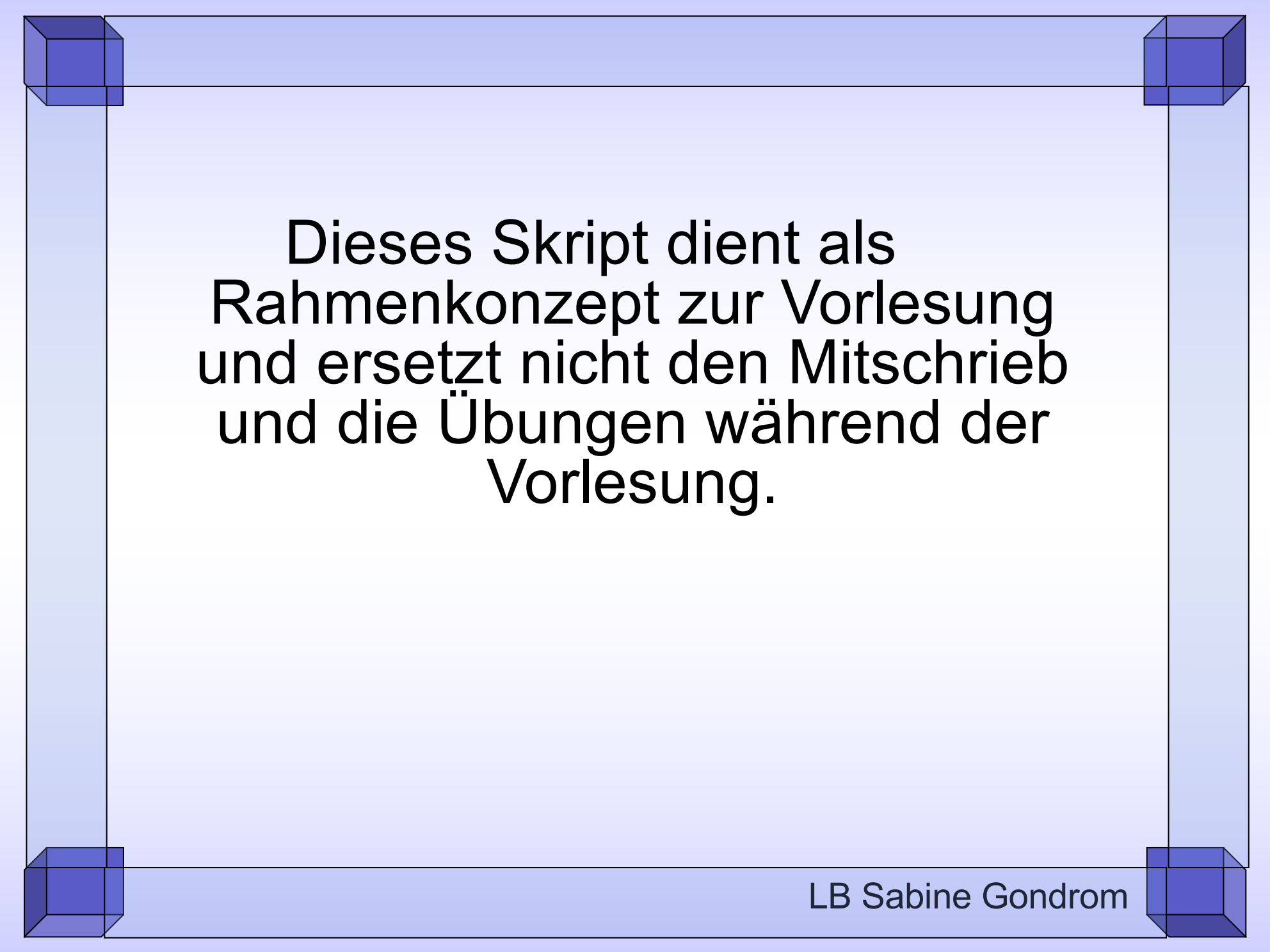


# **STATISTIK**

## **Teil 1**

LB Sabine Gondrom



Dieses Skript dient als  
Rahmenkonzept zur Vorlesung  
und ersetzt nicht den Mitschrieb  
und die Übungen während der  
Vorlesung.

## Literatur:

Josef Puhani: **Statistik, Einführung mit praktischen Beispielen**, 11. Auflage

Josef Puhani: **Kleine Formelsammlung zur Statistik**, 10. Auflage

## Inhaltsübersicht

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von der Stichprobe auf die Grundgesamtheit

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## **Exkurs: Ablauf einer statistischen Untersuchung (I)**

### **1. Die Planung**

Sachliche Abgrenzung des Untersuchungsgegenstandes

Methodische Abgrenzung bzw. Wahl des statistischen Verfahrens

-> wird beeinflusst vom Ziel der Untersuchung

Räumliche und zeitliche Abgrenzung

Organisatorische Vorbereitung der Untersuchung

### **2. Die Erhebung**

Gewinnung der Daten

Primärstatistische Erhebung:

Daten werden eigens für die Untersuchung erhoben,

z. B. durch:

Schriftliche Befragung, mündliche Befragung,  
Beobachtungen, automatische Verfahren,  
Experimente

Sekundärstatistische Erhebung:

Daten werden aus bereits vorhandenem Datenmaterial gewonnen.

## **Deskriptive Statistik**

Typisierung und Darstellung von Daten

Mittelwerte

Streuungsmaße

Indexzahlen

Korrelation und Regression

Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

Kombinatorik

Wahrscheinlichkeitsrechnung

Theoretische Verteilungen

Schluss von d. Stichprobe auf d. Grundgesamtheit

# **Exkurs: Ablauf einer statistischen Untersuchung (II)**

## **3. Die Aufbereitung**

Das erhobene Urmaterial wird geordnet, auf Fehler untersucht, verdichtet und somit für die Anwendung des statistischen Verfahrens vorbereitet

## **4. Die Analyse**

Untersuchung des aufbereitenden Datenmaterials mittels des ausgewählten statistischen Verfahrens

## **5. Die Interpretation**

Die Ergebnisse der Analyse werden sachbezogen interpretiert, wobei meist noch weitere Informationen, die nicht in die Untersuchung miteinbezogen wurden, heranzuziehen sind

Die Interpretation der Ergebnisse hängt maßgeblich vom Untersuchungszweck ab

## **Deskriptive Statistik**

### **• Typisierung und Darstellung von Daten**

- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlich- keitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlich-keitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# **Typisierung und Darstellung von Daten**

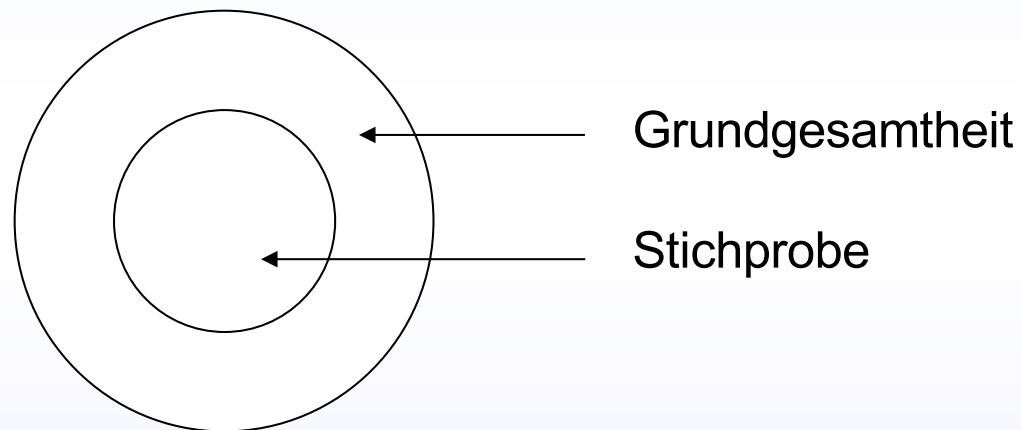
### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

