

Übungszettel 1: Grundbegriffe der Statistik

Onlineübung

Aufgabe 1.1 Für eine Studie zur Lebenssituation von Studierenden der FH Dortmund sollen folgende Merkmale abgefragt werden:

- Wohnform (alleine, WG, Studentenwohnheim, bei den Eltern, sonstiges)
 - Geschlecht
 - monatliche Miete
 - Entfernung zwischen Wohnung und Fachhochschule
 - eigene Beurteilung der bisherigen Studienleistungen (1=sehr gut, ..., 5=sehr schlecht)
- (a) Was sind die Merkmalsträger? Geben Sie für jedes Merkmal das Skalenniveau und den Wertebereich an.
- (b) Die Mentorinnen des Fachbereichs Informatik erheben diese Merkmale an 20 Studierende in Ihrer Sprechstunde. Erläutern Sie die Begriffe Grundgesamtheit und Stichprobe in diesem Zusammenhang. Stellen Sie Überlegungen zur Repräsentativität der Befragung an.

Aufgabe 1.2 Erstellen Sie ein R-Skript mit den Befehlen und ergänzen Sie erläuternde Kommentare.

```
gewicht<-c(65.4,78.3,80.5,68.5)
gewicht

geschlecht<-c("m","m","w","m")
is.numeric(geschlecht)
is.logical(geschlecht)
is.character(geschlecht)

c(4,1.4/0,NA,0/0,-exp(1000))
```

Hinweis: Zu jeder in R implementierten Funktion existiert eine Dokumentation in Form einer Hilfeseite. Diese lässt sich durch `?Funktionsname` bzw. `help(Funktionsname)` aufrufen.

Aufgabe 1.3 Was passiert, wenn Sie folgende Befehle ausführen?

- | | |
|-------------------------------|---|
| (a) <code>x <- 4:10</code> | (b) <code>x<- seq(1,3,length=8)</code> |
| <code>x</code> | <code>x[4]</code> |
| <code>y <- c(2,9)</code> | <code>x[-8]</code> |
| <code>y</code> | <code>x[c(1,3,7)]</code> |
| <code>x / y</code> | <code>x[x > 1.5]</code> |

Aufgabe 1.4 Denken Sie sich einen Datensatz zu Aufgabe 1.1 mit Beobachtungen an 15 Studierenden aus. Erstellen Sie für jedes Merkmal in R einen Vektor mit den von Ihnen ausgedachten Beobachtungen. Erzeugen Sie dann mit Hilfe dieser Vektoren den ausgedachten Datensatz in Form eines `data.frame`. Extrahieren Sie dann einmal alle weiblichen Studierenden, einmal alle männlichen Studierenden, und einmal alle Studierenden mit einer Miete unter 300 Euro aus Ihrem Datensatz.

Weitere Aufgaben zur selbständigen Bearbeitung

Aufgabe 1.5 In einem Krankenhaus wurden 16 Herzoperation-Patienten, die nach einer neuen Methode operiert wurden, auf die Überlebenszeit nach der Operation untersucht.

Folgende Merkmale wurden unter anderem erhoben:

- Alter
- Geschlecht
- Nierenprobleme (keine, sporadisch, chronisch)
- Herz-Kreislauf-Vorerkrankungen (Herzklappenfehler, Herzrhythmusstörungen, Herzinfarkt etc.)
- Überlebenszeit in Monaten nach der Operation

Ihnen liegt ein Auszug aus den Daten vor:

<i>Patienten-ID</i>	<i>Nierenprobleme</i>	<i>Überlebenszeit in Monaten</i>
1	keine	4
2	keine	15
3	sporadisch	29
4	sporadisch	10
5	chronisch	4
6	keine	29
7	keine	30
8	chronisch	15
9	chronisch	2
10	keine	15
11	keine	32
12	chronisch	7
13	keine	15
14	sporadisch	36
15	keine	29
16	keine	39

- (a) Erläutern Sie am Merkmal Nierenprobleme die Begriffe Merkmalsträger, Merkmalsausprägungen, Wertebereich und Beobachtungen.
- (b) Welches Skalenniveau besitzen die fünf erhobenen Merkmalen?
- (c) Geben Sie die Grundgesamtheit und die Stichprobe an. Nennen Sie mögliche Bedenken bezüglich der Repräsentativität der Untersuchung. Warum ist es bei solchen Studien wichtig Merkmale wie Alter, Geschlecht und Vorerkrankungen zu erfassen?
- (d) Erstellen Sie in R in einem `data.frame` einen Datensatz mit den Werten der 16 Patienten für die Merkmale Nierenprobleme und Überlebenszeit in Monaten.

Aufgabe 1.6

- a) Informieren Sie sich über die Funktionen `seq()` und `rep()` in R durch Aufrufen der entsprechenden Hilfeseiten (`?seq` bzw. `help(rep)`).

Erzeugen Sie nun die folgenden Vektoren mittels einer der beiden genannten Funktionen:

$(6, 7, 6, 7, 6, 7)'$, $(0.8, 1.6, 2.4, 3.2, 4)'$, $(1, 1, 1, 2, 2, 2, 3, 3, 3)'$.

Konstruieren Sie den Vektor $(8, 6, 4, 2, 8, 6, 4, 2, 8, 6, 4, 2)'$ mit einer Kombination aus `seq()` und `rep()`.

- b) Finden Sie heraus, wie eine Zahl (oder die Zahlen in einem Vektor) auf zwei Nachkommastellen gerundet wird und tun Sie das mit

```
bitterunden1 <- 0.59487387
```

und

```
bitterunden2 <- seq(1,2,length=4).
```