

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

Übung 6

30. November 2009

Dutter

21. Betrachten Sie 4 Altersgruppen der Patientendaten "Werner Blood Chemistry Data": (18,25], (25,32], (32,42], (42,55]. Die Anzahl pro Gruppe sollte sein: 52, 48, 42, 46.

Testen Sie auf Einfluss des Alters (der Altersgruppen) auf Cholesterinwerte. Erklären Sie die Voraussetzungen zu diesem Test. Kommentieren Sie das Resultat.

22. Verwenden Sie von den Altersgruppen der Patientendaten nur die ersten 3 bzw. die mittleren 2. Wie ändern sich die Signifikanzen? Diskutieren Sie Ihr Resultat.

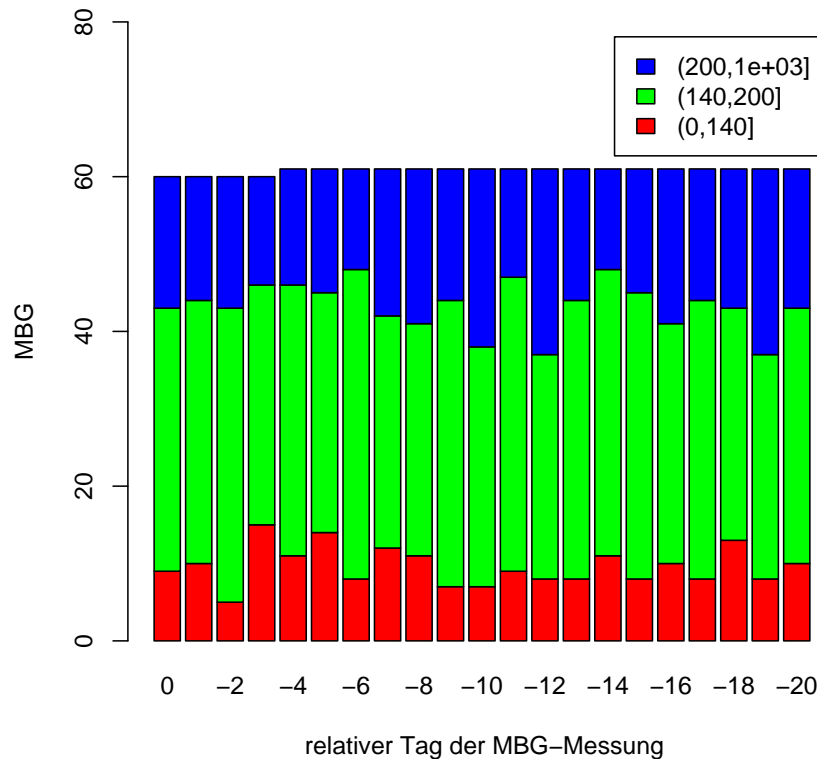
23. Analysieren Sie die Daten des Vienna City Marathons 2009 mit dem Computer Program System .

Empfehlung: Laden Sie die (korrigierten) Daten nochmals aus dem Internet und exekutieren:

```
> f.mara_read <- function() {
+   mara09 <- read.table("mara_t_09", sep = ":")
+   colnames(mara09) <- c("Kl-Sex", "1.Zw.Zeit_h", "1.Zw.Zeit_m",
+     "1.Zw.Zeit_s", "2.Zw.Zeit_h", "2.Zw.Zeit_m", "2.Zw.Zeit_s",
+     "Endzeit_h", "Endzeit_m", "Endzeit_s")
+   mara09[, 1] <- as.character(mara09[, 1])
+   dmar09 <- mara09[, 8] + mara09[, 9]/60 + mara09[, 10]/3600
+   dmar09 <- cbind(mara09[, 2] + mara09[, 3]/60 + mara09[,
+     4]/3600, dmar09)
+   kmar09 <- mara09[, 1]
+   dimnames(dmar09) <- list(kmar09, c("1. Zwischenz.", "Endzeit"))
+   datasex09 <- substr(kmar09, 1, 1)
+   datakl09 <- substr(kmar09, 3, 5)
+ }
> f.mara_read()
```

Man betrachte nur die Altersklassen H, 30, 35, 40, 45, 50, 55, 60, 65 und 70. Haben diese Klassen einen signifikanten Einfluss auf die Laufzeiten der Männer? Das Gleiche bei den Frauen?