- Statistische Untersuchungen vollziehen sich an sachlich, räumlich und zeitlich abgegrenzten Mengen von Elementen.
  - > Grundgesamtheit
- Teilmenge der Grundgesamtheit soll repräsentativ für die Grundgesamtheit sein
  - > Stichprobe



- Die Elemente der Grundgesamtheit heißen Merkmalsträger oder statistische Einheiten
- Merkmalsträger besitzen Eigenschaften
  - > Merkmale



## Deskriptive Statistik

### • Typisierung und Darstellung von Daten

- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Typisierung von Merkmalen

### • Diskrete Merkmale

können endliche viele oder abzählbar unendlich viele Ausprägungen annehmen

### • Stetige Merkmal

können jeden reellen Wert zumindest innerhalb eines Intervalls annehmen

## Deskriptive Statistik

### • Typisierung und Darstellung von Daten

- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Skalierung von Merkmalen

- metrisch skalierte Merkmale
  - genau messbar
  - Aussagen über Abstand möglich
- ordinal skalierte Merkmale
  - nicht genau messbar
  - keine Aussagen über Abstand möglich
  - Rangordnung feststellbar
- nominal skalierte Merkmale
  - nicht genau messbar
  - keine Aussagen über Abstand möglich
  - Rangordnung nicht feststellbar

## Deskriptive Statistik

### • Typisierung und Darstellung von Daten

- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Notation

$X$  - > Merkmal

$x_i$  - > Merkmalsausprägung  $i$

$N$  - > Anzahl der Elemente in der Grundgesamtheit

$n$  - > Anzahl der Elemente in der Stichprobe  
(Stichprobenumfang)

$n_i$  - > Anzahl der Elemente der Merkmalsausprägung  $x_i$   
in der Stichprobe

$\frac{n_i}{n}$  - > relative Häufigkeit

$\frac{n_i}{n} \cdot 100[\%]$  - > prozentuale Häufigkeit

## Deskriptive Statistik

### • Typisierung und Darstellung von Daten

- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Verdichten des Urmaterials

3<sup>-></sup> Beispiel Urliste

18	5	15	
2	23	5	17
20	7	29	20
15	12	15	30
9	5	11	2

### Häufigkeitsverteilung

$X_i$	$x_1$	$x_2$	$x_3$
$n_i$	$n_1$	$n_2$	$n_3$

-> weiteres Verdichten durch Bilden von Klassen möglich

# **STATISTIK**

## **Teil 2**

LB Sabine Gondrom

## Deskriptive Statistik

### • Typisierung und Darstellung von Daten

- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Graphische Darstellung (II)

### von Häufigkeitsverteilungen diskreter Merkmale

- auf der Abszisse werden die Merkmalsausprägungen abgetragen
- auf der Ordinate werden die jeweils zugeordneten absoluten/ relativen Häufigkeiten abgetragen
- Längen der Linien/ Streifen sind proportional zu den Häufigkeiten
- Liniendiagramm
  - bei metrisch skalierten Merkmalen
- Streifendiagramm
  - bei ordinal oder nominal skalierten Merkmalen

## **Deskriptive Statistik**

### **• Typisierung und Darstellung von Daten**

- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Graphische Darstellung (III)

## von Häufigkeitsverteilungen stetiger Merkmale

### • Histogramm

- Unmittelbar aneinander angrenzende Rechtecke
- Rechtecksgrenzen i. d. R. eine halbe Maßeinheit rechts und links des Messwertes
- Flächen der einzelnen Rechtecke sind den Häufigkeiten proportional

### • Häufigkeitspolygon

- alternativ oder ergänzend zum Histogramm
- wird aus einem Histogramm gewonnen, indem die Mitten der oberen Rechtecksbegrenzungen linear miteinander verbunden werden
- Fläche unter dem Häufigkeitspolygon entspricht der Fläche des Histogramms

## **Deskriptive Statistik**

### **• Typisierung und Darstellung von Daten**

- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Weitere Darstellungsformen

- Kreisdiagramm
- Staffelfeld
- Pictogramm



## **Deskriptive Statistik**

### **• Typisierung und Darstellung von Daten**

- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Summenhäufigkeitsfunktion

- Gibt die Anzahl/ den Anteil derjenigen Untersuchungseinheiten an, die eine Merkmalsausprägung von höchstens  $x$  haben
- Durch Aufsummieren der absoluten/ relativen Häufigkeiten
- Bei stetigen Merkmalen bzw. metrisch skalierten klassierten diskreten Merkmalen wird der Verlauf zwischen den Klassengrenzen durch eine Gerade (lineare Interpolation) approximiert
- Bei metrisch skalierten diskreten Merkmalen, die nicht klassiert sind, erhält man eine Summenhäufigkeitsfunktion in Form einer Treppenfunktion

## **Deskriptive Statistik**

### **• Typisierung und Darstellung von Daten**

- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Konzentrationskurve (Lorenzkurve)

- Bringt die Konzentration, d. h. die Stärke der Ungleichheit in der Verteilung der Merkmalssumme auf die Merkmalsträger zum Ausdruck
- Bei metrisch skalierten Merkmalen mit positiven Merkmalsausprägungen
- Lorenzkurve:
  - Aufsummieren (kumulieren) der nach zunehmender Größe der Merkmalsausprägungen geordneten relativen oder prozentualen Häufigkeiten
  - Zuordnen der jeweils kumulierten Anteile der Merkmalssumme zu diesen Summenhäufigkeiten
  - Summenhäufigkeiten werden auf der Abzisse und kumulierte Anteile auf der Ordinate abgetragen

## Deskriptive Statistik

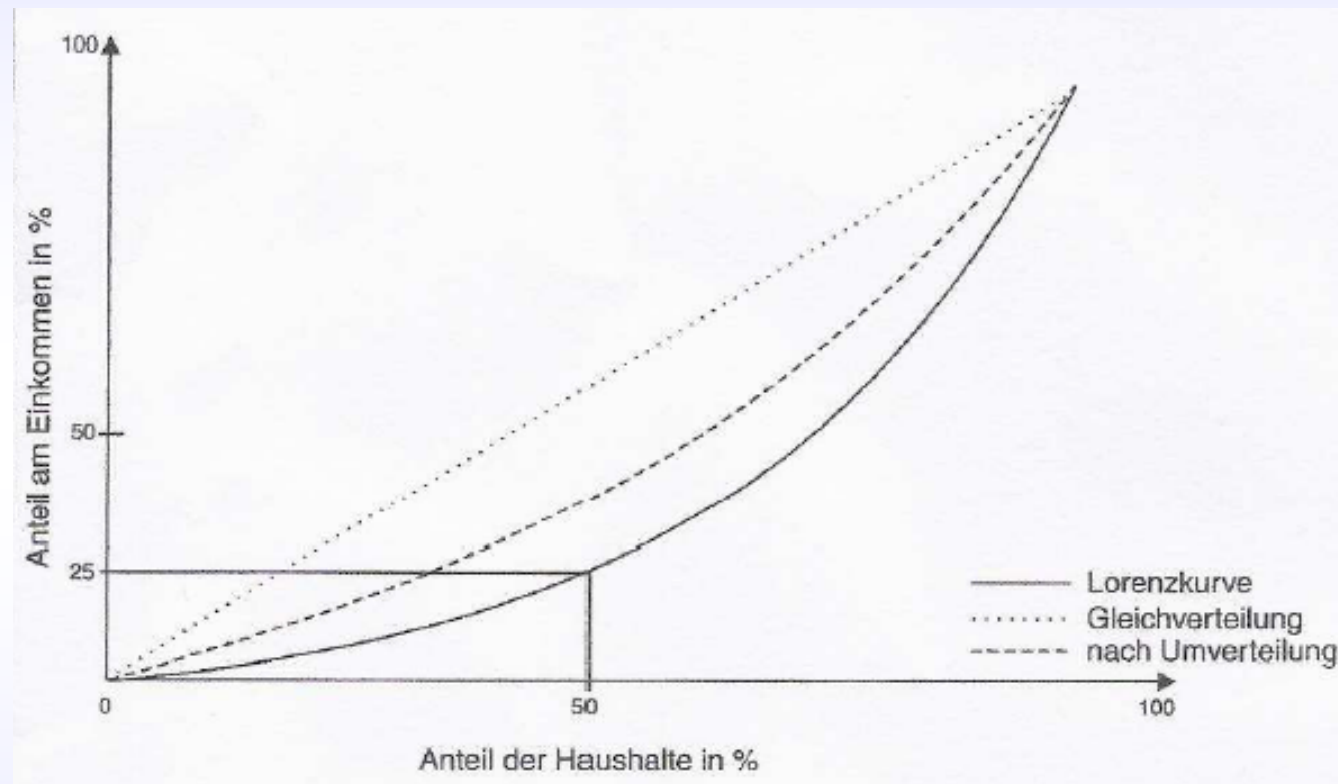
### • Typisierung und Darstellung von Daten

- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Beispiel Lorenzkurve



## Spitzenverdiener zahlen ein Viertel der Lohnsteuer

Deutschlands Top-Verdiener stemmen einen Großteil der Steuerlast. Sie zahlen ein Viertel der Abgaben, dabei ist ihre Zahl sehr gering.

ARTIKEL TEIL



Empfänger-E-M



FOTO: PICTURE ALLIANCE / MAXPPP

Hohe Einkommen, hohe Abgaben: Deutschlands Topverdiener tragen einen Großteil der Steuerlast

Die 383.000 Spitzenverdiener mit Einkünften von mehr als 172.000 Euro in Deutschland zahlen ein Viertel der Lohn- und Einkommensteuer. Dabei machen sie lediglich rund ein Prozent der erfassten Lohn- und Einkommensteuerpflichtigen aus. Zu diesen Ergebnissen kommt das Statistische Bundesamt in seiner Auswertung der Einkommensteuerstatistik 2007, die alle drei Jahre vorgelegt wird.

Auf die 19,2 Millionen **Geringverdiener** mit weniger als 22.500 Euro entfallen 3,6 Prozent der festgesetzten Einkommensteuer. Insgesamt wurden 38,4 Millionen Steuerpflichtige erfasst, die Einkünfte von 1,2 Billionen Euro erzielten. Die von den Finanzbehörden festgesetzte Lohn- und Einkommensteuer betrug rund 211 Milliarden Euro. Dies waren 30,1 Milliarden Euro oder 16,7 Prozent mehr als 2004.

### VIDEOS





## Dieter Nuhr

Welt: 10% Bestverdiener zahlen 60% der Steuer. Untere 50% zahlen 3,6% der Steuern. Gut! Aber wo bleibt da die Ungerechtigkeit?

Also zahlen nur 63,6 Prozent der Deutschen überhaupt Steuern? :)

13. Oktober um 12:26 · Gefällt mir · 4 Personen

10%+60%+50%+3,6% ergeben gerade mal 98% - ich glaub, die von der Welt können nicht rechnen. ?!

13. Oktober um 12:26 · Gefällt mir · 1 Person

Und was ist mit den Restlichen 36,4%, Statistiken sind doch fürn Arsch...:D

13. Oktober um 12:28 · Gefällt mir · 2 Personen

...ich komme auf einen Überschuss von 23,6%... waren das jetzt Steuern oder Steuerzahler?

13. Oktober um 12:32 · Gefällt mir

die ungerechtigkeit liegt darin, dass es nur 10% bestverdiener gibt!

13. Oktober um 12:37 · Gefällt mir · 8 Personen

■ D.h. also

- Die 1% Bestverdienenden zahlen 25% der Einkommenssteuer
- Die 50% Schlechtverdienenden zahlen 3,6% der Einkommenssteuer

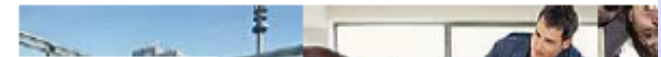
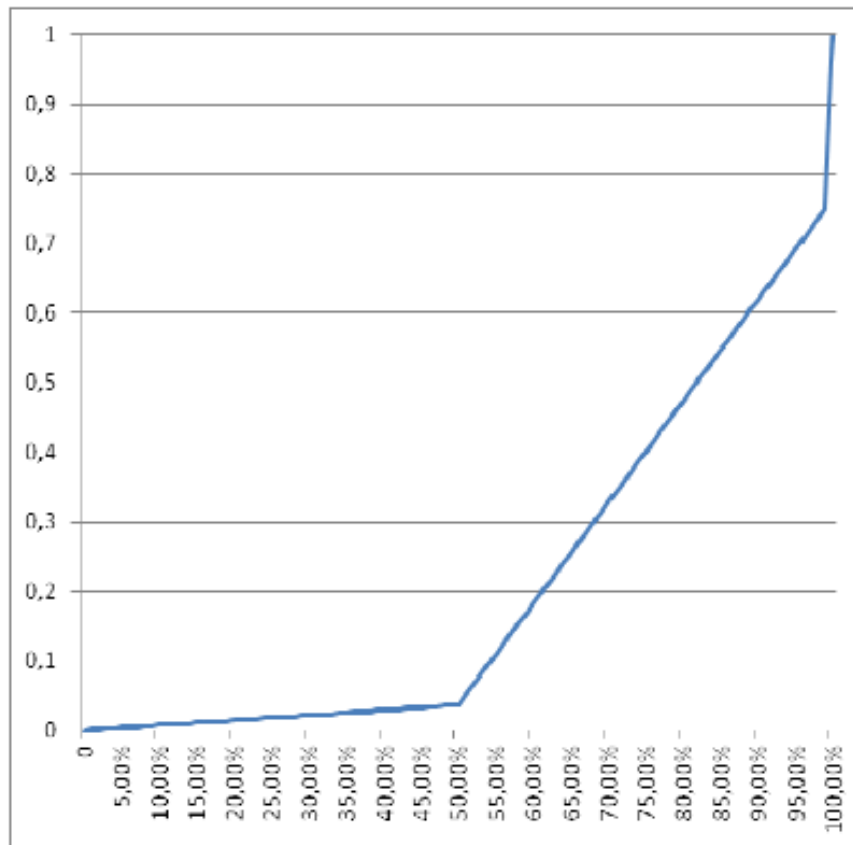
Klasse	Verdiener	Einkommenssteuer
1	50,00%	3,60%
2		
3	1,00%	25,00%

■ Die Mittelschicht sieht also folgendermaßen aus

Klasse	Verdiener	Einkommenssteuer
1	50,00%	3,60%
2	49,00%	71,40%
3	1,00%	25,00%

## ■ Die Lorenzkurve benötigt kumulierte Daten

Klasse	Verdiener	Kumulierte Verdiener	Einkommenssteuer	Kumulierte Einkommenssteuer
1	50,00%	50,00%	3,60%	3,60%
2	49,00%	99,00%	71,40%	75,00%
3	1,00%	100,00%	25,00%	100,00%





# **STATISTIK**

## **Teil 3**

LB Sabine Gondrom



## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- **Mittelwerte**
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Mittelwerte

## Deskriptive Statistik

- Typisierung und Darstellung von Daten
- **Mittelwerte**
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Arithmetisches Mittel bei Einzelwerten

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

## Deskriptive Statistik

- Typisierung und Darstellung von Daten
- **Mittelwerte**
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Gewogenes arithmetisches Mittel (Arithmetisches Mittel bei einer Häufigkeitsverteilung)

$$\bar{x} = \frac{x_1 n_1 + x_2 n_2 + \dots + x_k n_k}{n} = \frac{1}{n} \sum_{i=1}^k x_i n_i = \sum_{i=1}^k x_i \frac{n_i}{n}$$

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- **Mittelwerte**
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Arithmetisches Mittel für klassierte Merkmalsausprägungen

$$\bar{x} \approx \frac{1}{n} \sum_{i=1}^k x_i^* n_i = \sum_{i=1}^k x_i^* \frac{n_i}{n}$$

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- **Mittelwerte**
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Zentralwert (Median)

- Merkmalsausprägung, die in der Mitte der in eine Reihenfolge gebrachten Einzelausprägungen der n Elemente steht
- bei ungerader Anzahl der Elemente n ist der Zentralwert an der Stelle:

$$\frac{n+1}{2}$$

- bei gerader Anzahl der Elemente n ist der Zentralwert an der Stelle:

$$ZW_1: \frac{n}{2}$$

$$ZW_2: \frac{n}{2} + 1$$

## Häufigster Wert (Modus)

- Merkmalsausprägung, die am häufigsten vorkommt

## Deskriptive Statistik

- Typisierung und Skalierung von Daten
- **Mittelwerte**
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Weitere Verfahren in der Marktforschung

# Beispiel: Höhen von Flaschen



Flaschen-Nr.	Höhe in cm ( $x_i$ )	Anzahl der Flaschen ( $n_i$ )
1	18,1	1
2	20,5	3
3	22,3	6
4	19,7	4
• Arithmetisches Mittel?	18,9	2

- Median?
- Modus?

