

# Hypothesen, Fehlerarten, Teststärke und p-Werte

statistik-online.ch

Bei einem **Hypothesentest** wird mit Stichprobendaten entschieden, ob eine Aussage über die Population plausibel ist. Dabei können Fehler entstehen. Dieses Dokument erklärt die wichtigsten Begriffe:  $H_0$ ,  $H_1$ , Signifikanzniveau, Alpha-Fehler, Beta-Fehler, Teststärke und p-Wert.

## 1 Grundidee eines Hypothesentests

**Ein Hypothesentest vergleicht eine beobachtete Stichprobe mit einer Annahme über die Population.**

Die Annahme, die zuerst gilt, ist die Nullhypothese. Nur wenn die Daten stark genug gegen diese Annahme sprechen, wird sie verworfen.

Begriff	Bedeutung
$H_0$	Nullhypothese: Standardannahme, zum Beispiel “kein Unterschied” oder “kein Zusammenhang”.
$H_1$	Alternativhypothese: Aussage, die man zeigen oder prüfen möchte.
$\alpha$	Signifikanzniveau: vorher festgelegte Grenze für den Alpha-Fehler.

### Beispiel

Eine Schule möchte prüfen, ob ein neues Lernprogramm die durchschnittliche Punktzahl von 70 Punkten erhöht.

$$H_0 : \mu = 70$$

$$H_1 : \mu > 70$$

**Inhaltlich:**

- $H_0$ : Das Lernprogramm erhöht die durchschnittliche Punktzahl nicht.
- $H_1$ : Das Lernprogramm erhöht die durchschnittliche Punktzahl.

## 2 Entscheidung und Fehler

Bei einem Test gibt es zwei mögliche Entscheidungen:

- $H_0$  wird verworfen. ( $H_1$  wird angenommen.)
- $H_0$  wird beibehalten. ( $H_1$  wird verworfen.)

Gleichzeitig kann in Wirklichkeit entweder  $H_0$  wahr sein oder  $H_1$  wahr sein. Daraus entstehen vier Fälle:

	$H_0$ ist wahr	$H_1$ ist wahr
$H_0$ verwerfen	Alpha-Fehler	richtige Entscheidung
$H_0$ beibehalten	richtige Entscheidung	Beta-Fehler

## 3 Alpha-Fehler

**Ein Alpha-Fehler bedeutet: Man verwirft  $H_0$ , obwohl  $H_0$  in Wirklichkeit wahr ist.**

Das Signifikanzniveau  $\alpha$  legt fest, wie gross dieses Risiko maximal sein soll. Häufig wird verwendet:

$$\alpha = 0.05$$

Das bedeutet: Unter der Bedingung, dass  $H_0$  wahr ist, beträgt die Wahrscheinlichkeit,  $H_0$  fälschlicherweise zu verwerfen, 5%.

### Beispiel

Ein Medikament hat in Wirklichkeit keine Wirkung. Ein Test kommt aber zum Schluss, dass das Medikament wirkt. Dann wurde ein Alpha-Fehler gemacht.

## 4 Beta-Fehler

**Ein Beta-Fehler bedeutet: Man behält  $H_0$  bei, obwohl  $H_1$  in Wirklichkeit wahr ist.**

Der Beta-Fehler wird mit  $\beta$  bezeichnet. Er ist das Risiko, einen tatsächlichen Effekt zu übersehen.

### Beispiel

Ein Medikament wirkt in Wirklichkeit. Ein Test erkennt diese Wirkung aber nicht. Dann wurde ein Beta-Fehler gemacht.

## 5 Teststärke

Die Teststärke ist die Wahrscheinlichkeit, einen tatsächlichen Effekt zu erkennen.

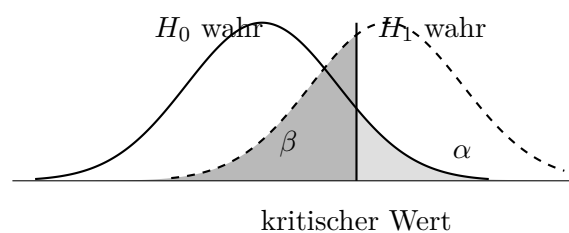
Die Teststärke wird auch **Power** genannt.

$$\text{Teststärke} = 1 - \beta$$

Wenn  $\beta = 0.20$ , dann gilt:

$$\text{Teststärke} = 1 - 0.20 = 0.80$$

Das bedeutet: Wenn der Effekt tatsächlich existiert, wird er mit einer Wahrscheinlichkeit von 80% erkannt.



### Was erhöht die Teststärke?

- eine grössere Stichprobe
- ein grösserer tatsächlicher Effekt
- kleinere Streuung in den Daten
- ein höheres Signifikanzniveau  $\alpha$

**Wichtig:** Ein höheres  $\alpha$  kann die Teststärke erhöhen, aber gleichzeitig steigt auch das Risiko für einen Alpha-Fehler.

