

STATISTIK 2 (107.325) WS 2009
COMPUTERSTATISTIK (107.258) WS 2009

Übung 8

14. Dezember 2009
Dutter

29. Analysieren Sie die Daten des Vienna City Marathons in den letzten 10 Jahren 2000 bis 2009 mit dem Computer Program System .

Lesen Sie alle 10 Datensätze ein und merken Sie sich die Klasseneinteilung, die Endzeiten und wo möglich (nicht vom Jahre 2004!) die (bei mehrfachen) höheren Zwischenzeiten.

29a:

Verwenden Sie ähnlich wie im Beispiel 23 nur die Altersklassen H, 30, 35, 40, 45 und 50.

Testen Sie Hypothesen im linearen Modell „Endzeiten abhängig von Altersklasse und Geschlecht (ohne Wechselwirkung)“ und kommentieren Sie das Ergebnis.

Der Computeroutput sollte etwa so ausschauen:

```
> anova(aov(endZeiten ~ agekl + sexkl, data = zeiten))
```

Analysis of Variance Table

Response: endZeiten

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
agekl	5	122.5	24.5	70.904	< 2.2e-16 ***
sexkl	1	1310.8	1310.8	3792.375	< 2.2e-16 ***
Residuals	63086	21805.4	0.3		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

- 29b:** Versuchen Sie eine kleinere Stichprobe, indem Sie z.B. aus den etwa 63.000 betrachteten Laufzeiten zufällig 1000 auswählen und die Varianzanalyse wiederholen.

Kommentieren Sie die Unterschiede.

Der Computeroutput könnte etwa so ausschauen:

```
> zzeiten <- zeiten[sample(1:nrow(zeiten), 1000), ]
> anova(aov(endZeiten ~ agekl + sexkl, data = zzeiten))
```

Analysis of Variance Table

Response: endZeiten

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
agekl	5	2.68	0.54	1.4522	0.203
sexkl	1	37.24	37.24	100.8620	<2e-16 ***
Residuals	993	366.63	0.37		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

30. Machen Sie eine Klasseneinteilung der Zwischenzeiten ZZ:

	ZZ	≤	1.5
1.5	<	ZZ	≤ 2
2	<	ZZ	≤ 2.5
2.5	<	ZZ	≤ 3
3	<	ZZ	

Schliessen Sie dabei die Nullen der Zwischenzeiten aus, d.h. der kleinste Wert sollte etwa .9 Stunden sein.

30a: Modellieren Sie linear die Endzeiten in Abhängigkeit von den Zwischenzeitklassen und Altersklassen (noch immer ohne Wechselwirkungen). Gibt es Abhängigkeiten?

30b: Bitte um ähnliche Analyse bezüglich Abhängigkeiten von Zwischenzeitklassen und Geschlecht.

30c: Modellieren Sie linear die Endzeiten in Abhängigkeit der Zwischenzeitklassen, Geschlecht und Altersklassen (noch immer ohne Wechselwirkungen). Gibt es Abhängigkeiten?

31. Untersuchen Sie die Abhängigkeiten wie in den Beispielen 29 und 30, aber unter Einbezug von möglichen Wechselwirkungen.

32. Man betrachte die Studie über die Wirksamkeit von Schmerztabletten (ähnlich dem Beispiel im Skriptum): Insgesamt testen 20 Personen 4 verschiedene Schmerzmittel. Die Daten sind in Tabelle 1 dargestellt. Die Personen „sind in den Medikamenten geschachtelt“, weil jede Person nur ein Medikament testen kann. Man prüfe die unterschiedliche Wirkung der Medikamente mit Hilfe des „geschachtelten Versuchsplanes“. Die Daten stehen unter dem Namen „schmerz.txt“ auf der Vorlesunghomepage <http://www.statistik.tuwien.ac.at/public/dutt/vorles/>.

Bitte senden Sie die Ausarbeitung in Form eines pdf-Files (nicht mehr als 3 Seiten) mit den Resultaten (Outputs plus textliche Kommentare) und Listing des

		P1	P2	P3	P4	P5
M1	1	8.1	10.6	7.5	14.8	11.0
	2	7.5	11.5	12.2	11.7	9.1
	3	14.5	11.2	8.1	3.4	13.4
M2	1	10.9	11.0	13.8	13.5	12.8
	2	13.8	13.3	11.2	5.0	12.9
	3	10.8	10.5	6.6	9.6	12.3
M3	1	16.1	11.7	13.2	11.8	7.9
	2	10.8	10.8	11.8	15.3	14.3
	3	11.5	11.2	14.1	13.7	9.9
M4	1	10.9	14.1	15.3	12.7	15.6
	2	14.2	11.2	14.0	9.6	17.3
	3	18.9	11.9	9.9	14.7	12.6

Tabelle 1: Wirkungsdauer von Schmerztabletten

Programmkodes (Funktion) an

`R.Dutter@tuwien.ac.at`

bis zum Freitag-Abend, 8 Jänner 2010.

Bitte den Namen des pdf-File folgendermaßen:

name_exer_8.pdf

wobei ‘name’ für den Familiennamen steht.

Empfehlenswert ist es, den Bericht mit „Sweave“ zu erstellen. (Vorlage siehe <http://www.statistik.tuwien.ac.at/public/dutt/vorles/> .)