Übersicht

- liefert Entscheidungsregel, für welche Stichproben wird
  - H_0 akzeptiert
  - H_0 zugunsten von H_1 verworfen
- kritische Bereich
  - $$K \subset \mathbb{R}^n$$
  - Menge der Stichproben
  - für die H_0 verworfen wird
- Beispiele
  - Münzwurf
    - Es ist intuitiv nachvollziehbar, die Nullhypothese

      $$\mathcal{H}_0 : \theta \leq 0.5$$

      zugunsten von

      $$\mathcal{H}_1 : \theta > 0.5$$

      zu verwerfen, wenn „zu oft“ Kopf ($x_i = 1$) beobachtet wird. Damit lautet der kritische Bereich

      $$K = \Big\{(x_1, \ldots, x_n) \in \{0,1\}^n : x_1 + \cdots + x_n \geq k\Big\},$$

      für ein $k \geq 0$.
  - Medikament
    - In diesem Fall ist es intuitiv, die Nullhypothese

      $$\mathcal{H}_0 : \mu = 0$$

      zugunsten von

      $$\mathcal{H}_1 : \mu \neq 0$$

      zu verwerfen, falls das empirische Mittel „weit“ von 0 abweicht. Dieser Fall wäre ein Anzeichen dafür, dass das Medikament eine Wirkung hat. Der entsprechende kritische Bereich lautet

      $$K = \Big\{(d_1, \ldots, d_n) \in \mathbb{R}^n : |\bar{d}| \geq k\Big\},$$

      für ein $k > 0$.

Ausgänge

	$\mathcal{H}_0$ akzeptiert	$\mathcal{H}_0$ verworfen
$\mathcal{H}_0$ wahr	korrekte Entscheidung	Fehler 1. Art
$\mathcal{H}_0$ falsch	Fehler 2. Art	korrekte Entscheidung

- **Fehler 1. Art**: $\mathcal{H}_0$ wird verworfen, obwohl $\mathcal{H}_0$ wahr ist.
- **Fehler 2. Art**: $\mathcal{H}_0$ wird akzeptiert, obwohl $\mathcal{H}_0$ falsch ist.
- Level-α-Test
  - α beschränkt Fehler 1. Art
  - Ein Test hat **Signifikanzniveau** $\alpha \in [0, 1]$, falls

    $$P_\theta(X \in K) \leq \alpha \qquad \forall\, \theta \in \Theta_0.$$
  - [[p-Wert]]
- Beispiele
  - Medikament
    * Wie zuvor sei D_i die Differenz im Blutdruck einer Person i vor und nach Behandlung mit einem speziellen Medikament.

      Zum Nachweis der Wirksamkeit des Medikaments wird ein statistischer Test durchgeführt. Dafür soll nachgewiesen werden, dass

      $$\mu = \mathrm{E}(D_i) > 0.$$

      Gelingt der Nachweis, wird das Medikament auf den Markt gebracht.

1. Ansatz: Man setzt

$$\mathcal{H}_0 : \mu > 0$$

und bringt das Medikament auf den Markt, wenn $\mathcal{H}_0$ von einem geeigneten Test *nicht* verworfen wird.

2. Ansatz: Man setzt

$$\mathcal{H}_0 : \mu \leq 0$$

und bringt das Medikament auf den Markt, wenn $\mathcal{H}_0$ von einem geeigneten Test verworfen wird.

* Die beiden Ansätze sind *nicht* äquivalent!

  Im **1. Ansatz** wird dem Medikament fälschlicherweise eine positive Wirkung zugeschrieben, falls ein *Fehler 2. Art* begangen wird.

  Im **2. Ansatz** wird dem Medikament fälschlicherweise eine positive Wirkung zugeschrieben, falls ein *Fehler 1. Art* begangen wird.

  Mit dem Signifikanzniveau ist allerdings nur der Fehler 1. Art unter Kontrolle. Daher ist der 2. Ansatz zu bevorzugen.

  Salopp formuliert: Verwenden Sie das Gegenteil von dem was gezeigt werden soll als Nullhypothese.

*

– Münzwurf

  Wir betrachten wieder n-unabhängige Münzwürfe und zweifeln daran, dass die Münze fair ist. Deshalb testen wir

$$\mathcal{H}_0 : \quad \theta = 0.5 \qquad \text{vs} \qquad \mathcal{H}_1 : \quad \theta \neq 0.5.$$

  Wird $\mathcal{H}_0$ durch einen Level–α–Test verworfen, besteht

$$100 \cdot (1 - \alpha)\%$$

  Gewissheit, dass die Münze unfair ist.

* Die Wahrscheinlichkeit eine richtige Nullhypothese zu verwerfen (also einen Fehler 1. Art zu begehen) ist durch α beschränkt.

Vorgehensweise

- 1. Formulieren der Null- bzw. Alternativhypothese.

  2. Wahl des Tests bzw. der Teststatistik.

  3. Festlegen des Signifikanzniveaus α und kritischen Bereiches.

  4. Berechnungen anhand der Stichprobe.

  5. Entscheidung ob die Nullhypothese verworfen wird.

Einstichprobenproblem

- statistische Tests um Eigenschaften einer [[Stichprobe]] von [[Zufallsvariable]]n zu studieren
- Konkret testen wir, ob der Erwartungswert normalverteilter Daten ungleich (oder größer bzw. kleiner) einem gegebenen Wert ist.
- Unterscheidung in zwei Fälle
  - [[Gauß-Test]], wenn Varianz σ^2 bekannt
  - [[t-Test]], wenn Varianz σ^2 unbekannt

Zweistichprobenproblem

Nun vergleichen wir zwei Verteilungen miteinander. Angenommen,

$$X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathrm{N}(\mu_X, \sigma_X^2), \qquad Y_1, \ldots, Y_m \overset{\text{iid}}{\sim} \mathrm{N}(\mu_Y, \sigma_Y^2).$$

- Außerdem seien $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ unabhängig.
- Untersuchung von Hypothesenpaar für $\omega_0 \in \mathbb{R}$
  - $$\mathcal{H}_0: \quad \mu_X - \mu_Y = \omega_0 \qquad \text{vs} \qquad \mathcal{H}_1: \quad \mu_X - \mu_Y \neq \omega_0,$$
- Unterscheidung in zwei Fälle
  - [[Zweistichproben-Gauß-Test]], wenn Varianzen bekannt
  - [[Zweistichproben-t-Test]], wenn Varianzen unbekannt

[[Testen]]