Statistik und Qualität - Ausarbeitung der ersten Übung

Stefan Dünser

# A) TECHNISCHE FERTIGKEITEN

### A-1) Handelt es sich bei den vorliegenden statistischen Gesamtheiten um Bestands- oder Bewegungsgrößen?

* **a) Studierende an einer Hochschule.**
  + Bestandsgröße
* **b) Hochzeiten am Standesamt einer Gemeinde.**
  + Bewegungsgröße
* **c) Bei der Behörde gemeldete Personenkraftwagen.**
  + Bestandsgröße
* **d) Maschinenausfälle in einer Werkstatt.**
  + Bewegungsgröße
* **e) Wartende Kunden vor einem Abfertigungsschalter.**
  + Bestandsgröße

### A-2) Im Servicecenter eines Unternehmens werden über einen Zeitraum eines Tages die eingehenden Anrufe aufgezeichnet. Gezählt wird die Anzahl der pro 10-Minuten-Zeitintervall eingehenden Anrufe. Für 40 derartige Zeitintervalle erhält man folgende Ergebnisse:

Liste.vec <- c(0, 0, 1, 3, 4, 1, 2, 2, 1, 1, 1, 2, 3, 0, 2, 0, 1, 3, 1, 2, 2, 0, 1, 1, 6, 1, 0, 2, 3, 1, 1, 4, 2, 3, 2, 0, 3, 0, 1, 2)  
Liste.vec

## [1] 0 0 1 3 4 1 2 2 1 1 1 2 3 0 2 0 1 3 1 2 2 0 1 1 6 1 0 2 3 1 1 4 2 3 2 0 3 0  
## [39] 1 2

##### **a) Was stellt bei dieser Fragestellung die statistische Grundgesamtheit dar? Was sind die beobachteten Merkmale der statistischen Einheiten und wie sind sie skaliert?**

* Grundgesamtheit
  + Die Grundgesamtheit setzt sich auch sen statistischen Einheiten zusammen. Die Grundgesamtheit sind die Anzahl der 10-Minuten-Zeitintervalle - (40).
* Merkmale
  + Anzahl der 10-Minuten-Zeitintervalle (40)
  + Anzahl der Anrufe über alle Zeitintervalle (65)
* Skalierung
  + Metrische Skalierung (Kardinalskala), die verhältnisskaliert ist

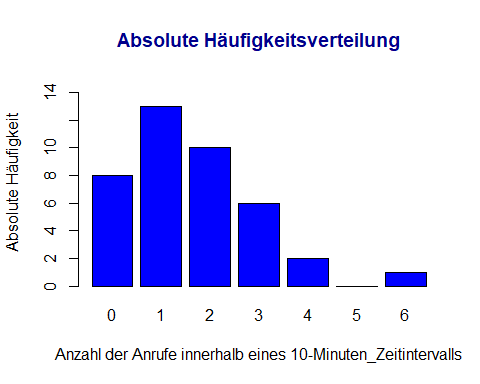
##### **b) Ermittle die absolute und relative Häufigkeitstabelle der eingehenden Anrufe und stelle die Häufigkeitsverteilung und Summenhäufigkeit grafisch dar.**

##### **Absolute Häufigkeit:**

table(Liste.vec)

## Liste.vec  
## 0 1 2 3 4 6   
## 8 13 10 6 2 1

barplot(table(factor(Liste.vec,levels=c(0,1,2,3,4,5,6))), ylim = c(0,15), xlab = "Anzahl der Anrufe innerhalb eines 10-Minuten\_Zeitintervalls", ylab = "Absolute Häufigkeit", main = "Absolute Häufigkeitsverteilung", col.main = "darkblue", col = "blue")

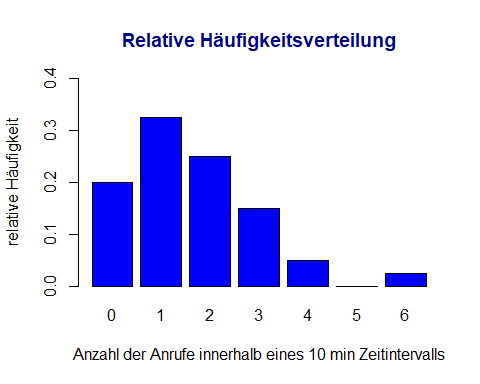


##### **Relative Häufigkeit:**

table(Liste.vec)/length(Liste.vec)

## Liste.vec  
## 0 1 2 3 4 6   
## 0.200 0.325 0.250 0.150 0.050 0.025

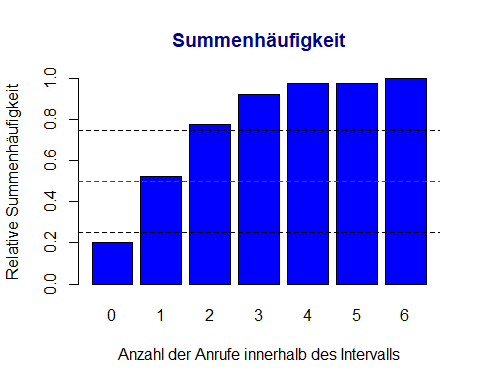
barplot(table(factor(Liste.vec,levels=c(0,1,2,3,4,5,6)))/sum(table(Liste.vec)), ylim = c(0,0.4), xlab = "Anzahl der Anrufe innerhalb eines 10 min Zeitintervalls",ylab = "relative Häufigkeit", main = "Relative Häufigkeitsverteilung", col.main = "darkblue", col = "blue")

