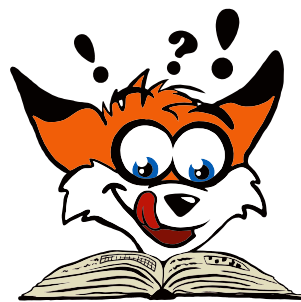


Deskriptive Statistik

Andreas Schneider

Version 1.0



Mathe**b**ibel

Inhaltsverzeichnis

Deskriptive Statistik	3
Merkmal	6
Lageparameter	9
Arithmetisches Mittel	11
Geometrisches Mittel	14
Harmonisches Mittel	17
Median	20
Modus	23
Streuungsparameter	25
Spannweite	26
Interquartilsabstand	30
Mittlere absolute Abweichung	35

Deskriptive Statistik

In diesem Kapitel schauen wir uns an, was die deskriptive Statistik ist.

Aufgabe der deskriptiven Statistik ist es, große Datenmengen auf einige wenige Maßzahlen zu reduzieren, um damit komplexe Sachverhalte übersichtlich darzustellen.

In der deskriptiven Statistik werden Daten erhoben, aufbereitet und analysiert.

Bei einer **Datenerhebung** werden ein oder mehrere **Merkmale** (Beobachtungsmerkmale) untersucht.

Beispiel:

Uns interessiert die Körpergröße aller Schüler in der Klasse.

→ Körpergröße = Beobachtungsmerkmal

Merkmale können verschiedene Werte annehmen, die **Merkmalsausprägungen** genannt werden.

Beispiele:

Merkmal	Merkmalsausprägung
Geschlecht	männlich, weiblich
Schulnote	1, 2, 3, 4, 5, 6
Körpergröße	x cm
Alter	x Jahre

Die Menge aller Elemente, auf die ein Untersuchungsziel in der Statistik gerichtet ist, heißt **Grundgesamtheit**.

Beispiel:

Uns interessiert die Körpergröße aller Schüler in einer Klasse.

→ Menge aller Schüler der Klasse = Grundgesamtheit

In unserem Beispiel untersuchen wir die Körpergröße von Schülern einer Klasse. In der Praxis beschäftigt sich die Statistik meist mit viel größeren Grundgesamtheiten. Stellen wir uns nur mal vor, wir würden uns für die Körpergröße aller 9. Klässler in Deutschland interessieren. Aus zeitlichen und nicht zuletzt finanziellen Gründen ist es fast unmöglich, hunderttausende Schüler nach ihrer Körpergröße zu befragen bzw. sie zu messen. In so einem Fall würde man die Untersuchung auf eine **Stichprobe** beschränken. Eine Stichprobe ist eine zufällige Teilmenge der Grundgesamtheit.

Beispiel:

Uns interessiert die Körpergröße (= Beobachtungsmerkmal) aller 9. Klässler in Deutschland (= Grundgesamtheit). Aus dieser Menge wählen wir zufällig 1000 Schüler aus (= Stichprobe).

Eine Datenerhebung der Grundgesamtheit nennt man **Vollerhebung**, wohingegen man eine Datenerhebung einer Stichprobe als **Stichprobenerhebung** bezeichnet.

Beispiel:

Eine Klasse hat 30 Schüler. Gegenstand der Untersuchung ist das Merkmal Körpergröße. Befragt man alle 30 Schüler der Klasse nach ihrer Körpergröße, handelt es sich um eine Vollerhebung. Wenn man dagegen nur einen Teil der Schüler (z.B. 5 Schüler) befragt, handelt es sich um eine Stichprobenerhebung.

Die in einer Stichprobe beobachteten Werte heißen **Stichprobenwerte** oder **Beobachtungswerte**.

Beispiel:

Wir messen die Körpergröße von 5 Schülern der Klasse.

Die Stichprobenwerte sind dann die gemessenen Körpergrößen, z.B:

172 cm, 164 cm, 167 cm, 175 cm, 159 cm;

Begriff	Mathematische Bezeichnung
Stichprobe vom Umfang N	(x_1, \dots, x_n)
Stichprobenwerte / Beobachtungswerte	x_1, \dots, x_n

Übrigens: Bei (x_1, \dots, x_n) handelt es sich um ein n -Tupel.

Statistische Maßzahlen

eindimensionaler Stichproben

Lageparameter

Arithmetisches Mittel

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$$

Geometrisches Mittel

$$\bar{x}_{\text{geom}} = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n}$$

Harmonisches Mittel

$$\bar{x}_{\text{harm}} = \frac{n}{\frac{1}{x_1} + \dots + \frac{1}{x_n}}$$

Median

$$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} & \text{für } n \text{ ungerade} \\ \frac{1}{2} (x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & \text{für } n \text{ gerade} \end{cases}$$

Modus

$$\tilde{x}_d = \text{Häufigster Beobachtungswert}$$

Streuungsparameter

Spannweite

(engl. range)

$$R = x_{\max} - x_{\min}$$

Interquartilsabstand

(engl. interquartile range)

$$IQR = Q_{0,75} - Q_{0,25}$$

Mittlere absolute Abweichung

(engl. average absolute deviation)

$$D = \frac{1}{n} \cdot \sum_{i=1}^n |x_i - \bar{x}|$$
