Vorlesung 13

Inhalt

- Grundprinzipien von statistischen Tests
- Fehler 1. und 2. Art
- $p-$Werte

Lernziele

- Die Grundprinzipien von statistischen Tests kennen
- Fehler 1. und 2. Art kennen
- Die Rolle von p-Werten in der Statistik kenne
- Ein Beispiel für einen $t-$Test kennen

Benötigte Vorkenntnisse

- Wahrscheinlichkeiten, Verteilungen, Mittelwert, Varianz

Testen von Hypothesen

Vor dem Experiment: Aufstellen einer Hypothese (=Nullhypothese) betreffend der zu messenden Daten, z.B.

- Die Daten folgen einer Normalverteilung
- Der Erwartungswert ist höchstens 10.5
- Die Daten sind unkorreliert

Test: Nach Messung der Daten wird durch Rechnung überprüft, ob die gemessenen Daten der Hypothese widersprechen, d.h. ob unter der Annahme, dass die Hypothese gilt, die Daten sehr unwahrscheinlich sind.

Falls die Daten (unter der Hypothese) sehr unwahrscheinlich sind, wird die Hypothese verworfen. Andernfalls wird die Hypothese nicht verworfen (oder angenommen).

Testen von Hypothesen

Grundprinzip: Falls unter der Nullhypothese die beobachteten Daten sehr unwahrscheinlich sind, so verwirft man die Nullhypothese.

Was bedeutet "Daten sind sehr unwahrscheinlich?"

- Z.B. p-Wert
- Z.B. Konfidenzintervall
- etwas mehr dazu später

Wichtig: Dass die Nullhypothese angenommen wird, bedeutet nicht, dass sie wahr ist - und umgekehrt. Sie wird dann angenommen,wenn das Gegenteil sehr unwahrscheinlich wäre.

Fehler 1. und 2. Art

Bei einem statistischen Test sind zwei Typen von “Fehlern” möglich:

- Der Fehler 1. Art: Die Nullhypothese wird verworfen, obwohl sie eigentlich richtig wäre
- Der Fehler 2. Art: Die Nullhypothese wird angenommen, obwohl sie eigentlich falsch ist

Wichtig: (Nur) der Fehler 1. Art kann in der Konstruktion von statistischen Tests kontrolliert werden. Man kann sich ein Fehlerniveau $\alpha \in [0, 1]$ vorgeben, und dazu einen Annahmebereich konstruieren, so dass die Wahrscheinlichkeit, die Nullhypothese abzulehnen, obwohl sie eigentlich wahr ist, kleiner oder gleich α ist.

- Fehler 2. Art im Allgemeinen nicht explizit quantifizierbar
- Je kleiner der Fehler 1. Art, desto größer der Fehler 2. Art, und umgekehrt
- Wahl der Nullhypothese wichtig!

p-Wert

Seien $x_1, ..., x_n$ Messwerte. Der p-Wert zu diesen Daten bezeichnet die Wahrscheinlichkeit, dass (unter der Nullhypothese) Werte "gleich oder extremer als die Messwerte $x_1, ..., x_n$" gemessen werden.

Formal

$$p = p(x_1, ..., x_n) = \mathbb{P}(X \geq x_i, i = 1, ..., n \,|\, H_0)$$

- Ist der p-Wert zu den gemessenen Daten $p = p(x_1, ..., x_n)$ groß: Daten $(x_1, ..., x_n)$ sind "typische" (und nicht extreme) Ergebnisse
- Ist der p-Wert klein, so sind die gemessenen Daten unwahrscheinlich.
- Bei kleinem $p-$Wert wird man also die Nullhypothese ablehnen, bei großem $p-$Wert wird man sie annehmen.

p-Wert, Signifikanz

Was bedeutet groß oder klein beim $p-$Wert?

$p = 0.1$ bedeutet: (nur) 10% aller möglicher Messungen sind extremer als das, was wir gemessen haben. Unsere Daten sind also, falls die Nullhypothese wahr ist, eher unwahrscheinlich.

Annahmebereiche für statistische Tests werden oft so konstruiert, dass die Nullhypothese für $p \geq \alpha$ angenommen wird.

Dabei spricht man oft von einem signifikanten Ergebnis, wenn $p \leq 0.05$ ist, die Wahrscheinlichkeit, einen Fehler 1. Art zu begehen, also kleiner als 0.05 ist.

Zusammenfassung

- Kleiner $p-$Wert ($p \leq \alpha$): Das Ergebnis ist unter der Nullhypothese sehr unwahrscheinlich
- Deshalb wird die Nullhypothese abgelehnt. Der Fehler erster Art ist dabei kleiner als α (also mit kleiner Wahrscheinlichkeit ist die Nullhypothese doch wahr)
- Für $p \leq 0.05$ spricht man von einem signifikanten Ergebnis.

Beachte: $p \leq \alpha = 0.05$ bedeutet, dass die Nullhypothese abgelehnt wird!

