sie zwei Vektoren `x` und `y` mit jeweils 100 Elementen mit: $X \sim N(1, 1.4^2)$ und $Y \sim N(5, 1.4^2)$. Erzeugen Sie einen Dataframe mit den Variablen `id`, `group` und `score`, wobei `x` und `y` den Score für die beiden Gruppen bilden:

```
'data.frame':        200 obs. of  3 variables:
 $ id   : int  1 2 3 4 5 6 7 8 9 10 ...
 $ group: Factor w/ 2 levels "x","y": 1 1 1 1 1 1 1 1 1 1 ...
 $ score: num  0.6699 -0.0338 0.1212 0.9549 1.3541 ...
```

(a) Testen Sie die Hypothese "kein Gruppenunterschied" auf drei Arten: Mit (1) einem t-Test unter der Annahme $\sigma_X^2 = \sigma_Y^2$, (2) einer Varianzanalyse mit der Funktion `aov()` und (3) einer Regression (Zielgröße: `score`, Einflussgröße: `group`).

(b) Lassen Sie sich für den t-Test die Schätzungen der Mittelwerte von X und Y (`estimate`) ausgeben und für die Varianzanalyse und die Regression jeweils die Parameterschätzungen (`coefficients`). Vergleichen Sie außerdem die p-Werte der drei Methoden sowie die quadrierte t-Statistik mit den F-Statistiken der beiden anderen Methoden. Interpretieren Sie diese Ergebnisse.

(c) Verwenden Sie die Funktion `model.matrix()`, um die Designmatrizen der beiden geschätzten Modelle (Varianzanalyse und Regression) auszugeben. Stellen Sie die Kontraste auf Effektkodierung und schätzen Sie die Modelle erneut. Wie verändern sich unter Effektkodierung: die Ergebnisse aus (b), die Designmatrizen, die Interpretation der Parameterschätzungen?

Literatur

Bortz, J. (1999). *Statistik für Sozialwissenschaftler (5. Auflage).* Berlin: Springer.