

**STATISTIK 2 (107.325) WS 2011**  
**COMPUTERSTATISTIK (107.258) WS 2011**

**Übung 5**

**23. November 2011**  
**Dutter**

Analysieren Sie die Daten des Vienna City Marathons 2011 mit dem Computer Program System  $\text{\textcircled{R}}$ . Sie stehen (anonymisiert) auf <http://www.statistik.tuwien.ac.at/public/dutt/vorles/> (ganz unten!).

```
> f.mara_read <- function() {
+   mara11 <- read.table("http://www.statistik.tuwien.ac.at/public/dutt/vorles/data/mara_
+     sep = ":", header = T)
+   mara11[, 1] <- as.character(mara11[, 1])
+   dmar11 <- mara11[, 8] + mara11[, 9]/60 + mara11[, 10]/3600
+   dmar11 <- cbind(mara11[, 2] + mara11[, 3]/60 + mara11[,
+     4]/3600, dmar11)
+   kmara11 <- mara11[, 1]
+   dimnames(dmar11) <- list(kmara11, c("1. Zwischenz.", "Endzeit"))
+   datasex11 <- substr(kmara11, 1, 1)
+   datakl11 <- substr(kmara11, 3, 5)
+ }
> f.mara_read()
```

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

Gibt es einen signifikanten Unterschied der Gesamt-Laufzeitmittel der Hauptklasse „H“-Männer und der Klasse „50“ der Männer, d.h. Alter zwischen 50 und 55?

18. Verwenden Sie die Patientendaten “Werner Blood Chemistry Data“. Betrachten Sie den „Body Mass Index“, d.h.

$$bmi = (\text{Gew. in kg}) / (\text{Größe in m})^2$$

(In unseren Daten könnten Sie die ‘pounds’ ungefähr mit 1/2 und die ‘inches’ mit 2.54/100 umrechnen, z.B. mit dem Befehl

```
bmi <- (werner_bcd[, "WEIGHT"]/2)/(werner_bcd[, "HEIGHT"] * 2.54/100)^2
.)
```

Präsentieren Sie die Werte von *bmi* grafisch.

Teilen Sie die Daten in 4 etwa gleich große Teile bezüglich *bmi* und stellen Sie für Kontrollzwecke die wirklichen Häufigkeiten in diesen Teilen dar.

19. Testen Sie auf Unabhängigkeit des Mittels von CHOLSTRL bezüglich dieser Gruppierung. Führen Sie das Gleiche für die anderen Variablen AGE, ALBUMIN, CALCIUM, URICACID durch.

20. Stellen Sie eine „Versuchsplanmatrix“  $\mathbf{X}$  für das vorige Modell „CHOLSTRL hängt linear von den Klassen von `bmi` ab“. Präsentieren Sie (zumindest auszugsweise) die Matrix.

Hinweis: Die Funktion `model.matrix` kann hilfreich sein.

Bitte, stellen Sie die Ausarbeitung in Form eines pdf-Files (nicht mehr als 3 Seiten) mit den Resultaten (Outputs plus textliche Kommentare) und Kurz-Listing des Programmcodes (Funktion) in die TUWEL-Seite

<https://tuwel.tuwien.ac.at/course/view.php?idnumber=107258-2011W>  
bis zum 28. November 2011.

Bitte den Namen des pdf-File folgendermaßen:

**name\_exer\_5.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/> .)

**Kreuzen Sie außerdem bitte im Internet** jede Nummer des Übungsbeispiels an, das Sie dann in der Übungsstunde an der Tafel (mit Beamer-Unterstützung) vorrechnen wollen und können. Der Termin ist üblicherweise 2 Stunden vor der Übung, d.h. Mi., 12:00. Siehe <http://www.statistik.tuwien.ac.at/cgi-bin/uebbsp/bspstart.cgi> .