

Lösung zu Kapitel 6: Beispiel 4

In `noten.dat` finden wir die Variable `note` in der alle Ergebnisse einer Prüfung eingetragen sind. Die Frage ist, ob sich diese Daten von der Verteilung vorheriger Jahre signifikant unterscheiden.

- Die Rohdaten werden aus `noten.dat` eingelesen und in das Objekt `noten` gespeichert. Danach erstellen wir eine Häufigkeitstabelle der Noten in `note`, berechnen die (auf 3 Dezimalstellen gerundeten) Prozentwerte in `proz` und geben die relativen Häufigkeiten der vergangenen Prüfungen in `popu` an.

Diese drei Objekte fassen wir mit `cbind` spaltenweise zusammen und speichern das neue Objekt in `noten`.

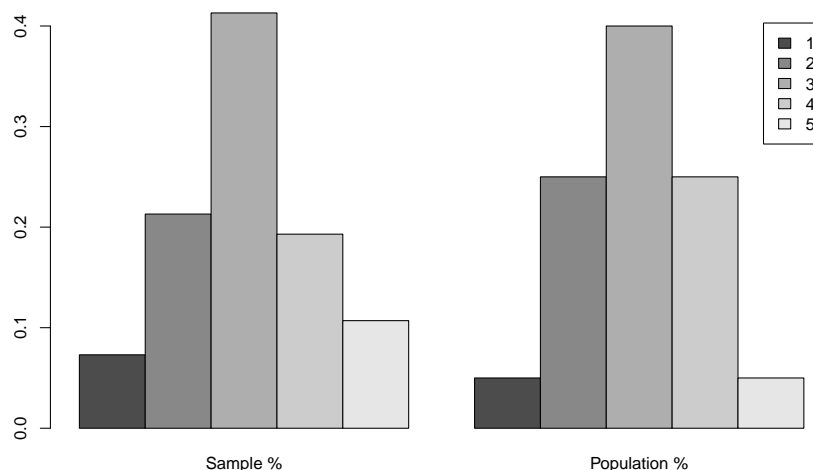
R

```
> noten <- read.table("noten.dat", header = TRUE)
> note <- table(noten$note)
> proz <- round(note/sum(note), digits = 3)
> popu <- c(0.05, 0.25, 0.4, 0.25, 0.05)
> noten <- cbind(`abs. H.` = note, `Sample %` = proz, `Population %` = popu)
```

- Um die zwei Verteilungen grafisch gegenüberzustellen fertigen wir je ein Balkendiagramm der Prozentwerte und der relativen Häufigkeiten unserer vorgegebenen Verteilung an.

R

```
> barplot(noten[, 2:3], beside = TRUE, legend = 1:5)
```



- Zur Prüfung der Forschungsfrage kommt ein χ^2 -Test zum Einsatz bei dem wir einerseits beobachteten Häufigkeiten (`note`) gegen die Verteilung von `popu` vergleichen. Wir sehen dass $p = .007$, d.h. das Ergebnis ist signifikant weshalb wir die Nullhypothese verwerfen und die Alternativhypothese – die beobachtete Verteilung weicht von der spezifizierten ab – annehmen.

R

```
> chisq.test(x = note, p = popu)
```

Chi-squared test for given probabilities

data: note

X-squared = 14.0667, df = 4, p-value = 0.007085