### **Deskriptive Statistik**

- Typisierung und Skalierung von Daten
- **Mittelwerte**
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

### **Weitere Verfahren in der Marktforschung**

## Geometrisches Mittel

$$\sqrt[n]{X_1 \cdot X_2 \cdot \dots \cdot X_n}$$

$X_1, X_2, \dots, X_n$  Wachstumsfaktoren

Durchschnittliche Wachstumsrate =  
Durchschnittlicher Wachstumsfaktor - 1

### **Deskriptive Statistik**

- Typisierung und Skalierung von Daten
- **Mittelwerte**
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

### **Weitere Verfahren in der Marktforschung**

## Beispiel Geometrisches Mittel

Gegeben sind folgende Wachstumsraten des Umsatzes in den letzten 4 Jahren:  
+12%, +7%, -2%, +5%.

Wie viel Prozent beträgt die durchschnittliche Wachstumsrate?



## Deskriptive Statistik

- Typisierung und Darstellung von Daten
- **Mittelwerte**
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Harmonisches Mittel

$$HM = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

### **Deskriptive Statistik**

- Typisierung und Skalierung von Daten
- **Mittelwerte**
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

### **Weitere Verfahren in der Marktforschung**

## Beispiel Harmonisches Mittel

Es werden zu vier Zeitpunkten für jeweils 10 000 EUR Anteile von einem Fond gekauft.

Die Einzelkurse des Fonds pro Zeitpunkt sind:  
79,5/ 72,30/ 68,70/ 60,30

Gesucht ist der mittlere Kaufkurs

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- **Streuungsmaße**
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Streuungsmaße

## **Deskriptive Statistik**

Typisierung und  
Darstellung von  
Daten

Mittelwerte

## **Streuungsmaße**

Indexzahlen

Korrelation und  
Regression

Elemente d.  
Zeitreihenanalyse

## **Wahrscheinlich- keitsrechnung und schließende Statistik**

Kombinatorik

Wahrscheinlich-  
keitsrechnung

Theoretische  
Verteilungen

Schluss von d.  
Stichprobe auf d.  
Grundgesamtheit

# Spannweite

$$SW = x_{\max} - x_{\min}$$

beschreibt den gesamten Streubereich der Beobachtungsreihe

für Maßstabsbetrachtungen auf der x-Achse wichtig

### **Deskriptive Statistik**

- Typisierung und Skalierung von Daten
- Mittelwerte
- **Streuungsmaße**
- Indexzahlen
- Korrelation und Regression

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

### **Weitere Verfahren in der Marktforschung**

## Beispiel Spannweite

$$x_i = 1; 3; 3; 4; 4; 4; 5; 5; 8; 10$$

Wie groß ist die Spannweite?

$$10 - 1 = 9$$

Wir haben doch eine Spannweite von 9 Einheiten

## Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- **Streuungsmaße**
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Varianz

auch für Hochrechnung auf umfangreiche Grundgesamtheiten einsetzbar

- bei Einzelwerten

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

- bei einer Häufigkeitsverteilung

$$s^2 = \frac{1}{n-1} \sum_{i=1}^k (x_i - \bar{x})^2 n_i$$

- bei klassierten Merkmalsausprägungen

$$s^2 \approx \frac{1}{n-1} \sum_{i=1}^k (x_i^* - \bar{x})^2 n_i$$

## Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- **Streuungsmaße**
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Standardabweichung

- Die Standardabweichung ist die Wurzel aus der Varianz und besitzt die gleiche Dimension wie die Beobachtungswerte  $x_i$  und der Mittelwert  $\bar{x}$ .

$$S = \sqrt{S^2}$$

### Deskriptive Statistik

- Typisierung und Skalierung von Daten
- Mittelwerte
- **Streuungsmaße**
- Indexzahlen
- Korrelation und Regression

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

### Weitere Verfahren in der Marktforschung

## Beispiel Varianz und Standardabweichung

Gesucht ist die Varianz und die Standardabweichung der Stichprobe:

$x_i$	1	2	3	4	5
$n_i$	10	22	27	25	16



# **STATISTIK**

## **Teil 4**

LB Sabine Gondrom

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- **Korrelation und Regression**
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# **Korrelation und Regression**

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- **Korrelation und Regression**
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Korrelation

Untersuchung von mehreren Merkmalen

Bsp: 2 Merkmale X und Y

- Ausprägungen für X:  $x_1, x_2$
- Ausprägungen für Y:  $y_1, y_2$
- Urliste:

$(x_1; y_1)$

$(x_1; y_2)$

- Es soll untersucht werden, ob ein Zusammenhang zwischen X und Y besteht.
- Statistisches Verfahren, welches die Stärke der Beziehung zwischen zwei Merkmalen beschreibt

- Korrelationsanalyse

behandelt beide Merkmale symmetrisch

- Regressionsanalyse

betrachtet eine Variable (Merkmal) als unabhängige Variable und die andere als abhängige Variable

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- **Korrelation und Regression**
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Korrelationskoeffizient Bravais Pearson

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

$$r = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\left[ n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2 \right] \left[ n \sum_{i=1}^n y_i^2 - \left( \sum_{i=1}^n y_i \right)^2 \right]}}$$

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- **Korrelation und Regression**
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Rangkorrelationskoeffizient nach Spearman

$$r_s = 1 - \frac{6 \cdot \sum_{i=1}^n d_i^2}{n \cdot (n^2 - 1)}$$

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- **Korrelation und Regression**
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

### • Bravais-Pearson-Korrelationskoeffizient

- misst einen linearen Zusammenhang
- für metrisch skalierte Merkmale
- kann Werte zwischen  $-1$  und  $1$  annehmen
- $r > 0$  (z. B. ab  $r > 0,5$ )  $\rightarrow$  Merkmale sind positiv korreliert;

je näher bei  $1$ , desto stärker positiv korreliert

- $r < 0$  (z. B. ab  $r < -0,5$ )  $\rightarrow$  Merkmale sind negativ korreliert;

je näher bei  $-1$ , desto stärker negativ korreliert

- $r=0 \rightarrow$  Merkmal unkorreliert

### • Rangkorrelationskoeffizient nach Spearman

- für zwei ordinal skalierte Merkmale  $X$  und  $Y$
- bei metrisch skalierten Merkmalen müssen die Merkmalsausprägungen zuerst der Größe nach sortiert werden
- Zuordnung von Rangwerten  $x_i$  und  $y_i$

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- **Korrelation und Regression**
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Regressionsrechnung

- X erklärende/ unabhängige Variable -> Regressor
- Y erklärte/ abhängige Variable -> Regressand
- Y wird durch X erklärt
- nur für metrisch skalierte Merkmalsausprägungen
- der korrelative Zusammenhang soll mittels einer linearen Funktion (= Gerade) beschrieben werden
- Gerade:  $\hat{y} = b_1 + b_2 x$
- die Gerade passt am besten, wenn die Summe

$$\text{der } \varepsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow \min$$