## 6 p-Wert

Der p-Wert ist die Wahrscheinlichkeit, unter der Annahme einer wahren Nullhypothese ein mindestens so grosses Ergebnis wie das beobachtete zu erhalten.

$$p = P(\text{mindestens so grosses Ergebnis wie beobachtet} \mid H_0 \text{ ist wahr})$$

Kleine p-Werte sprechen gegen  $H_0$ , da es unwahrscheinlich ist, dass man den gefundenen Effekt erhalten würde, falls  $H_0$  in Wirklichkeit gilt. Grosse p-Werte bedeuten nicht, dass  $H_0$  bewiesen ist. Sie bedeuten nur, dass die Daten nicht stark genug gegen  $H_0$  sprechen.

### Entscheidungsregel mit p-Wert

$$p \leq \alpha \implies H_0 \text{ verwerfen}$$

$$p > \alpha \implies H_0 \text{ beibehalten}$$

## Beispiel

Es wird mit  $\alpha = 0.05$  getestet. Der Test ergibt:

$$p = 0.032$$

Da  $0.032 \leq 0.05$ , wird  $H_0$  verworfen.

Wenn stattdessen  $p = 0.18$  wäre, würde  $H_0$  beibehalten, weil  $0.18 > 0.05$ .

## 7 Häufige Missverständnisse

- Ein p-Wert ist **nicht** die Wahrscheinlichkeit, dass  $H_0$  wahr ist, sondern die Wahrscheinlichkeit, unter  $H_0$  ein mindestens so grosses Ergebnis wie das beobachtete zu erhalten.
- Ein nicht signifikanter Test beweist **nicht**, dass es keinen Effekt gibt, sondern bedeutet, dass die Daten nicht stark genug gegen  $H_0$  sprechen.
- $\alpha = 0.05$  bedeutet **nicht**, dass der konkrete Befund mit 95% Wahrscheinlichkeit richtig ist, sondern dass man unter einer wahren Nullhypothese ein Risiko von 5% für einen Alpha-Fehler akzeptiert.
- Teststärke bezieht sich auf das Erkennen eines tatsächlichen Effekts.

## 8 Zusammenfassung

Begriff	Bedeutung
$H_0$	Nullhypothese, die nur bei genügend starker Evidenz verworfen wird.
$H_1$	Alternativhypothese, die gegen $H_0$ steht.
$\alpha$	Signifikanzniveau; maximal akzeptiertes Risiko für einen Alpha-Fehler.
Alpha-Fehler	$H_0$ wird verworfen, obwohl $H_0$ wahr ist.
Beta-Fehler	$H_0$ wird beibehalten, obwohl $H_1$ wahr ist.
Teststärke	Wahrscheinlichkeit, einen tatsächlichen Effekt zu erkennen; $1 - \beta$ .
p-Wert	Bedingte Wahrscheinlichkeit für ein mindestens so grosses Ergebnis, unter der Annahme, dass $H_0$ wahr ist.

## 9 Aufgaben

### Aufgabe 1: Hypothesen formulieren

Eine Firma behauptet, dass ihre Batterien im Durchschnitt länger als 10 Stunden halten. Formulieren Sie  $H_0$  und  $H_1$ .

### Aufgabe 2: Fehlerart bestimmen

Ein neues Training verbessert die Leistung in Wirklichkeit nicht. Eine Studie kommt aber zum Schluss, dass das Training wirksam ist. Welche Fehlerart liegt vor?

### Aufgabe 3: Teststärke berechnen

Ein Test hat einen Beta-Fehler von  $\beta = 0.15$ . Wie gross ist die Teststärke?

### Aufgabe 4: p-Wert interpretieren

Es wird mit  $\alpha = 0.05$  getestet. Der p-Wert beträgt  $p = 0.08$ . Wird  $H_0$  verworfen?

### Aufgabe 5: Multiple Choice

Kreuzen Sie alle richtigen Aussagen an.

- Ein Alpha-Fehler bedeutet, dass  $H_0$  fälsch verworfen wird.
- Die Teststärke ist  $1 - \beta$ .
- Ein p-Wert von 0.03 bedeutet, dass  $H_0$  mit 3% Wahrscheinlichkeit wahr ist.
- Eine grössere Stichprobe kann die Teststärke erhöhen.

## 10 Lösungen

### Lösung 1

$$H_0 : \mu = 10$$

$$H_1 : \mu > 10$$

**Interpretation:**  $H_0$  sagt, dass die Batterien im Durchschnitt nicht länger als der Referenzwert getestet werden.  $H_1$  entspricht der Behauptung der Firma.

### Lösung 2

Es liegt ein **Alpha-Fehler** vor.

Begründung: Das Training wirkt in Wirklichkeit nicht. Trotzdem wird ein Effekt angenommen.  $H_0$  wird also fälsch verworfen.

### Lösung 3

$$\text{Teststärke} = 1 - \beta = 1 - 0.15 = 0.85$$

**Antwort:** Die Teststärke beträgt 85%.

### Lösung 4

$$p = 0.08 \quad \alpha = 0.05$$

Da  $0.08 > 0.05$ , wird  $H_0$  beibehalten.

### Lösung 5

- ✓ Richtig – Ein Alpha-Fehler ist das falsche Verwerfen von  $H_0$ .
- ✓ Richtig – Die Teststärke ist  $1 - \beta$ .
- ✗ Falsch – Der p-Wert beschreibt ein Ergebnis unter der Annahme, dass  $H_0$  wahr ist.
- ✓ Richtig – Eine grössere Stichprobe kann die Teststärke erhöhen.