# 19. Vorlesung: Hypothesentests

Nikolas Tapia

24. Juni 2024, Stochastik für Informatik(er)

# Grundprinzipen

Vor der Messung der Daten wird eine Hypothese (**Nullhypothese** $H_0$) aufgestellt, die dann überprüft wird.

Beispielweise:

1. Die Daten folgen eine Normalverteilung.
2. Der Mittelwert der Daten ist 0.
3. Die Daten sind unabhängig.

Noch konkreter:

1. Vergleich der Wirkung von zwei Medikamenten.
2. Überprüfung der Qualität von Produkten, usw.

# Fehler 1. und 2. Art

## Definition 19.1

Zwei Typen von “Fehler” sind möglich:

1. **Fehler 1. Art**: $H_0$ wird abgelehnt, obwohl sie wahr ist.
2. **Fehler 2. Art**: $H_0$ wird angenommen, obwohl sie falsch ist.

## Definition 19.2

Das **Signifikanzniveau** $\alpha \in (0,1)$ ist der maximal zulässige Fehler 1. Art, d.h.

$$\mathbb{P}(H_0 \text{ wird abgelehnt} \mid H_0 \text{ gilt}) \leq \alpha.$$

# Vorgehen

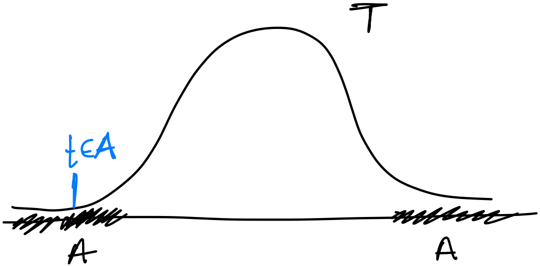
1. Nullhypothese aufstellen.
   - Behauptung, die mittels Daten geprüft werden kann.
2. Teststatistik $T$ aufstellen.
   - Die Verteilung von $T$ unter $H_0$ ist bekannt, z.B. $T \sim \mathcal{N}(0,1)$ oder $T \sim t_n$.
3. Signifikanzniveau $\alpha \in (0,1)$ wählen.
   - Maximal zulässige Fehler 1. Art.
   - Üblicherweise $\alpha = 0.05$ oder $\alpha = 0.01$.
4. Experiment durchführen und Daten erhalten.
   - Messwerte $x_1, \dots, x_n$ ergeben einen Testwert $t$ für die Teststatistik.
5. $H_0$ wird entweder abgelehnt oder nicht.
   - Neyman-Pearson-Entscheidungsregel: $H_0$ wird abgelehnt, wenn $t$ im Ablehnungsbereich liegt. Sonst ist $H_0$ nicht abgelehnt.

# Hypothesentest

**Definition 19.3**

Ein **Ablehnungsbereich** (oder **kritischer Bereich**) zum Fehlerniveau $\alpha$ für $H_0$ ist eine Teilmenge $A \subseteq \mathbb{R}$, sodass

$$\underbrace{\mathbb{P}(T \in A \mid H_0)}^{\text{Fehler 1. Art.}} \leq \alpha.$$



# Testarten

1. **原假设** ($H_0$)：这是初始假设，通常表示没有效应或没有差异的情况。它是研究者在检验开始时假定为真的假设。在进行统计检验时，我们试图通过数据来验证这个假设是否成立。例如，在一个单侧检验中，原假设可能是一个参数小于或等于某个特定值 (如 $H_0 : \theta \leq \theta_0$)，在双侧检验中，原假设可能是参数等于某个特定值 (如 $H_0 : \theta = \theta_0$)。
2. **备择假设** ($H_1$)：这是与原假设相对立的假设，通常表示存在效应或存在差异的情况。如果数据提供了足够的证据使我们怀疑原假设不成立，我们就会拒绝原假设，转而接受备择假设。例如，在一个单侧检验中，备择假设可能是一个参数大于某个特定值 (如 $H_1 : \theta > \theta_0$)，在双侧检验中，备择假设可能是参数不等于某个特定值 (如 $H_1 : \theta \neq \theta_0$)。

## Definition 19.4

Sei $(\mathbb{P}_\theta)_{\theta\in\Theta}$ ein statisches Modell für eine Grundgesamtheit.