Deshalb ist die Wahl der Nullhypothese wichtig. Bei Wirksamkeitstests wählt man deshalb als Nullhypothese die Annahme, dass keine Wirksamkeit vorliegt. Falls diese Nullhypothese dann zum Fehlerniveau $\alpha = 0.05$ abgelehnt wird, spricht man dann von einem signifikanten Ergebnis, man erwartet also eine Wirksamkeit (!).

Testen von Hypothesen

Vorgehen: Situation: $x_1, \ldots, x_n$ Messwerte, Grundannahmen festhalten.

1. Formuliere die Nullhypothese H_0. (Bsp).
2. Konstruiere den Annahmebereich (bzw. Ablehnungsbereich).
3. Führe das Experiment durch.
4. Überprüfe, ob die Daten im Annahmebereich liegen.

Ergebnis von 4.:

- Falls nein: Nullhypothese verwerfen
- Falls ja: Nullhypothese nicht verwerfen, bzw. annehmen

Falls Nullhypothese verworfen: Alternativhypothese annehmen.

Konstruktion von Annahmebereichen mittels p-Werten oder Quantilen.

$t-$Test

Der t-Test (einseitig für ein Merkmal) ist ein konkreter Test, welcher dem allgemeinen Vorgehen folgt und testet, ob die Daten der Annahme, dass der Erwartungswert (Mittelwert) einer Größe gleich (kleiner gleich, größer gleich) einer gegebenen Zahl μ_0 ist, widersprechen oder nicht.

Gegeben: Daten $x_1, ..., x_n$

Annahme: Daten entsprechen unabhängigen Messungen der gleichen Größe mit endlicher Varianz.

Nullhypothese: Der Mittelwert/Erwartungswert ist $\leq \mu_0$.

Die Wahl der Nullhypothese ist dabei wichtig!

$t-$Test: Beispiel

Beispiel: Wirkung eines fiebersenkenden Medikamentes

Nr. Patient	Temp. Einnahme	Temp. 3 h später	Differenz
i	u_i	v_i	$x_i = u_i - v_i$
1	39.1	38.7	0.4
2	38.3	38.1	0.2
3	37.6	37.9	-0.3
4	38.0	37.5	0.5
5	40.1	39.2	0.9
6	39.5	39.1	0.4
7	38.7	38.7	0.0
8	37.9	37.5	0.4
9	39.2	38.2	1.0
10	38.0	37.4	0.6

Nullhypothese: Das Medikament hat keine fiebersenkende Wirkung, der Mittelwert von x_i ist ≤ 0. (!!!)

Ein Merkmal, einseitiger t-Test: Vorgehen

1. Stelle die Nullhypothese auf: H_0: $\mu \leq \mu_0$
2. Wähle Fehlerniveau $\alpha \in (0,1)$
3. Erhebe die Daten $x_1, \ldots, x_n$.
4. Berechne $\bar{\mu}_n$ und $\bar{\sigma}_n = \sqrt{\bar{\sigma}_n^2}$ aus den Daten
5. Berechne den Testwert

$$t = \sqrt{n} \cdot \frac{\bar{\mu}_n - \mu_0}{\bar{\sigma}_n}$$

6. Berechne den p-Wert $p(t) = \mathbb{P}(T \geq t)$ zum Testwert t unter Verwendung der $t-$Verteilung zum Parameter $f = 1 - n$.
   Entscheidungsregel:
   - Ist $p(t) \leq \alpha$ so wird H_0 abgelehnt
   - Andernfalls wird H_0 nicht abgelehnt (angenommen).

   Falls $p(t) \leq \alpha$ ist, betrachten wir also die Wirkung des Medikaments als signifikant!

t-Verteilung

Eine Wahrscheinlichkeitsverteilung, welche im $t-$Test wichtig ist. Sie ist der Normalverteillung relativ ähnlich.

Ihre Werte und Quantile werden in Tabellen aufgelistet, bzw. können z.B. in Excel aufgerufen werden:
=TVERT(t, n-1,1) gibt in der Situation des obigen Beispiels den gesuchten $p-$Wert $p(t) = \mathbb{P}(T \geq t)$ an.

Weiterführende Literatur

- Statistik für Mediziner und Pharmazeuten (Philip Rowe, Wiley)
- Mathematik in den Life Sciences (Gerhard Keller, Ullmer)
- Einführung in die mathematische Behandlung der Naturwissenschaften (Hans Heiner Storrer, Birkhäuser)
- Stochastik für Informatiker (Noemi Kurt, Springer Vieweg)
- Mehr finden Sie in Ihrer Fachbibliothek!

Dringend empfohlen: Statistik für Biowissenschaften, VL von Prof. Dr. Gaby Schneider

Und als Bettlektüre: Mai Thi Nguyen-Kim: Die kleinste gemeinsame Wirklichkeit (Droemer).