-> Methode der kleinsten Quadrate (C. F. Gauß)

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- **Korrelation und Regression**
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Stichprobenregressionsfunktion

$$\hat{y} = b_1 + b_2 x$$

$$y = b_1 + b_2 x + e$$



$$b_1 = \bar{y} - b_2 \bar{x}$$

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- **Korrelation und Regression**
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Regressionskoeffizienten

$$b_1 = \bar{y} - b_2 \bar{x}$$

$$b_1 = \frac{\sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i - \sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}$$

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- **Korrelation und Regression**
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Regressionskoeffizienten

$$b_2 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$b_2 = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2}$$

# **STATISTIK**

## **Teil 5**

LB Sabine Gondrom

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- **Elemente d. Zeitreihenanalyse**

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# **Elemente der Zeitreihenanalyse**

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- **Elemente d. Zeitreihenanalyse**

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## **Komponenten der Zeitreihe (I)**

- Ursprungswerte (Beobachtungswerte) einer aus ökonomischen Daten bestehenden Zeitreihe können durch Einflussgrößen beeinflusst werden:
  - langfristiger Wachstumspfad
  - zyklische Bewegungen mit einer Periodenlänge von mehreren Jahren
  - zyklische Bewegungen, die sich innerhalb eines Jahres vollziehen
  - nicht regelmäßige Einflüsse
- Einflussgrößen werden mit mehr oder weniger Gewicht zum Zustandekommen der Ursprungswerte  $y$  beitragen
- Bei der Zeitreihenanalyse sollen nicht nur die Veränderungen der Ursprungswerte festgestellt werden, sondern einzelne Komponenten isoliert erfasst werden
- Komponenten:
  - Trendkomponente (tk)
  - Konjunkturkomponente (kk)
  - Saisonkomponente (sk)
  - irreguläre Komponente (ik)

## Deskriptive Statistik

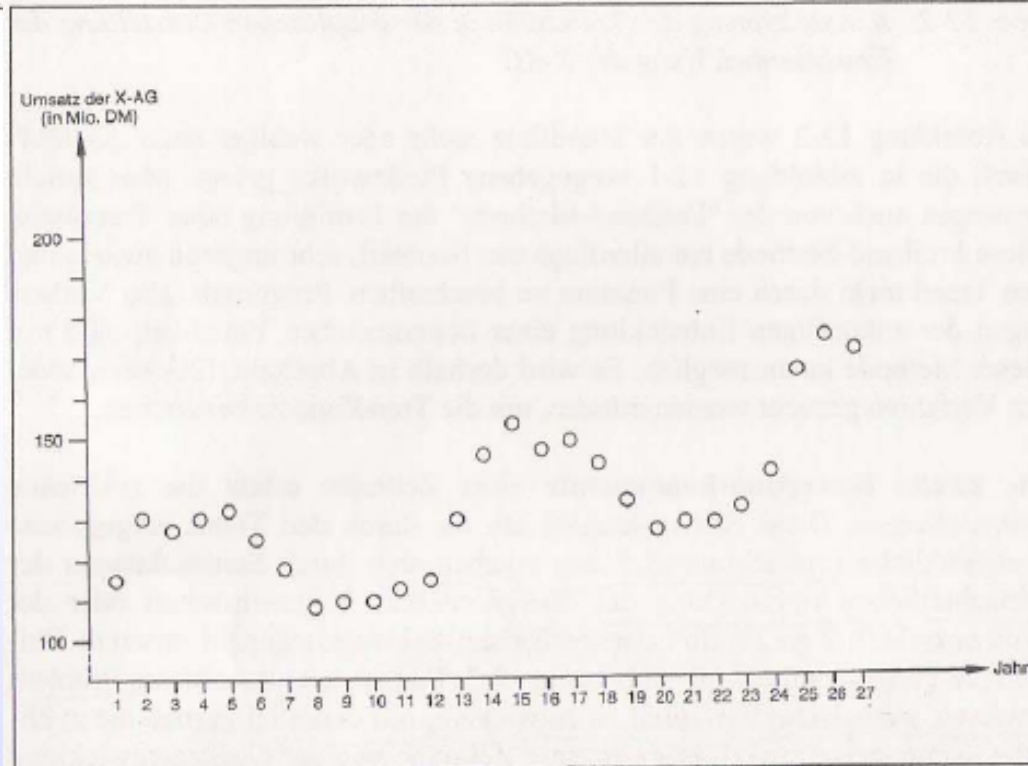
- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- **Elemente d. Zeitreihenanalyse**

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Beispiel:

Beobachtungs-jahr	Jahres-umsatz (Mio.DM)	Beobachtungs-jahr	Jahres-umsatz (Mio.DM)	Beobachtungs-jahr	Jahres-umsatz (Mio.DM)	Beobachtungs-jahr	Jahres-umsatz (Mio.DM)
1	115	8	109	15	152	22	128
2	130	9	110	16	146	23	131
3	127	10	110	17	148	24	140
4	130	11	112	18	142	25	165
5	131	12	115	19	134	26	173
6	124	13	130	20	126	27	170
7	118	14	145	21	128		



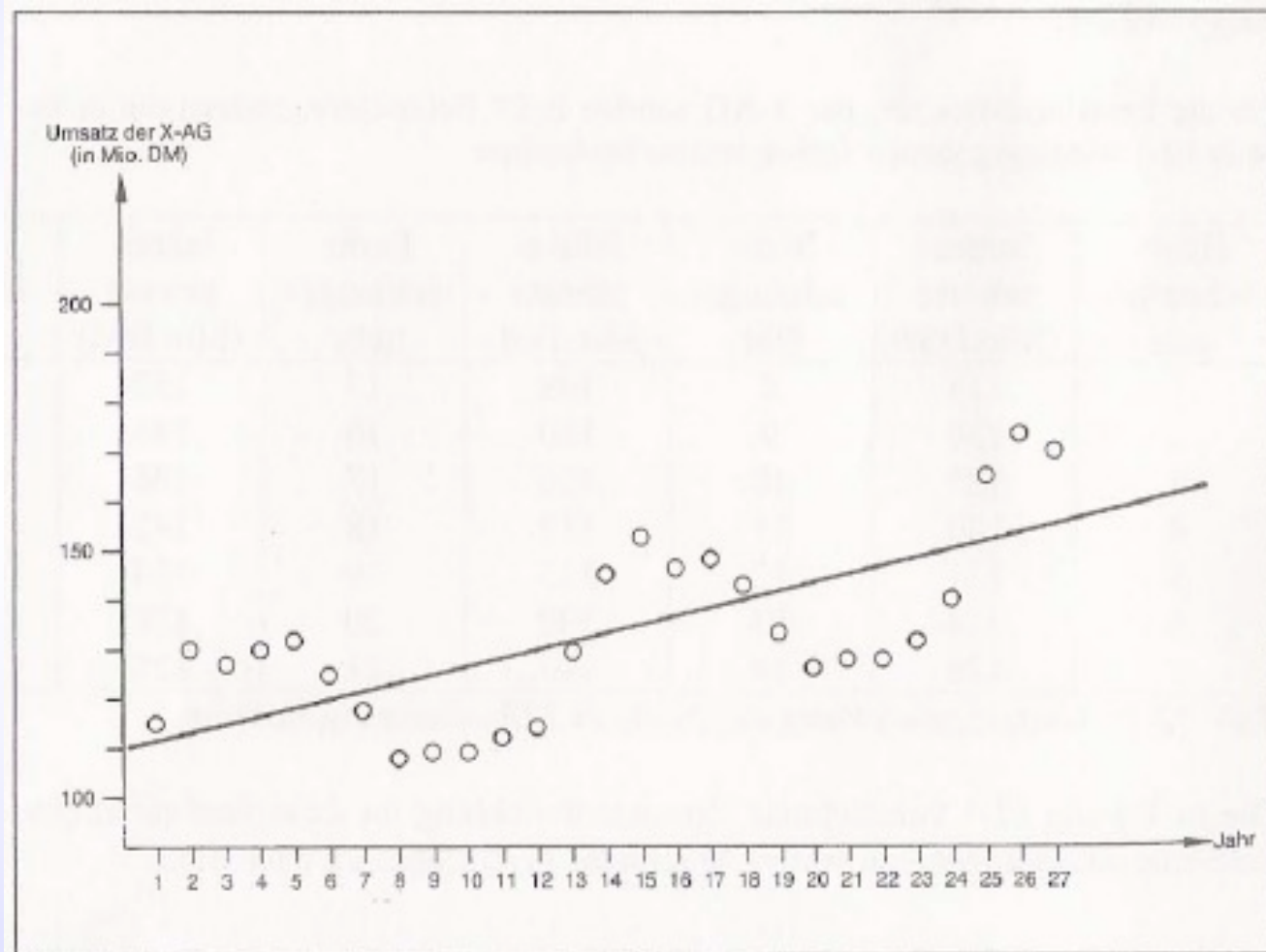
### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- **Elemente d. Zeitreihenanalyse**