24. Man betrachte Diabetes-Daten: Werte von HbA_{1c} (Hämoglobin A1c) und den entsprechenden MBG-Wert (mittlere Blutglukose des Tages) in $[mg/dl]$ (am gleichen Tag, Tag davor, 2 Tage davor, ..., 20 Tage davor). Ein Balkendiagramm der MBG-Werte über die Tage (relativ zum Messtag von HbA_{1c}) wird in der naechst Grafik dargestellt. Die Werte selbst sind aber in 3 Klassen geteilt: niedrig (unter 140), normal (140 – 200) und hoch (höher als 200).



Man untersuche jeweils den Einfluss von MBG am gleichen Tag und an den betrachteten Tagen vor der Messung von HbA_{1c} auf den Wert von HbA_{1c} mittels einfacher Varianzanalyse. D.h. man mache eine einfache Varianzanalyse mit den MBG-Werten als Einfluss-Faktor am Tag j vor der Messung und j läuft von 0 bis 20. Man fasse die wesentlichen Ergebniss übersichtlich in Tabellenform zusammen.

Vorschlag:

```
> hh <- anova(aov(HbA1C$HbA1C ~ d[, 1]))[1, ]
> for (j in 2:21) {
+   h <- anova(aov(HbA1C$HbA1C ~ d[, j]))[1, ]
+   hh <- rbind(hh, h)
+ }
```

```
> attributes(hh)$row.names = dimnames(d)[[2]]
> print(hh)
```

Analysis of Variance Table

Response: HbA1C\$HbA1C

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
MBG	2	11.0209	5.5104	5.3489	0.007431	**
MBG...1.T	2	13.8529	6.9264	7.0641	0.001816	**
MBG...2.T	2	6.3640	3.1820	2.8618	0.065412	.
MBG...3.T	2	5.2709	2.6354	2.3300	0.106493	
MBG...4.T	2	2.8824	1.4412	1.2483	0.294584	
MBG...5.T	2	1.4734	0.7367	0.6249	0.538859	
MBG...6.T	2	1.8894	0.9447	0.8063	0.451438	
MBG...7.T	2	1.0001	0.5000	0.4213	0.658201	
MBG...8.T	2	2.4370	1.2185	1.0485	0.357022	
MBG...9.T	2	5.2009	2.6004	2.3332	0.106022	
MBG...10.T	2	12.0415	6.0207	6.0413	0.004137	**
MBG...11.T	2	3.0461	1.5231	1.3225	0.274390	
MBG...12.T	2	7.4284	3.7142	3.4514	0.038349	*
MBG...13.T	2	0.7510	0.3755	0.3152	0.730887	
MBG...14.T	2	8.5775	4.2887	4.0601	0.022373	*
MBG...15.T	2	0.8092	0.4046	0.3399	0.713220	
MBG...16.T	2	7.3073	3.6536	3.3886	0.040567	*
MBG...17.T	2	2.7186	1.3593	1.1745	0.316203	
MBG...18.T	2	0.8822	0.4411	0.3710	0.691664	
MBG...19.T	2	7.7091	3.8545	3.5981	0.033650	*
MBG...20.T	2	11.4327	5.7164	5.6762	0.005606	**

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

Bitte kommentieren Sie das Ergebnis.

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 Programmkodes (Funktion) an

R.Dutter@tuwien.ac.at

bis zum Samstag-Abend, 5. Dezember 2009.

Bitte den Namen des pdf-File folgendermaßen:

name_exer_6.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/> .)