

Lösung zu Kapitel 10: Beispiel 1

In `teenagework.csv` wird die Hausarbeit von Teenagern (`stunden`) inklusive der Berufstätigkeit der Mutter (`mutter`) und dem Geschlecht (`sex`) erfasst. Es soll geprüft werden, ob sich die Anzahl der Stunden zwischen Männern und Frauen unterscheidet.

- Die Datei `teenagework.csv` wird in das Objekt `work` eingelesen und mit `attach` kann man direkt auf die Variablen zugreifen. Mit `factor` erstellen wir ein Objekt `sx`, in dem das Geschlecht der Teenager mit Labels erfasst ist.

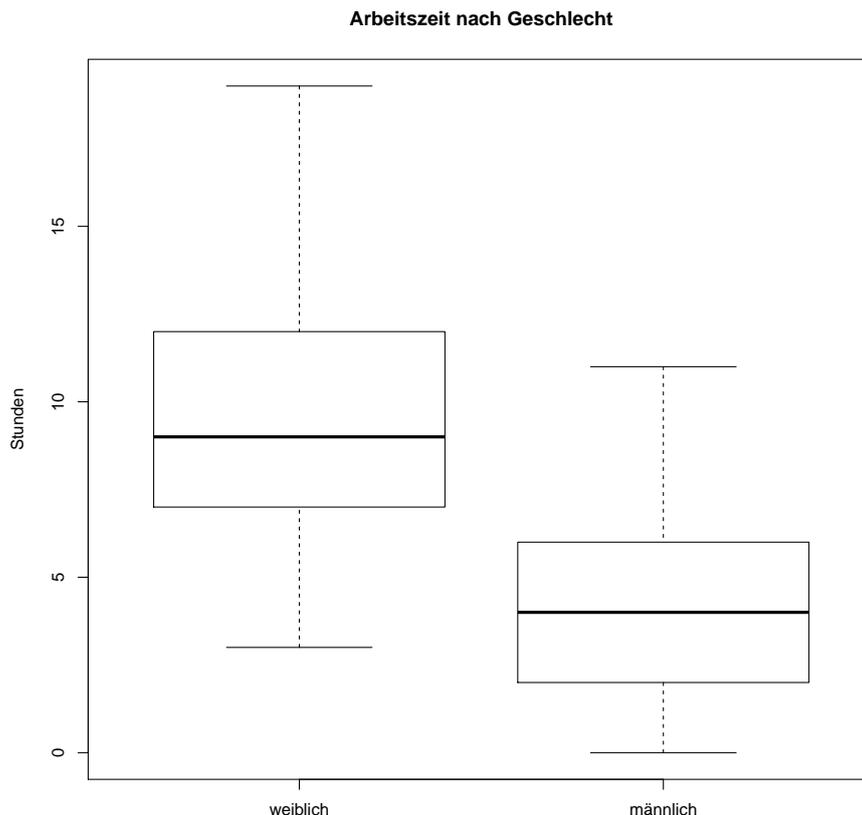
R

```
> work <- read.csv2("teenagework.csv", header = TRUE)
> attach(work)
> sx <- factor(sex, labels = c("weiblich", "männlich"))
```

- Um die Unterschiede zu visualisieren verwenden wir einen Boxplot.

R

```
> boxplot(stunden ~ sx, ylab = "Stunden", main = "Arbeitszeit nach Geschlecht")
```



- Zur deskriptiven Analyse der Daten berechnen wir zuerst Mittelwerte und Standardabweichungen der Gruppen. Mit `tapply` berechnen wir jeweils `mean` (bzw. `sd`) für alle Kategorien von `sx` und speichern die Ergebnisse in `mws` (bzw. `sds`) ab. Zur Darstellung der Ergebnisse fassen wir `mws` und `sds` mit `rbind` zeilenweise zusammen.

Ein t -Test zeigt mit $p < .001$, dass es Unterschiede gibt und Frauen signifikant mehr im Haushalt arbeiten als Männer.

Da die Verteilungen, wie im Boxplot ersichtlich, schief sind kann man statt einem t -Test auch einen Rangsummentest nach Wilcoxon rechnen, der dasselbe Ergebnis liefert ($p < .001$).

R

```
> mws <- tapply(stunden, sx, mean)
> sds <- tapply(stunden, sx, sd)
> rbind(Mittelwerte = mws, Standardabw = sds)
```

```
          weiblich männlich
Mittelwerte 9.366337 4.439560
Standardabw 3.233335 2.729301
```

R

```
> t.test(stunden ~ sx)
```

Welch Two Sample t-test

```
data:  stunden by sx
t = 11.4432, df = 189.219, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 4.077495 5.776057
sample estimates:
mean in group weiblich mean in group männlich
          9.366337              4.439560
```

R

```
> wilcox.test(stunden ~ sx)
```

Wilcoxon rank sum test with continuity correction

```
data:  stunden by sx
W = 8080.5, p-value < 2.2e-16
alternative hypothesis: true location shift is not equal to 0
```

R

```
> detach(work)
```