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## **Trendlinie in der graphischen Darstellung der Umsatzentwicklung**



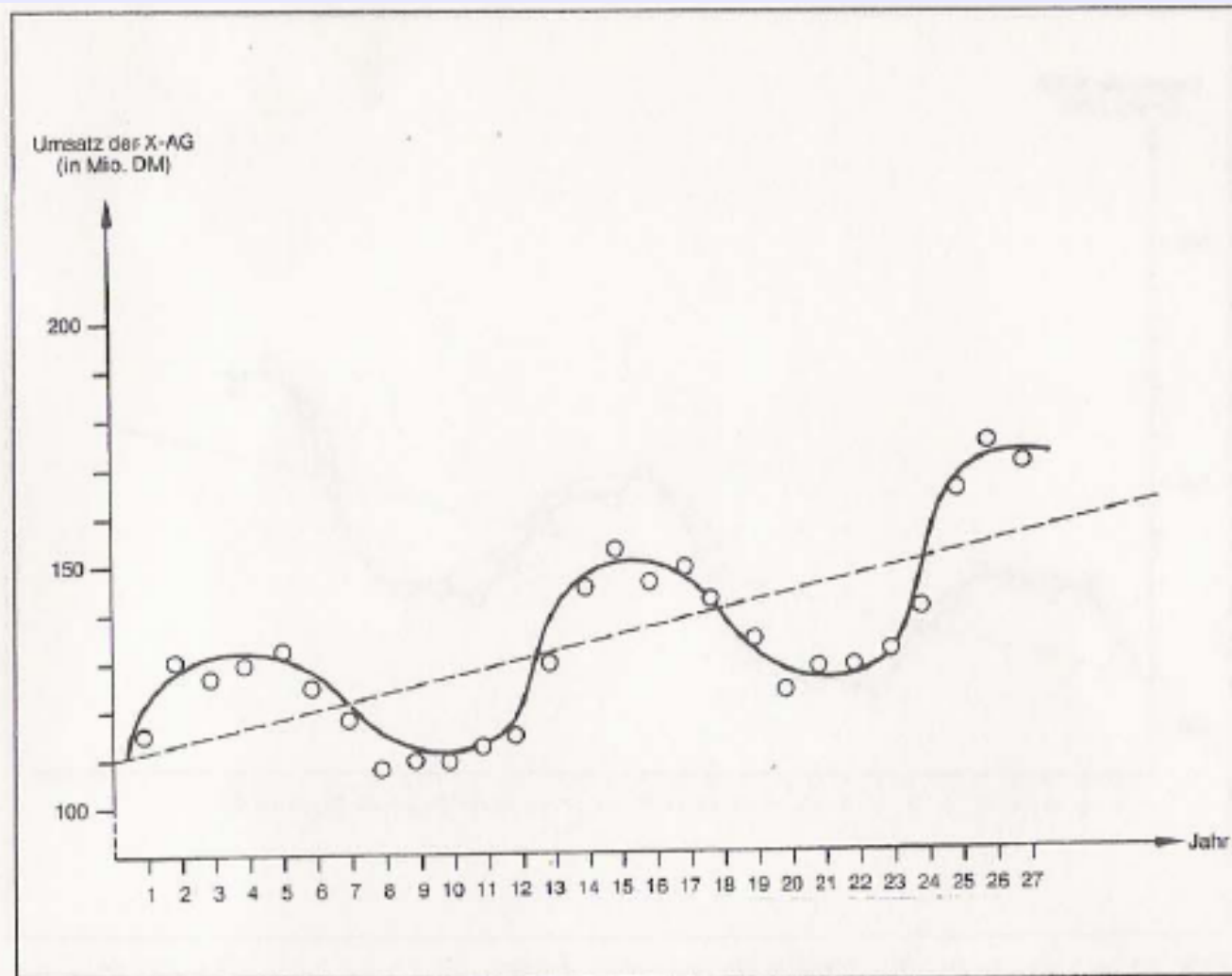
### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- **Elemente d. Zeitreihenanalyse**

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Einfluss von zyklischen Bewegungen





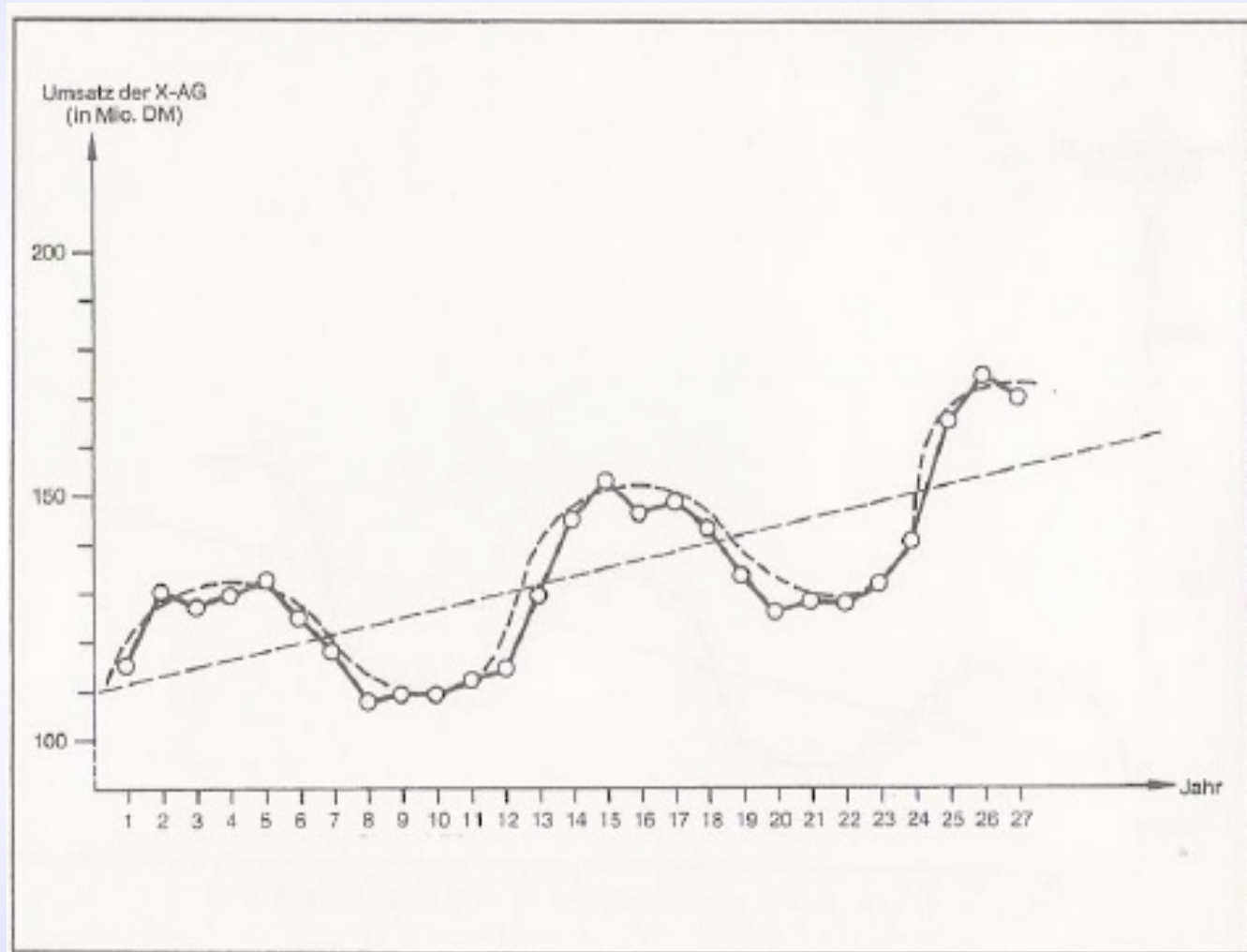
### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- **Elemente d. Zeitreihenanalyse**