 ##### **Summenhäufigkeit**

cumsum(table(Liste.vec))

## 0 1 2 3 4 6   
## 8 21 31 37 39 40

barplot(cumsum(table(factor(Liste.vec,levels = c(0,1,2,3,4,5,6))))/sum(table(Liste.vec)), xlab = "Anzahl der Anrufe innerhalb des Intervalls", ylab = "Relative Summenhäufigkeit", main = "Summenhäufigkeit", col.main = "darkblue", col = "blue")  
abline(0.25, 0, lty = "dashed")  
abline(0.5, 0, lty = "dashed", col = "red")  
abline(0.75, 0, lty = "dashed")



##### **c) Schätze das arithmetische Mittel aus der grafischen Darstellung für die Häufigkeit und den Median aus der grafischen Darstellung für die Summenhäufigkeit.**

**Arithmetisches Mittel:** Das arithmestische Mittel kann als Schwerpunkt der Häufigkeitsverteilung angesehen werden. In diesem Fall beträgt das arithmetische Mittel nach Abschätzung rund 1,5, da somit links und rechts der x-Achse in etwa gleich viele Werte sind. **Median:** Der Median wird über die y-Achse ermittelt. Der Median befindet sich dann auf der x-Achse an dem Punkt, an dem 50% des y-Werts erreicht sind. In diesem Fall ist der Median bei 1.

##### **d) Berechne die möglichen Lageparameter (Zentralmaße und Streumaße)**

**Zentralmaße:**  
***Modus***

getmode <- function(Liste.vec) {  
 uniqv <- unique(Liste.vec)  
 uniqv[which.max(tabulate(match(Liste.vec, uniqv)))]  
}  
mode <- getmode(Liste.vec)  
mode

## [1] 1

**Median**

median(Liste.vec)

## [1] 1

**Arithmetisches Mittel**

mean(Liste.vec)

## [1] 1.625

***Streumaße:*** Minimum

Minimum <- min(Liste.vec)  
Minimum

## [1] 0

**Maximum**

Maximum <- max(Liste.vec)  
Maximum

## [1] 6

**Spannweite**

Spannweite <- range(Liste.vec)  
Spannweite

## [1] 0 6

**Quantile**

quantile(Liste.vec)

## 0% 25% 50% 75% 100%   
## 0 1 1 2 6

**Mittlere absolute Abweichung**

mean(abs(Liste.vec-mean(Liste.vec)))

## [1] 1.05625

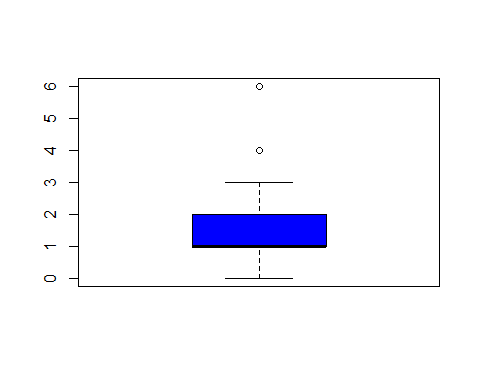
**Standardabweichung**

sd(Liste.vec)

## [1] 1.333734

##### **e) Stelle die Daten in einem Boxplot dar**

boxplot(Liste.vec, col = "blue")

  
Besonders Auffällig ist bei diesem Boxplot bzw. den Werten aus der Liste, dass der Median und das erste Quartil (25% Quantil) zusammenfallen. Die Werte 4 Anrufe sowie 6 Anrufe innerhalb eines 10-Minuten\_Zeitintervalls werden hier als Ausreißer aufgefasst und als Kreis dargestellt.

### A-3) Eine Anzahl von 1000 Kleinmotoren weist folgende Lebensdauer auf:

AnzahlMotoren <- c(33,276,404,237,50)  
Lebensdauer <- c(0:10)

##### **a) Stelle die Häufigkeitsverteilung und Summenhäufigkeit grafisch dar.**

##### **b) Schätze das arithmetische Mittel aus der grafischen Darstellung für die Häufigkeit und den Median aus der grafischen Darstellung für die Summenhäufigkeit.**

##### **c) Bestimme der Anteil der Motoren mit über 6 Jahren Lebensdauer.**

##### **d) Berechne die möglichen Lageparameter (Zentralmaße und Streumaße).**

### A-4) Ein technisches Servicecenter zeichnet an 100 Tagen die Häufigkeit der Einsätze auf. Es ergibt sich folgende Tabelle:

AnzahlTage <- c(16,48,27,9)  
AnzahlEinsaetze <- c(0:80)

##### **a) Stelle die Häufigkeitsverteilung und Summenhäufigkeit grafisch dar.**

##### **b) Schätze das arithmetische Mittel aus der grafischen Darstellung für die Häufigkeit und den Median aus der grafischen Darstellung für die Summenhäufigkeit.**

##### **c) Bestimme der Anteil der Tage mit über 20 Einsätzen.**

##### **d) Berechne die möglichen Lageparameter (Zentralmaße und Streumaße). Welcher Aspekt könnte hier problematisch sein? Warum?**

# B) VERSTÄNDNISFRAGEN

### B-1) Zeige, dass das arithmetische Mittel unter dem Schwerpunkt der Häufigkeitsfunktion liegt.

### B-2) Verschiebungssatz zur Berechnung der Standardabweichung: Das arithmetische Mittel ist gegeben durch und die Varianz ist definiert als $\sigm^2$. Zeige die direkte Umformung, dass daraus folgt (Verschiebungssatz): $$\sigm^2\ =\ $$