

Statistik Übung 3

Michael Rynkiewicz

21/03/2019

Aufgabe 17

```
traveledDistance <- c(39, 35, 25, 37, 15, 36, 50, 52, 37, 51, 39)
traveledDistance.length <- length(traveledDistance)
traveledDistance.mean <- sum(traveledDistance) / traveledDistance.length
traveledDistance.table <- table(traveledDistance)
traveledDistance.sorted <- sort(traveledDistance)

traveledDistance.modus <- names(traveledDistance.table[which(traveledDistance.table == max(traveledDistance.table))])
traveledDistance.firstQuartil <- traveledDistance.sorted[ceiling(traveledDistance.length * 0.25)]
traveledDistance.secondQuartil <- traveledDistance.sorted[ceiling(traveledDistance.length * 0.50)]
traveledDistance.thirdQuartil <- traveledDistance.sorted[ceiling(traveledDistance.length * 0.75)]

traveledDistance.variance <- sum((traveledDistance - traveledDistance.mean) ^ 2) / traveledDistance.length
traveledDistance.sd <- sqrt(traveledDistance.variance)
traveledDistance.gamma <- sum(((traveledDistance - traveledDistance.mean) / traveledDistance.sd)^3) / traveledDistance.length

traveledDistance
```

```
## [1] 39 35 25 37 15 36 50 52 37 51 39
```

```
traveledDistance.sorted
```

```
## [1] 15 25 35 36 37 37 39 39 50 51 52
```

a)

Modus

```
paste("{ ", paste(traveledDistance.modus, collapse = ", "), " }")
```

```
## [1] "{ 37, 39 }"
```

Mean

```
traveledDistance.mean
```

```
## [1] 37.81818
```

b)

First Quartil

```
traveledDistance.firstQuartil
```

```
## [1] 35
```

Second Quartil

```
traveledDistance.secondQuartil
```

```
## [1] 37
```

Third Quartil

```
traveledDistance.thirdQuartil
```

```
## [1] 50
```

c)

Spannweite

```
max(traveledDistance) - min(traveledDistance)
```

```
## [1] 37
```

Standardabweichung

```
traveledDistance.sd
```

```
## [1] 10.54702
```

Empirische Schiefe

```
traveledDistance.gamma
```

```
## [1] -0.5472377
```

D.h. die Radfahrerin hat in den letzten Tagen der Radtour größere Strecken zurückgelegt.

Aufgabe 18

```
xi <- c(2, 3, 4, 5, 6)
fi.a <- c(3, 6, 3, 2, 4)
fi.b <- c(4, 1, 5, 4, 6)
```

```
xi.fi.a <- xi * fi.a
xi.fi.b <- xi * fi.b
```

Arithmetisches Mittel für fi-A

```
sum(xi.fi.a) / sum(fi.a)
```

```
## [1] 3.888889
```

Arithmetisches Mittel für fi-B

```
sum(xi.fi.b) / sum(fi.b)
```

```
## [1] 4.35
```

Das arithmethisches Mittel ist gleich für diskrete und stetige Merkmale.

Aufgabe 19

```
apple.prices <- c(1.14, 1.32, 1.45)
apple.amount <- c(30, 10, 5)
apple.wheights <- apple.amount / sum(apple.amount)
```

```
apple.unweighted <- sum((apple.prices * apple.amount)) / sum(apple.amount)
apple.wheighted <- sum((apple.prices * apple.wheights)) / sum(apple.wheights)
```

Preis alleine

```
sum(apple.prices) / length(apple.prices)
```

```
## [1] 1.303333
```

Ungewichtet

```
apple.unweighted
```

```
## [1] 1.214444
```

Gewichtet

```
apple.wweighted
```

```
## [1] 1.214444
```

Aufgabe 20

```
years <- c(2011, 2013, 2014, 2015, 2016, 2017, 2018)
```

```
turnover <- c(121, 152, 136, 140, 122, 131, 136)
```

```
turnover.average <- sum(turnover) / length(turnover)
```

```
turnover.growthrate <- prod(turnover[1:length(turnover) - 1] / turnover[2:length(turnover)])^(1/(length
```

Durchschnittlicher Umsatz

```
turnover.average
```

```
## [1] 134
```

Umsatzwachstumsrate [%]

```
round(turnover.growthrate, 2)
```

```
## [1] -0.02
```

Aufgabe 21

```
income.employees <- c(1390, 1780, 1511, 1688, 1621, 1410, 1320)
```

```
income.employer <- c(9735)
```

```
income <- sort(c(income.employees, income.employer))
```

```
income.average <- mean(income)
```

```
income.median <- 0
```

```
if (length(income) %% 2 == 0){
```

```
  income.median <- (income[length(income) / 2] + income[(length(income) / 2) + 1]) / 2
```

```
} else {
```

```
  income.median <- income[(length(income) + 1) / 2]
```

```
}
```

Sortiertes Einkommen

```
income
```

```
## [1] 1320 1390 1410 1511 1621 1688 1780 9735
```

Durchschnitt

```
income.average
```

```
## [1] 2556.875
```

Median

```
income.median
```

```
## [1] 1566
```

Der Median ist hier die bessere Lagemaßzahl, da das Einkommen der Geschäftsinhaberin ein Ausreißer ist und damit die Statistik verfälscht.

