

1. Einführung zur Statistik

Hajo Holzmann

Philipps-Universität Marburg

Was ist Statistik?

Statistik = Wissenschaft vom Umgang mit Daten

Statistik = Wissenschaft vom Umgang mit Daten

Phasen einer **statistischen Studie**

① Studiendesign

Welche Daten sollen erhoben werden? Wie sollen diese erhoben werden?

Bsp.: Entwerfen von Fragebögen, um viele, präzise Antworten zu bekommen.

② Datenerhebung / verarbeitung

③ Datenauswertung

- alle (potentiell) benötigten Daten erfassen
→ auch **Meta-Daten** wie Erfassungszeit, -ort und -methode.

- alle (potentiell) benötigten Daten erfassen
→ auch **Meta-Daten** wie Erfassungszeit, -ort und -methode.
- **Rohdaten** (einzelne Messwerte) digitalisieren
→ erst später zu benötigten Größen (Mittelwerte, Indizes)
zusammenfassen

- alle (potentiell) benötigten Daten erfassen
→ auch **Meta-Daten** wie Erfassungszeit, -ort und -methode.
- **Rohdaten** (einzelne Messwerte) digitalisieren
→ erst später zu benötigten Größen (Mittelwerte, Indizes) zusammenfassen
- **Qualitätsmanagement** bei Datenerfassung
→ zwei Personen geben die selben Daten ein, Abgleichen.
→ liegen digitalisierte Daten im physikalischen Messbereich?

- **deskriptive (beschreibende) Statistik**

→ Zusammenfassen der Daten, um wesentliche Eigenschaften zu erkennen.

z.B. Kennzahlen für Lage der Daten, Streuung, Diagramme für Verteilung.

- **deskriptive (beschreibende) Statistik**

→ Zusammenfassen der Daten, um wesentliche Eigenschaften zu erkennen.

z.B. Kennzahlen für Lage der Daten, Streuung, Diagramme für Verteilung.

- **induktive (schließende, mathematische) Statistik**

Problem: Daten nur aus kleiner Stichprobe, mit Messfehlern, unpräzise oder natürlichen Schwankungen unterworfen.

→ Unsicherheit (Ungenauigkeit, Unwägbarkeit) in Daten mit berücksichtigen

- Beobachtungen x_1, \dots, x_n
- Interpretiere als Realisierung von unabhängigen Zufallsvariablen X_1, \dots, X_n
- Verteilung P_X der X_j unbekannt

- Beobachtungen x_1, \dots, x_n
- Interpretiere als **Realisierung** von **unabhängigen Zufallsvariablen** X_1, \dots, X_n
- Verteilung P_X der X_j unbekannt

→ ziehe Schlüsse über mögliche Verteilung P_X anhand der Beobachtungen

auf diese Weise **quantifiziere Unsicherheit** in den Beobachtungen x_j .

→ Hilfsmittel: **Wahrscheinlichkeitstheorie**