

Einführung B (11.03.2009)

# **MET1: Statistik**

# Einführung B: Übersicht

1. Wozu Statistik?
2. Sozialwissenschaftliche Statistik
3. Teilbereiche der Statistik
4. Ablauf einer statistischen Analyse
5. Methoden der Datengewinnung
6. Datenerfassung
7. Probleme in der Praxis
8. Übersicht Statistik Software

# Wozu Statistik?

- Weil es im Studienplan steht ;-)
- Sozialwissenschaftliche Forschung ein Qualitätsmerkmal von KWM am Standort Hagenberg darstellt
- Weil man in der beruflichen Praxis permanent mit Statistiken konfrontiert wird

# Wozu Statistik?

- *„Insbesondere der Fortschritt auf dem Gebiet der elektronischen Datenverarbeitung hat die Anwendung statistischer Methoden bei empirischen Untersuchungen praktisch jedem interessierten Sozialwissenschaftler möglich gemacht. Doch gerade wegen der leichten Anwendbarkeit immer komplexerer statistischer Verfahren ist zu beklagen, dass der mathematisch-statistische Ausbildungsstand der Forscher oft nicht den eingesetzten Methoden entspricht.“*

Atteslander (2008): Methoden der empirischen Sozialforschung. S. 229. Berlin: Erich Schmidt Verlag

# SOWI Statistik (1)

- ◎ **Naturwissenschaften:** oft direkte Messwerte
- ◎ **Sozialwissenschaften:** Umsetzen von Meinungen, Verhaltensweisen in Zahlen

# SOWI Statistik (2)

- Interessierte Merkmale variieren stark, daher kommt es „nur“ zu Wahrscheinlichkeitsaussagen
- Es wird der Schluss auf die „Gesamtgruppe“ gezogen

# SOWI Statistik (3)

## ○ Beispiel NAWI:

*„Wasser besteht aus  $H_2O$ -Molekülen, diese wiederum aus zwei Atomen Wasserstoff und einem Atom Sauerstoff. Die Massen des Wasserstoff- und Sauerstoffatoms sind bekannt. Kennt man die Eigenschaften eines Moleküls, so kennt man die Eigenschaften aller Wassermoleküle, heute wie in Zukunft.“*

Atteslander (2008): Methoden der empirischen Sozialforschung. S. 230. Berlin: Erich Schmidt Verlag

# SOWI Statistik (4)

## ◎ **Beispiel SOWI:**

Wir wollen die Größe österreichischer Familien ermitteln. Eine Familie zu betrachten genügt hier nicht; idealerweise werden wir alle Familien erfassen, um daraus die durchschnittliche Größe zu berechnen.

➔ **Einfacher: Wir ziehen eine Stichprobe ...**  
(siehe Themenblock Nr. 6)

# SOWI Statistik (5)

- ◎ Grenzen der SOWI Statistik:
  - Keine absolut gültigen und zeitlich unbegrenzten Erkenntnisse
  - Ergebnisse sind genau begrenzt
- ◎ Möglichkeiten der SOWI Statistik:
  - Kann Einblick in die Konformität von Werthaltungen geben
  - Kann aber auch den Nachweis der Unrichtigkeit vorhandener Überzeugungen erbringen

# SOWI Statistik (6)

- ◉ Wahrscheinlichkeitstheorie:  
Wenn das Eintreten bestimmter Ereignisse nicht mehr deterministisch ist, spricht man von stochastischen (zufälligen) Ereignissen.
- ◉ Chance für einen 6er im Lotto:  
1:14 Millionen (Deutschland)

# SOWI Statistik (7)

- ⊙ Klassischer Wahrscheinlichkeitsbegriff:

$$\text{Wahrscheinlichkeit } P(A) = \frac{\text{Anzahl der interessierten Ereignisse}}{\text{Anzahl aller möglichen Ereignisse}}$$

- ⊙ Statistischer Wahrscheinlichkeitsbegriff:

$$\text{Wahrscheinlichkeit } P(A) = \frac{\text{Häufigkeit des Auftretens eines Ereignisses } A}{\text{Häufigkeit des Auftretens aller Ereignisse}}$$

# SOWI Statistik (8)

- ⊙ Statistischer Wahrscheinlichkeitsbegriff:

$$= \frac{f(A)}{n}$$

$$n \rightarrow \infty$$

$$P(\text{Frau}) = \frac{526}{1000} = 0,526$$

# Teilbereiche der Statistik (1)

- **Deskriptive Statistik:**

Die beschreibende Statistik ist der Ausgangspunkt jeder Datenanalyse. Daten werden in Tabellen und Grafiken aufbereitet (vgl. Informationsdesign) und einfache statistische Kennzahlen berechnet. Mit ihrer Hilfe kann ein erster Überblick geschaffen werden.

# Teilbereiche der Statistik (2)

- ◎ **Induktive Statistik:**

Der erhobene Datensatz ist eine repräsentative Stichprobe.

Rückschlüsse von der Stichprobe auf die Grundgesamtheit können durch Methoden der schließenden (induktiven) Statistik hergestellt werden. Die Unsicherheit lässt sich abschätzen.

# Teilbereiche der Statistik (3)

- **Explorative Datenanalyse:**  
Inzwischen ein eigener Bereich zwischen deskriptiver und schließender Statistik. Sie wird zur Suche nach Strukturen und nach möglichen Hypothesen verwendet. Die Überprüfung erfolgt dann mittels induktiver Statistik.

