

Lösung zu Kapitel 6: Beispiel 5

`unfaelle.dat` enthält in `auto` die Unfallhäufigkeiten nach Autoklassen. Wir kennen die Verteilung in der Population und wollen wissen, ob gewisse Autoklassen öfters in Unfällen verwickelt sind.

- Die Daten werden aus `unfaelle.dat` eingelesen und in das Objekt `unf` gespeichert. In `auto` speichern wir eine Häufigkeitstabelle, in `proz` die (auf 3 Dezimalstellen gerundeten) Prozentwerte aus `auto` und in `popu` speichern wir die relativen Häufigkeiten der Zulassungen. Diese Information fassen wir mittels `cbind` in `unfaelle` zusammen und vergeben zur besseren Übersicht noch entsprechende Zeilenamen.

R

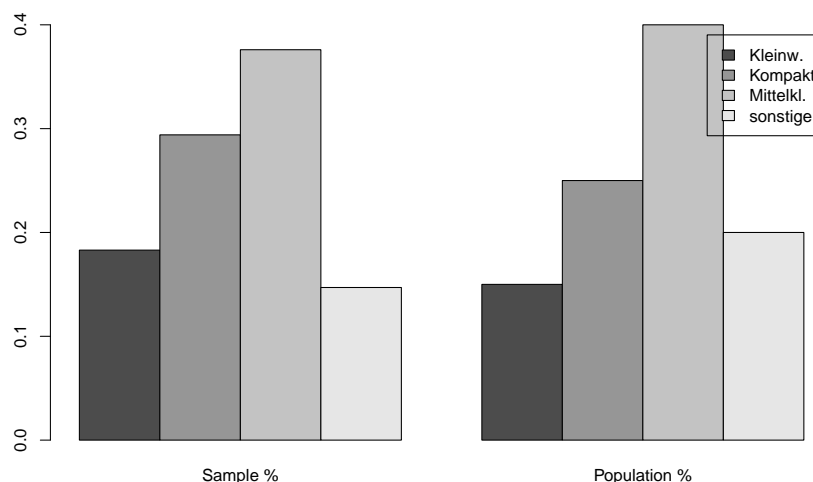
```
> unf <- read.table("unfaelle.dat", header = TRUE)
> auto <- table(unf$auto)
> proz <- round(auto/sum(auto), digits = 3)
> popu <- c(0.15, 0.25, 0.4, 0.2)
> unfaelle <- cbind(`abs. H.` = auto, `Sample %` = proz, `Population %` = popu)
> rownames(unfaelle) <- c("Kleinw.", "Kompakt", "Mittelkl.", "sonstige")
> unfaelle
```

	abs. H.	Sample %	Population %
Kleinw.	36	0.183	0.15
Kompakt	58	0.294	0.25
Mittelkl.	74	0.376	0.40
sonstige	29	0.147	0.20

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

R

```
> barplot(unfaelle[, 2:3], beside = TRUE, legend = rownames(unfaelle))
```



- Mit einem χ^2 -Test prüfen wir, ob die Häufigkeiten in `auto` signifikant von der Verteilung in `popu` abweichen und sehen dass dies mit $p = .112$ nicht der Fall ist. Wir verwerfen daher die Nullhypothese nicht und bleiben bei der Annahme, dass keine Autoklasse häufiger in Unfälle verwickelt ist.

R

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

Chi-squared test for given probabilities

data: auto

X-squared = 6, df = 3, p-value = 0.1116