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Einfluss der irregulären Komponente



### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- **Elemente d. Zeitreihenanalyse**

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## **Komponenten der Zeitreihe (II)**

- Wenn Komponenten vom Trendniveau der Zeitreihe unabhängig sind  
-> additive Überlagerung von Trend (tk)- Konjunktur- (kk), Saison- (sk) und irregulärer Komponente (ik)

$$y = tk + kk + sk + ik$$

wobei  $tk+kk=gk$  (glatte Komponente)

- Wenn Komponenten mit steigendem Trendniveau zunehmen, d. h. sich proportional zum Trend verhalten  
-> multiplikative Verbundenheit von Trend (tk)- Konjunktur- (kk), Saison- (sk) und irregulärer Komponente (ik)

$$y = tk \cdot kk \cdot sk \cdot ik$$

## Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- **Elemente d. Zeitreihenanalyse**

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Saisonbereinigung

- Saisonschwankungen sind zyklische Bewegungen, die sich im Jahresrhythmus ziemlich regelmäßig wiederholen und ein typisches Bild für jeden Jahresabschnitt zeigen
- mögliche Ursachen:
  - Feiertage
  - Witterung
  - Von der Jahreszeit abhängige Nachfragewellen
- Verfahren zur Eliminierung dieser Saisonkomponente
  - > **Methode der gleitenden Durchschnitte**
  - Beispiel siehe nächste Folie**
- Um Informationsverluste an den Enden der Reihe zu vermeiden, muss man typische Saisoneinflüssen für gleichnamigen Jahresabschnitte berechnen
- Differenz zwischen Ursprungswerten der Zeitreihe und jeweils zugehörigen Saisonkomponenten
  - > **Saisonbereinigte Werte**

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- **Elemente d. Zeitreihenanalyse**

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Gleitende Dreier-Mittel für die Originalbeobachtungen des Beispiels

Beobachtungsjahr	Zeitreihe der Original Beobachtungswerte	Abgeleitete Zeitreihe der gleitenden Dreier-Mittel	Zugeordnet dem Beobachtungsjahr
1	115	$\frac{115 + 130 + 127}{3} = 124$	
2	130	$\frac{130 + 127 + 130}{3} = 129$	2
3	127	$\frac{127 + 130 + 131}{3} = 129,3$	3
4	130		4
5	131	$\frac{130 + 131 + 124}{3} = 128,3$	5
6	124		

## **Deskriptive Statistik**

Typisierung und Darstellung von Daten

Mittelwerte

Streuungsmaße

Indexpzahlen

Korrelation und Regression

## **Elemente d. Zeitreihenanalyse**

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

Kombinatorik

Wahrscheinlichkeitsrechnung

Theoretische

Verteilungen

Schluss von d. Stichprobe auf d. Grundgesamtheit

# **Methode der gleitenden Durchschnitte**

Hierdurch wird eine Glättung der Kurve der ursprünglichen Beobachtungswerte zu einer Trendlinie erreicht.

Aus mehreren aufeinanderfolgenden Beobachtungswerten wird ein arithmetisches Mittel gebildet.

Das arithmetische Mittel wird als abgeleiteter Beobachtungswert demjenigen Beobachtungsjahr zugeordnet, das in Folge der Originalbeobachtungen genau in der Mitte steht, aus denen das arithmetische Mittel berechnet wurde.

Nachteile dieser Methode:

Führt nicht zu einer mathematisch fassbaren Funktion, sondern zu einer geglätteten, aber unregelmäßigen Bewegungslinie. Dadurch ist diese Methode für Prognosezwecke ungeeignet.

Will man die Trendlinie möglichst schwankungsfrei erhalten, muss man die Anzahl der Beobachtungswerte pro Mittelwert immer weiter erhöhen. Dadurch wird aber die Trendlinie immer kürzer und weniger aussagekräftig.

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- **Elemente d. Zeitreihenanalyse**

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## **Bestimmung der Trendkomponente**

- Trendschätzung mit Hilfe der Methode der kleinsten Quadrate, wenn anzunehmen ist, dass Trend linear verläuft.
- Methode der Trendbestimmung ist nichts anderes als der Sonderfall der linearen Einfachregression.  
-> anstelle der x-Werte treten Zeitwerte t
- Vorteil: eine von Schwankungen völlig bereinigte Funktion, die Prognosen der weiteren grundlegenden Entwicklung ermöglicht.

# **STATISTIK**

## **Teil 6**

LB Sabine Gondrom

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- **Indexzahlen**
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Indexzahlen (I)

- Indexzahlen sollen zeitliche Veränderungen mehrerer Größen durch eine einzige Zahl beschreiben.
- Berechnung einer Messzahl, die die Veränderung des Preises eines Gutes beschreibt:

Preis pro Mengeneinheit des Gutes 1 in der Berichtszeit

Preis pro Mengeneinheit des Gutes 1 in der Basiszeit

- zwei beobachtete Merkmalsausprägungen desselben Merkmals verschiedener Zeitpunkte oder –räume werden zueinander ins Verhältnis gesetzt
- Zähler: Ausprägung der Berichtszeit
- Nenner: Ausprägung der Basiszeit



### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- **Indexzahlen**
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Indexzahlen (II)

- Um die Veränderung des Preises aller Güter mittels einer Zahl auszudrücken, müssen die Messzahlen aller Produkte zusammengefasst werden, wobei diese entsprechend ihrer Bedeutung zu gewichten sind.
- Bildung eines gewogenen arithmetischen Mittels dieser Messzahlen -> Indexzahl

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- **Indexzahlen**
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Preisindex

- Die Preismesszahl beschreibt als das Verhältnis von End- zu Ausgangspreis die relative Preisänderung eines Einzelproduktes.
- Um die Preisentwicklung mehrerer Produkte in einer Zahl abzubilden, müssen auch hier die Preismesszahlen der einzelnen Produkte zusammengefasst werden, wobei auch diese entsprechend ihrer Bedeutung zu gewichten sind.

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- **Indexzahlen**
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Preisindex nach Paasche: Berichtsgewichtung

$$P_P = \frac{\sum_{i=1}^n p_{1i} q_{1i}}{\sum_{i=1}^n p_{0i} q_{1i}}$$

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- **Indexzahlen**
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Preisindex nach Laspeyres: Basisgewichtung

$$L_P = \frac{\sum_{i=1}^n p_{1i} q_{oi}}{\sum_{i=1}^n p_{oi} q_{oi}}$$

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- **Indexzahlen**
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Verbraucherpreisindex (VPI) für Deutschland (I)

Der Verbraucherpreisindex misst die durchschnittliche Preisentwicklung aller Waren und Dienstleistungen, die von privaten Haushalten zu Konsumzwecken gekauft werden.

- bildet die Veränderung der Verbraucherpreise umfassend ab
- berücksichtigt werden Güter des täglichen Bedarfs, langlebige Gebrauchsgüter und Dienstleistungen
- Zweck:
  - Beurteilung der Geldwertstabilität/Inflationsmaßstab
  - Wertsicherungsklauseln
  - Grundlage für die Deflationierung von Wertgrößen in der volkswirtschaftlichen Gesamtrechnung

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- **Indexzahlen**
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Verbraucherpreisindex (VPI) für Deutschland (II)

Zur Berechnung des VPI wird der Preisindex nach **Laspeyres** genutzt.