Geeignete Streuungsmaßzahl
Standardabweichung

```
sd(income)
```

```
## [1] 2904.687
```

Die Spannweite ist nicht aussagekräftig.

Die Standardabweichung ist besser als die Varianz, sollte ein Mitarbeiter eine Gehaltserhöhung erhalten.

Aufgabe 22

```
phoneBill <- c(71.92, 74.20, 81.6, 218.4, 72.46, 73.68, 60.88, 63.36, 81.08, 63.20, 74.32, 77.88)
phoneBill.mean <- mean(phoneBill)
```

Mittlere Monatliche Kosten

```
phoneBill.mean
```

```
## [1] 84.415
```

Verschleierung eines hohen Telefonmittelwertes:

“Eine Berechnung aller Monate für den Mittelwert ist nicht notwendig, daher wird nur jeder Zweite verwendet”

```
phoneBill.deception1 <- phoneBill[seq(from=1, to=length(phoneBill), by = 2)]
mean(phoneBill.deception1)
```

```
## [1] 73.71
```

“Man verwendet den Wert in der Mitte anstelle des Mittelwertes”

Verwenden des Median anstelle des Mittelwertes

```
median(phoneBill)
```

```
## [1] 73.94
```

Aufgabe 23

$$576p_1 + 496p_2 = 517p_1 + p_2 = 1$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 576 & 496 & 517 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -80 & -59 \end{pmatrix}$$

$$\Rightarrow p_2 = 0.737 \Rightarrow p_1 = 0.2625$$

26,25% aller Kunden reisten nach Nordstrand.

73,75% aller Kunden reisten nach Neßmersiel.

100,0% aller Kunden reisten in eine erholsame Zeit.

Aufgabe 24

```
weeklyMeals <- c(0, 1, 2, 3, 4, 5)
weeklyMeals.props <- c(0.1, 0.3, 0.3, 0.1, 0.2, 0.0)
weeklyMeals.mean <- sum(weeklyMeals * weeklyMeals.props)
weeklyMeals.table <- matrix(c(weeklyMeals, weeklyMeals.props), ncol = 2)
weeklyMeals.sorted <- weeklyMeals[order(weeklyMeals.props)]
weeklyMeals.median <- median(weeklyMeals.sorted)
weeklyMeals.variance <- sum((weeklyMeals - weeklyMeals.mean)^2) / length(weeklyMeals)
weeklyMeals.sd <- sqrt(weeklyMeals.variance)
```

Durchschnittliche Anzahl an Mittagessen in der Betriebskantine

```
weeklyMeals.mean
```

```
## [1] 2
```

Median

```
weeklyMeals.median
```

```
## [1] 2.5
```

Varianz

```
round(weeklyMeals.variance, 2)
```

```
## [1] 3.17
```

Standardabweichung

```
round(weeklyMeals.sd, 2)
```

```
## [1] 1.78
```

Aufgabe 25

```
generic.breaks <- c(1, 6, 11, 16, 21, 26, 31, 36)
generic.xi <- (generic.breaks[1:length(generic.breaks)-1] + generic.breaks[2:length(generic.breaks)]) /
generic.a <- c(0, 10, 26, 28, 25, 11, 0)
generic.b <- c(3, 5, 24, 40, 20, 4, 4)
generic.c <- c(0, 4, 24, 40, 20, 4, 8)
generic.names <- c("A", "B", "C")

generic.means <- c(
  sum(generic.a * generic.xi) / sum(generic.a),
  sum(generic.b * generic.xi) / sum(generic.b),
  sum(generic.c * generic.xi) / sum(generic.c)
)
generic.medians <- c(median(generic.a), median(generic.b), median(generic.c))

names(generic.means) <- generic.names
names(generic.medians) <- generic.names

generic.data <- data.frame(
  "Arithmetisches Mittel" = generic.means,
  "Median" = generic.medians,
  "Erstes Quartil" = c("-", median(generic.b[generic.b < generic.medians[2]]), "-"),
```

```

"Drittes Quartil" = c("-", median(generic.b[generic.b > generic.medians[2]]), "-"),
"Varianz" = c(sum((generic.a * generic.xi - generic.means[1])^2) / sum(generic.a) , "-", "-")
)
generic.data

```

	Arithmetisches.Mittel	Median	Erstes.Quartil	Drittes.Quartil	Varianz
## A	18.55	11	-	-	7757.767175
## B	18.35	5	4	24	-
## C	19.50	8	-	-	-

Aufgabe 26

Gesamtpreis Vorjahr: 1

Computerpreis Vorjahr: 0.6

Softwarepreis Vorjahr: 0.4

10% Erhöhung des Gesamtpreises

20% Senkung des Computerpreises

Gesamtpreis: 1.10

Computerpreis: $0.6 * 0.8 = 0.48$

Softwarepreise: $1.10 - 0.48 = 0.62$

=>

Der Softwarepreis ist um 55% gestiegen im Vergleich zum Vorjahr.