

# Statistik für Linguisten

## Grundbegriffe, Skalenniveaus

Andreas Blombach



FRIEDRICH-ALEXANDER  
UNIVERSITÄT  
ERLANGEN-NÜRNBERG

PHILOSOPHISCHE FAKULTÄT  
UND FACHBEREICH THEOLOGIE

# Teilbereiche der Statistik



FRIEDRICH-ALEXANDER  
UNIVERSITÄT  
ERLANGEN-NÜRNBERG

PHILOSOPHISCHE FAKULTÄT  
UND FACHBEREICH THEOLOGIE

# Deskriptive/beschreibende Statistik

## **Beschreibung** von Daten:

- Aufbereitung und Darstellung (in Form von Tabellen, Diagrammen und Maßzahlen)
- dabei ggf. auch Informationsverdichtung
- Datenvalidierung

# Induktive/inferentielle/schließende Statistik

- Was sagen die Daten einer Stichprobe über die Grundgesamtheit aus?
- Hypothesenprüfung
- Grundlage: Wahrscheinlichkeitstheorie

## Explorative Statistik

- Entdeckung unbekannter Muster, Strukturen und Zusammenhänge in Daten (oft visuell)
- Hypothesengenerierung

# Grundbegriffe



FRIEDRICH-ALEXANDER  
UNIVERSITÄT  
ERLANGEN-NÜRNBERG

PHILOSOPHISCHE FAKULTÄT  
UND FACHBEREICH THEOLOGIE

## Einige Grundbegriffe (1)

Die untersuchten Einzelobjekte einer Erhebung (Beobachtung, Befragung, Experiment), also z.B. die Versuchspersonen, nennt man **statistische Einheiten** oder **Merkmalsträger**. Die Menge aller statistischen Einheiten mit übereinstimmenden Identifikationskriterien (z.B. „alle Germanistikstudenten im ersten Semester“) nennt man **Grundgesamtheit** oder **Gesamtpopulation**.

Da es aus praktischen Gründen oft nicht möglich ist, die Grundgesamtheit zu untersuchen, die einen interessiert, zieht man eine **Stichprobe** (z.B. 100 Germanistikstudenten im ersten Semester). Diese erlaubt (unter gewissen Vorbehalten) Rückschlüsse auf die Grundgesamtheit, sofern sie groß genug ist und methodisch sauber gezogen wurde (im Idealfall: echte Zufallsstichprobe).

## Einige Grundbegriffe (2)

Die Eigenschaften der statistischen Einheiten, die untersucht werden, nennt man **Merkmale** oder **Variablen** (z.B. Größe, Geschlecht). Sie können verschiedene **Ausprägungen** haben (z.B. 1,70 m, weiblich).

Merkmale unterscheidet man danach, ob sie **qualitativ** (kategorial) oder **quantitativ** (metrisch) sind. Erstere lassen sich nur der Art nach erfassen (z.B. Geschlecht), letztere mit Zahlen (z.B. Größe). Außerdem unterscheidet man bei quantitativen Merkmalen **diskrete**, die nur bestimmte Ausprägungen annehmen können (z.B. Anzahl Studenten), von **stetigen**, die alle Werte aus einem Intervall annehmen können (z.B. Größe, Temperatur).



## Alles verstanden?

- Was ist in folgenden Fällen Grundgesamtheit, statistische Einheit/Merkmalsträger, was Variable?
  - Ich möchte die Hypothese prüfen, dass Statistikkurse zu Depressionen führen.
  - Ich möchte die Hypothese prüfen, dass Kinder besser Englisch lernen, wenn sie dabei durch brennende Reifen springen.
  - Ich möchte die Hypothese prüfen, dass in belletristischen Texten mehr Adjektive verwendet werden als in wissenschaftlichen.
- Wie lassen sich die Variablen messen? Sind sie qualitativ oder quantitativ?

# Skalenniveaus



FRIEDRICH-ALEXANDER  
UNIVERSITÄT  
ERLANGEN-NÜRNBERG

PHILOSOPHISCHE FAKULTÄT  
UND FACHBEREICH THEOLOGIE

## Skalenniveaus

Abhängig von den untersuchten Merkmalen werden verschiedene **Skalen** verwendet. Vom Skalenniveau, auf dem sich Daten befinden, hängt ab, was sich daraus berechnen lässt und welche statistischen Verfahren überhaupt sinnvoll sind.

Häufig werden vier Skalenniveaus unterschieden (die aus der Psychologie stammen)\*:

- |                    |   |                       |
|--------------------|---|-----------------------|
| 1. Nominalskala    | } | qualitativ/kategorial |
| 2. Ordinalskala    |   |                       |
| 3. Intervallskala  | } | quantitativ/metrisch  |
| 4. Verhältnisskala |   |                       |

\* Es gibt aber auch feinere Unterteilungen!

## Nominalskala

- niedrigstes Skalenniveau
- Merkmalsausprägungen können nur anhand ihrer Ungleichheit unterschieden werden, sie lassen sich nicht sortieren (A kann nicht größer sein als B oder umgekehrt)
- Beispiele: Geschlecht, Muttersprache

## Ordinalskala

- Rangordnung
- z.B. Schulbildung: Hauptschulabschluss < Realschulabschluss < Abitur
- keine Aussage über die Abstände zwischen den Ausprägungen (Intervalle) – also auch nicht notwendigerweise gleich groß:

*Wie häufig sprechen Sie im Alltag Dialekt?  
immer – oft – gelegentlich – selten – nie*

*Wie sehr stimmen Sie folgenden Aussagen zu?  
[...]*

*stimme überhaupt nicht zu – stimme eher nicht zu – weder noch –  
stimme eher zu – stimme voll zu*

## Intervallskala

- metrische Skala
- Abstände zwischen Merkmalsausprägungen lassen sich genau bestimmen
- kein absoluter Nullpunkt
- z.B. Temperatur in Celsius

## Verhältnisskala

- wie Intervallskala, aber mit absolutem Nullpunkt
- daher Aussagen zu Verhältnissen möglich: „A ist doppelt so groß wie B“
- z.B. Temperatur in Kelvin, Alter

# Überblick

	Nominal- skala	Ordinal- skala	Intervall- skala	Verhältnis- skala
Ausprägungs- unterschiede	+	+	+	+
sortierbar	-	+	+	+
Abstände genau bestimmbar	-	-	+	+
absoluter Nullpunkt	-	-	-	+



## Aufgaben zu Skalenniveaus

- Welche Art von Skala haben folgende Daten?
  - Reaktionszeiten bei einem Experiment
  - Grammatikalitätsurteile auf einer Skala von „absolut richtig“ bis „nicht akzeptabel“
  - Noten in einer Linguistikklausur
  - Anzahl geäußerter Silben pro Minute
  - Pluralallomorphe ({-er}, {-en}, {-s}, {-n}, {-e})
  - Platzierungsliste bei einem Schnellesewettbewerb
  - Wortlänge in Anzahl an Buchstaben
  - Jahresangaben