# Ablauf einer stat. Analyse (1)

- Die Planung ist ein wichtiger Bestandteil einer statistischen Erhebung. Ein konkretes Forschungsziel ist entscheidend für die Qualität und Aussagekraft der Ergebnisse. Unzureichende Planung kann später nicht mehr korrigiert und nur sehr schwer kompensiert werden.

# Ablauf einer stat. Analyse (2)

- ⊙ Vorbereitung
- ⊙ Datengewinnung
- ⊙ Datenerfassung und –aufbereitung
- ⊙ Statistische Auswertung
- ⊙ Bericht

# Ablauf einer stat. Analyse (3)

- ◎ @ Vorbereitung
  - Zielformulierung
  - Zeitplan erstellen
  - Festlegen der Grundgesamtheit und ggf. von Merkmalen
  - Auswertungsmethode festlegen
  - Pretest (Probeerhebung und Korrekturen)

# Methoden der Datengewinnung (1)

- ◎ **Wie können Daten gewonnen werden?**  
(→ Diskussion)
- ◎ Schlecht oder unzureichend erhobene Daten können auch nicht mit den Methoden der Statistik „gerettet“ werden!

# Methoden der Datengewinnung (1)

- ◎ Primärstatistische Erhebung
- ◎ Sekundärstatistische Erhebung

# Datenerfassung (1)

- Informationen werden beispielsweise den ausgefüllten Fragebögen entnommen (mittels Dateneingabemaske oder direkt in eine Tabelle)
- **Codeplan:** Anleitung zur Übersetzung der Informationen in Zahlen

# Datenerfassung (2)

- Informationen werden nach Möglichkeit als Zahlen erfasst
- Erste Zeile enthält die Variablennamen
- Erste Spalte enthält eine laufende Nummer
- Jede Zeile = Erhebungseinheit
- Jede Spalte = Variable (Achtung bei Mehrfachnennungen!)

# Probleme in der Praxis

- ⊙ **Datenschutz und Anonymität**
  - Gesetze beachten
  - Rückschlüsse auf Einzelne möglich?
- ⊙ **Unzureichendes Studiendesign**
  - Stichprobenbildung (Repräsentativität)
  - Erreichbarkeit der Zielgruppe (Telefon- oder Onlineumfrage?)
- ⊙ **Sekundärstatistiken**
  - Wie wurden die Daten erhoben?
  - Was wurde erhoben?
  - Wann wurde erhoben?
  - Wie wurde codiert?

# Statistiksoftware

- ◎ **Excel** (kann mehr als man auf dem ersten Blick ahnt!)
- ◎ **SPSS** (<http://www.spss.com/de/>)
- ◎ **SAS** (<http://www.sas.com/offices/europe/germany/index.html>)
- ◎ **LISREL** (<http://www.ssicentral.com/>)
- ◎ **The R Project** (<http://www.r-project.org/>)
- ◎ **ALMO** (<http://www.almo-statistik.de/>)  
wurde an der JKU Linz entwickelt; hat zwar eine veraltete Oberfläche,  
beherrscht aber viele relevante stat. Verfahren

# SPSS

The screenshot displays the SPSS Statistics Data Editor window with a dataset named 'transactions.sav'. The data is organized into columns: Date, ID, Product Line, Amount, and Product Number. A dialog box titled 'Get It Analyzed from Transaction Data' is overlaid on the data, showing the following configuration:

- Variables:** Product Line (ProductLine), Product Number (ProductNumber)
- Transaction Date:** Purchase Date (Date)
- Transaction Amount:** Amount (Amount)
- Summary Method:** Total
- Customer Identifiers:** Customer ID (ID)

ID	Date	Product Line	Amount	Product Number
1	05/15/2004	500	192	500
2	04/05/2004	500	158	500
3	05/18/2004	500	196	500
4	06/04/2006	500	181	500
5	03/28/2005	500	150	500
6	08/01/2004	500	181	500
7	05/19/2006	500	181	500
8	07/21/2004	500	181	500
9	04/01/2004	500	181	500
10	09/21/2006	500	181	500
11	06/17/2005	500	181	500
12	08/15/2005	500	181	500
13	07/18/2005	500	181	500
14	06/24/2005	500	181	500
15	06/25/2004	500	181	500
16	10/11/2004	500	181	500
17	07/14/2004	500	181	500
18	05/07/2004	500	181	500
19	03/06/2006	500	181	500
20	09/16/2006	500	181	500
21	12/23/2006	500	181	500
22	05/28/2005	500	181	500
23	02/19/2006	500	181	500
24	12/13/2004	500	210	500
25	04/15/2005	500	211	500

# Onlineumfragen

- ◎ Limesurvey (Open Source)
  - <http://www.limesurvey.org/>
  - große Community
  - viele Funktionen
  - gleichwertig zu kommerziellen Angeboten
  - Leichte Anpassbarkeit (mittels PHP und CSS für Layout)
  - Datenexport als .xls .csv .doc und für SPSS

# Limesurvey

**LimeSurvey 1.8RC2 Online Demo!**

Administration -- Logged in as: admin

Surveys: New Survey

Survey New Survey (10:56176)

Groups: New Group

Group New Group (10:39)

Questions: 2: Please rate on a 10 point scale

Question Please rate on a 10 point scale how you feel about the quality of our products. (10:35)

Code: 2: (Mandatory question)

Question: Please rate on a 10 point scale how you feel about the quality of our products.

Help: Select a radio button that reflects your response.

Type: Array (10 point choice)

Mandatory: Yes

Other questions having conditions on this question: [QID: 37]

LimeSurvey  
Version 1.8RC2 (6062)

Like it? Donate to LimeSurvey