

Empirische Methoden zur Analyse gesprochener Sprache

Prinzip der Hypothesenprüfung

(am Beispiel des t-Tests für unabhängige
Stichproben)

Einordnung

Statistische Verfahren:

- Deskriptive (beschreibende) Statistik: Stichprobenbeschreibung (bzw. Beschreibung der gesamten Population)
 - Kennwerte
 - Diagramme
- Inferenz- (schließende) Statistik
 - Populationsbeschreibende Untersuchungen (Abschätzung von Populationsparametern durch Untersuchung einer Stichprobe)
 - Hypothesenerkundende Untersuchungen
 - **Hypothesenprüfende Untersuchungen**

Schritte der Hypothesenprüfung

- Formulierung von Hypothesen
- Festlegung der Stichprobenart
- Operationalisierung der Variablen
- Festlegung des Fehlerniveaus
- Auswahl des statistischen Tests
- Datenerhebung
- Berechnung des (Stichproben-) Kennwerts
- Berechnung des Testwerts
- Abgleich mit Testwertverteilung: Bestimmung der Fehlerwahrscheinlichkeit
- Entscheidung: Vergleich mit Fehlerniveau
- falls signifikant: Interpretation der Effektgröße

Hypothesen

Im Unterschied zu den zuvor genannten Untersuchungstypen, erfordert die Hypothesenprüfung

- **Vorwissen** (wenigstens theoriegeleitet) über den zu untersuchenden Gegenstand,
- welches als Annahmen über den Untersuchungsausgang in einem **Hypothesenpaar** formuliert wird
 - Nullhypothese (H_0)
 - Alternativhypothese (H_1)

Hypothesen

Annahmen über den Untersuchungsausgang können prinzipiell

- einen/keinen Zusammenhang (zwischen mindestens zwei Merkmalen)
- einen/keinen (Mittelwerts-) Unterschied bzgl. mindestens eines Merkmals zwischen (mindestens zwei) Untersuchungs-objektgruppen (bzw. auch Messzeitpunkten, Messverfahren, ...)

postulieren

Hypothesen

aus diesen Annahmen abgeleitete Hypothesen **müssen**

- sich gegenseitig ausschließen
- alle möglichen Ausgänge der Untersuchung abdecken

und **sollten** so genau wie möglich formuliert sein

Hypothesen

- Die **Forschungshypothese** postuliert mit Hilfe (klar definierter) theoretischer Konstrukte den Zusammenhang bzw. Unterschied (in der Population), den der Forscher bestätigen will, bzw. der Anlass für die Untersuchung ist.
- Die **operationale Hypothese** verfeinert dieses Postulat um die Angabe der Operationalisierung der Variablen.
- Die **statistische Hypothese** ist ein formalisiertes Postulat (z.B. $H_1: \mu_1 \neq \mu_2$)

Hypothesen

- **ungerichtete (Alternativ-) Hypothesen** sind Hypothesen, die die Richtung des postulierten Effekts (Zusammenhang oder Unterschied) nicht bestimmen; sie sind (relativ) informationsarm (z.B. $H_1: \mu_1 \neq \mu_2$)
- **gerichtete (Alternativ-) Hypothesen** sind Hypothesen, die die Richtung des postulierten Effekts bestimmen; sie sind (relativ) informationsreich (z.B. $H_1: \mu_1 > \mu_2$)

Hypothesen

im Falle einer gerichteten Formulierung der Hypothesen können spezifische und unspezifische Hypothesen unterschieden werden:

- **unspezifische Hypothesen** sind Hypothesen, bei denen die (minimal bedeutsame) Effektgröße nicht postuliert wird; (relativ) informationsarm
- **spezifische Hypothesen** sind Hypothesen, bei denen die (minimal bedeutsame) Effektgröße spezifiziert wird; (relativ) informationsreich

Fehlerarten

| | | in der Population gilt die | |
|--|----|--|---|
| | | H0 | H1 |
| Entscheidung aufgrund der Stichprobe zugunsten der | H0 | richtige Entscheidung | β -Fehler (F. zweiter Art; Type II Error) |
| | H1 | α -Fehler (F. erster Art; Type I Error) | richtige Entscheidung |

Test-Algorithmus und Fehlerabschätzung

- Forschungshypothese sollte der Alternativhypothese entsprechen (), da...
- aus mathematischen Gründen (Stichprobenkennwerteverteilung im Falle des Zutreffens der H_1 unbekannt) **kann** in der Regel (sofern keine minimale Effektgröße bekannt ist, also die H_1 unspezifisch formuliert ist) **nur der α -Fehler berechnet werden**, aber nicht der β -Fehler

Signifikanzniveau

[auch α -Fehler-Niveau (α)]

- legt fest, **ab** welcher α -Fehler-Wahrscheinlichkeit (p) die H_0 zugunsten der H_1 verworfen wird ($\alpha \geq p$)
- bestimmt also die (inferenzstatistische (!)) Bedeutsamkeit, bzw. ab welcher Fehlerwahrscheinlichkeit ein empirisches Ergebnis nicht mehr mit der H_0 vereinbar sein soll

[inhaltliche Bedeutsamkeit eines Effektes wird durch die Effektgröße denotiert]

- ist (im Falle der Hypothesenprüfung) **vor** der Durchführung der Untersuchung zu bestimmen

Testwahl

die korrekte Wahl des Signifikanztests ist abhängig von

- der Art der Hypothesen: Zusammenhang oder Unterschied oder Veränderung
- des Skalenniveaus der (abhängigen) Variablen: Verfahren für nominal-, ordinal-, intervallskalierte Variablen
- der Relation der Variablen (bzw. Stichproben) zueinander: abhängig oder unabhängig
- der Anzahl der Variablen: UV: ein- oder mehrfaktoriell, AV: uni- oder multivariat
- der Erfüllung von (mathematischen) Voraussetzungen des Testalgorithmus bzgl. der (multivariaten) Verteilung der Daten

Berechnung des Stichprobenkennwerts

- Stichprobenkennwert ist hier (im Falle des t-Tests für unabhängige Stichproben) die Differenz der arithmetischen Mittelwerte der beiden zu vergleichenden Gruppen (Stichproben)

$$\Delta = \bar{x}_1 - \bar{x}_2$$

Berechnung des Testwerts

- der Testwert ist (im Vergleich zum Kennwert) in der Regel „normalisiert“ (um an einer Standardverteilung beurteilt werden zu können)
- hier: t ist eine Mittelwertsdifferenz, die am (empirisch abgeschätzten) Standardfehler der Mittelwertsdifferenzen relativiert wird

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\hat{\sigma}_{(\bar{x}_1 - \bar{x}_2)}}$$

Bestimmung der Fehlerwahrscheinlichkeit

erfolgt über Abgleich des empirisch ermittelten Testwerts mit der Testwerteverteilung (-stabelle)

- im Falle des t-Tests ist die t-Werte-Verteilung einschlägig

Vergleich mit Fehlerniveau

- die so ermittelte empirische (α -) Fehler-Wahrscheinlichkeit wird mit dem Signifikanzniveau verglichen
 - falls $p < \alpha$: signifikant
- davon abgesehen gibt es die Sprachregelung:
 - $p < 0,05$: signifikant
 - $p < 0,01$: hoch-signifikant
 - $p < 0,001$: höchst-signifikant