

Topic 6: Kausalitaet, A/B-Tests und Logistische Regression

Block 3: Komplexitaet, Kausalitaet & Generalisierung

January 25, 2026

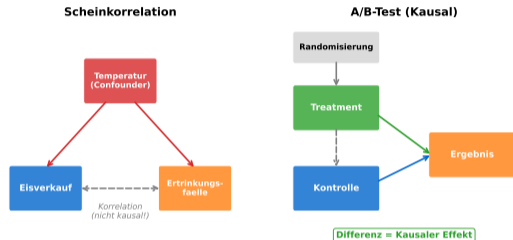
Korrelation bedeutet nicht Kausalitaet

- Zwei Variablen koennen zusammen variieren, ohne dass eine die andere verursacht
- **Confounder**: Dritte Variable, die beide beeinflusst
- Beispiel: Eisverkauf und Ertrinkungsfaelle (Confounder: Temperatur)

Kausale Aussagen erfordern:

- Zeitliche Abfolge
- Mechanismus/Theorie
- Randomisierung oder starke Annahmen

Korrelation vs. Kausalitaet



Beobachtete Zusammenhaenge sind selten kausal - Confounding ist die Regel, nicht die Ausnahme

Randomisiertes kontrolliertes Experiment

Design:

- **Gruppe A (Kontrolle):** Keine Intervention
- **Gruppe B (Treatment):** Intervention (z.B. neues Feature)
- **Zufaellige Zuteilung** zu A oder B

Warum funktioniert es?

- Randomisierung balanciert alle Confounder (auch unbeobachtete!)
- Einziger systematischer Unterschied: Treatment
- Differenz in Outcomes = **kausaler Effekt**

Beispiel: Button-Farbe

- A: Blauer Button (5% Click-Rate)
- B: Gruener Button (6% Click-Rate)
- Differenz: +1 Prozentpunkt

Interpretation:

- Gruener Button *verursacht* hoehere Click-Rate
- Nicht nur Korrelation, sondern Kausalitaet
- Voraussetzung: Ausreichende Sample Size

Randomisierung ist der Schluessel zu kausalen Aussagen - sie eliminiert Confounding durch Design

Systematischer Prozess fuer valide Tests

1. Hypothese formulieren

- Was erwarten wir? Warum?
- Beispiel: "Vereinfachte Checkout-Seite erhoht Conversion"

2. Primaere Metrik waehlen

- Eine Haupt-KPI (z.B. Conversion-Rate)
- Sekundaere Metriken zur Kontrolle
- Wichtig: *vor* dem Test festlegen!

3. Effektgroesse schaeetzen

- Minimum Detectable Effect (MDE)
- Was ist praktisch relevant?

4. Sample Size berechnen

- Abhaengig von MDE, Signifikanz-Level, Power
- Zu klein: Effekt nicht detektierbar
- Zu gross: Verschwendung

5. Laufzeit planen

- Wochenzyklen beruecksichtigen
- Mindestens 1-2 volle Wochen

Haeufige Fehler:

- Peeking (vorzeitiges Stoppen)
- Multiple Testing ohne Korrektur
- Selection Bias bei Zuteilung

Gute Planung ist entscheidend - ein schlecht designter Test ist schlimmer als kein Test

Problem mit linearer Regression

- Binäre Outcomes: $Y = 0$ oder 1
- OLS kann Werte < 0 oder > 1 vorhersagen
- Residuen nicht normalverteilt

Loesung: Logit-Modell

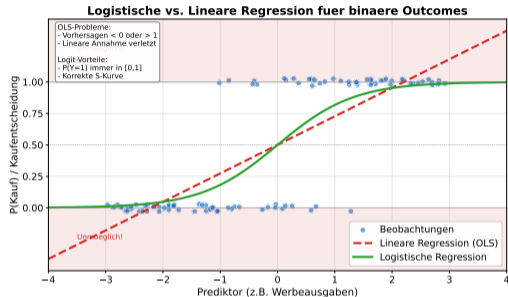
- Modelliert **Wahrscheinlichkeit**: $P(Y = 1|X)$
- Werte immer zwischen 0 und 1
- S-foermige Kurve (Sigmoid)

Mathematisch:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X_1 + \dots$$

- Linke Seite: Log-Odds
- Rechte Seite: Linear in X

Logistische Regression ist der Standard fuer binäre Klassifikation - interpretierbar und robust



Anwendungen:

- Churn-Vorhersage (Kunde bleibt/geht)
- Click-Prediction (klickt/klickt nicht)
- Kreditrisiko (Default ja/nein)
- Medizin (krank/gesund)

Von Koeffizienten zu Odds Ratios

Interpretation von β

- β = Veraenderung in Log-Odds
- Schwer direkt zu interpretieren
- Besser: **Odds Ratio (OR)**

Odds Ratio berechnen:

$$OR = e^{\beta}$$

Interpretation:

- $OR = 1$: Kein Effekt
- $OR > 1$: Erhoehte Chance ($Y = 1$)
- $OR < 1$: Verringerte Chance
- $OR = 2$: Verdopplung der Odds

Beispiel in R:

```
# Daten: Click (0/1), AdSpend
model <- glm(Click ~ AdSpend,
             data = data,
             family = binomial)
```

```
# Koeffizienten
coef(model)
# (Intercept)      AdSpend
#      -2.3         0.5
```

```
# Odds Ratios
exp(coef(model))
# (Intercept)      AdSpend
#      0.10         1.65
```

Interpretation:

- Pro 1 Euro mehr AdSpend: 65% hoehere Odds fuer Click