- Ein **einseitiges Test** ist ein Test, bei dem die Nullhypothese $H_0$ in der Form

$$H_0 : \theta \leq \theta_0, \quad H_1 : \theta > \theta_0$$

oder

$$H_0 : \theta \geq \theta_0, \quad H_1 : \theta < \theta_0$$

aufgestellt wird.

- Ein **beidseitiges Test** ist ein Test, bei dem die Nullhypothese $H_0$ in der Form

$$H_0 : \theta = \theta_0, \quad H_1 : \theta \neq \theta_0$$

$(p = 1/2) \qquad (p \neq 1/2)$

$\left( \begin{array}{l} H_0' : \theta \neq \theta_0 \\ H_1' : \theta = \theta_0 \end{array} \right)$

aufgestellt wird.



Leibniz Gemeinschaft

# Ablehnungsbereich

## Aussage 19.1

Sei $T$ eine Teststatistik, $t$ den entsprechenden Testwert, und $H_0$ eine Nullhypothese.

- Bei einem einseitigen Test ist der Ablehnungsbereich der Form $A_\alpha = \{t \mid t \leq k\}$, bzw. $A_\alpha = \{t \mid t \geq k\}$, wobei $k$ so gewählt wird, dass $\mathbb{P}(\underbrace{T \geq k}_{T \in A} \mid H_0) \leq \alpha$ bzw. $F_T(k) = \mathbb{P}(T \leq k \mid H_0) \leq \alpha$.

  $k$ = Quantil zum Niveau $\alpha$.

- Bei einem beidseitigen Test ist der Ablehnungsbereich der Form

$$A_\alpha = \{t \mid t \geq k_1\} \cup \{t \mid t \leq k_2\}, \qquad (k_1 > k_2)$$

wobei $k_1$ und $k_2$ so gewählt werden, dass $\mathbb{P}(T \geq k_1 \mid H_0) \leq \alpha/2$ und $\mathbb{P}(T \leq k_2 \mid H_0) \leq \alpha/2$.

$$\mathbb{P}(T \in A \mid H_0) = \mathbb{P}(T \geq k_1 \mid H_0) + \mathbb{P}(T \leq k_2) \leq \frac{\alpha}{2} + \frac{\alpha}{2} \leq \alpha.$$

# $p$-Wert

## Definition 19.5

Sei $T$ eine Teststatistik, $t$ den entsprechenden Testwert, und $H_0$ eine Nullhypothese.

1. Der **rechtsseitige $p$-Wert** ist
$$p_r = \mathbb{P}(T \geq t \mid H_0).$$
2. Der **linksseitige $p$-Wert** ist
$$p_l = \mathbb{P}(T \leq t \mid H_0).$$
3. Der **beidseitige $p$-Wert** ist
$$p_b = 2\min(p_r, p_l).$$

Bei linksseitige H_0

$p < \alpha \iff H_0$ abgelehnt

$\iff t \leq k$

$\iff t \in A$

## Aussage 19.2

Sei $T$ eine Teststatistik, $t$ den entsprechenden Testwert, und $H_0$ eine Nullhypothese.

- Bei einem einseitigen Test ist $H_0$ abgelehnt, wenn $p_r \leq \alpha$ bzw. $p_l \leq \alpha$.
- Bei einem beidseitigen Test ist $H_0$ abgelehnt, wenn $p_b \leq \alpha$.

Beispiel

Eine Münze wird $n$-mal geworfen. Dabei erhalten wir Daten $x_1, \ldots, x_n \in \{0, 1\}^n$. Die Nullhypothese ist, dass die Münze fair ist. Die Alternativhypothese ist, dass die Münze nicht fair ist. Unter $H_0$ ist die Teststatistik

$$T = \sum_{i=1}^{n} X_i \sim \text{Binomial}\left(n, \frac{1}{2}\right).$$

$H_0$ wird abgelehnt, wenn der Testwert $t$ im Ablehnungsbereich $\{t \leq k_1\} \cup \{t \geq k_2\}$ liegt. Äquivalent wenn der $p$-Wert $p_b$ kleiner als $\alpha$ ist.

H_0: p = 1/2
H_1: p ≠ 1/2