- Vorteil:

- nur noch Erfassung der aktuellen Preise nötig, da in der Vergangenheit festgelegtes Mengengerüst beibehalten wird (= Basisgewichtung)
- bessere Vergleichbarkeit, da Veränderungen der Indizes nicht durch Änderung der Gewichte zustande kommen können

- Nachteil:

- es können nur solche Produkte miteinbezogen werden, die sowohl im Basiszeitraum als auch im Berichtszeitraum produziert bzw. verbraucht wurden

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- **Indexzahlen**
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Mengenindex nach Paasche: Berichtsgewichtung

$$P_M = \frac{\sum_{i=1}^n q_{1i} p_{1i}}{\sum_{i=1}^n q_{0i} p_{1i}}$$

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- **Indexzahlen**
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Mengenindex nach Laspeyres: Basisgewichtung

$$L_M = \frac{\sum_{i=1}^n q_{1i} p_{0i}}{\sum_{i=1}^n q_{0i} p_{0i}}$$



## Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- **Indexzahlen**
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Wertindex

$$WI = \frac{\sum_{i=1}^n p_{1i} q_{1i}}{\sum_{i=1}^n p_{0i} q_{0i}}$$

# **STATISTIK**

## **Teil 7**

LB Sabine Gondrom

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- **Kombinatorik**
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# **Kombinatorik**

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

### **• Kombinatorik**

- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Fakultät und Binomialkoeffizient

## 1. Fakultät

$n!$  -> Abkürzung für das Produkt der ersten  $n$  aufeinander folgenden positiven ganzen Zahlen:

$$n! = 1 \cdot 2 \cdot \dots \cdot (n-1) \cdot n$$

## 2. Binomialkoeffizient

$$\binom{N}{n} = \frac{N!}{n! \cdot (N-n)!}$$

## Permutationen von $n$ verschiedenen Elementen

$n!$

gibt die Anzahl der Permutationen, d. h. die Anzahl der Möglichkeiten der Anordnung von  $n$  Elementen an

## Kombinationen

-> Tabelle der Kombinationen mit/ ohne Wiederholung und mit/ ohne Berücksichtigung der Reihenfolge

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- **Wahrscheinlichkeitsrechnung**
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# **Wahrscheinlichkeitsrechnung**

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- **Wahrscheinlichkeitsrechnung**
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## **Einführung in die Wahrscheinlichkeitsrechnung (I)**

Der Begriff der Wahrscheinlichkeit:

- Spezialgebiet der Mathematik
- versucht den Einfluss des Zufalls auf bestimmte Vorgänge zu erfassen

Bsp.:

Ein Freund wird versucht anzurufen.

Die Leitung kann besetzt oder frei sein. Wenn sie frei ist, kann es sein, dass der Freund nicht zu Hause ist oder das Telefon abgestellt hat ...

- Unternimmt man einen Versuch, dessen Ergebnis (wird im Folgendem als Ereignis bezeichnet) vom Zufall abhängt, dann führt man ein Zufallsexperiment durch.
- Ein Zufallsexperiment muss beliebig oft wiederholt werden können
- Vor Beginn des Zufallsexperiments müssen alle denkbaren Ereignisse des Versuchs angegeben werden können

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- **Wahrscheinlichkeitsrechnung**
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Einführung in die Wahrscheinlichkeitsrechnung (II)

- Die Gesamtheit aller möglichen unterscheidbaren Ereignisse  $x_1, x_2, x_3, \dots, x_n$  eines Zufallsexperimentes kann als endliche Menge

$$\Omega = \{x_1, x_2, x_3, \dots, x_n\}$$

oder bei  $n \rightarrow \infty$  als unendliche Menge

$$\Omega = \{x_1, x_2, x_3, \dots, \}$$

ausgedrückt werden

- Die Menge aller Ereignisse eines Zufallsexperimentes wird als Ereignisraum des Experimentes bezeichnet.
- In einem Ereignisraum ist jedes Element  $x_i$  eine mögliche Realisation des Zufallsexperimentes
- Aus dem Ereignisraum kann eine beliebige Menge  $M \subseteq \Omega$  abgeleitet werden, was als Zufallsereignis bezeichnet wird

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- **Wahrscheinlichkeitsrechnung**
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Einführung in die Wahrscheinlichkeitsrechnung (IV)

- Wenn alle denkbaren Ereignisse des Zufallsexperiments gleichermaßen möglich sind, kann die Wahrscheinlichkeit für das Eintreten eines interessierenden Ereignisses wie folgt berechnet werden:

$$W = \frac{\text{Anzahl der günstigen Ereignisse}}{\text{Anzahl aller gleichmöglichen Ereignisse}}$$



### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- **Wahrscheinlichkeitsrechnung**
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Grundregeln der Wahrscheinlichkeitsrechnung (I)

### Regel 1:

$$W(A \cup B) = W(A) + W(B)$$

-> Spezieller Additionssatz für sich ausschließende Ereignisse

### Regel 2:

$$W(A) = 1 - W(\bar{A})$$

### Regel 3:

$$W(A_1 \cup A_2 \cup \dots \cup A_n) = W(A_1) + W(A_2) + W(A_3) + \dots + W(A_n) = 1$$

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- **Wahrscheinlichkeitsrechnung**
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Grundregeln der Wahrscheinlichkeitsrechnung (II)

### Regel 4:

$$W(A \cup B) = W(A) + W(B) - W(A \cap B)$$

-> Allgemeiner Additionssatz bei Wahrscheinlichkeiten

### Regel 5:

$$W(A/B) = \frac{W(A \cap B)}{W(B)} \quad \text{bzw.} \quad W(B/A) = \frac{W(A \cap B)}{W(A)}$$

### Exkurs: stochastische Unabhängigkeit

### Regel 6:

$$W(A \cap B) = W(A) \cdot W(B)$$

-> Spezieller Multiplikationssatz für (stochastisch) unabhängige Ereignisse

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- **Wahrscheinlichkeitsrechnung**
- Theoretische Verteilungen
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Grundregeln der Wahrscheinlichkeitsrechnung (III)

### Regel 7:

$$W(A \cap B) = W(A) \cdot W(B / A) = W(B) \cdot W(A / B)$$

-> allgemeiner Multiplikationssatz der Wahrscheinlichkeitsrechnung

### Regel 8:

$$W(B) = \sum_{i=1}^n W(A_i) \cdot W(B / A_i)$$

-> Totale Wahrscheinlichkeit

# **STATISTIK**

## **Teil 8**

LB Sabine Gondrom

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# **Theoretische Verteilungen**

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# **Zufallsvariable**

- Die möglichen Ereignisse eines Zufallsexperimentes werden als Elementarereignisse bezeichnet
- Alle Elementarereignisse bilden den Ereignisraum  $\Omega$ .
- Die Elementarereignisse kann man nach bestimmten Regeln reellen Zahlen zuordnen.
- Mathematisch gesehen ist eine derartige Zuordnungsregel eine Funktion, deren Funktionswerte die zugeordneten Zahlen sind.
- Solch eine Funktion wird als Zufallsvariable  $X$  bezeichnet.

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

Für eine weitere Betrachtung ist die Unterscheidung in diskrete und stetige Zufallsvariablen nötig !!

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Wahrscheinlichkeits- und Verteilungsfunktion von diskreten Zufallsvariablen

- Wahrscheinlichkeitsfunktion

$$f(x_i) = W(X = x_i)$$

- Verteilungsfunktion

$$F(x) = W(X \leq x_i) = \sum_{x_i \leq x} f(x_i)$$



### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Dichte- und Verteilungsfunktion von stetigen Zufallsvariablen

- Dichtefunktion

$$f(x)$$

- Verteilungsfunktion

$$F(x) = W(X \leq x_i) = \int_{-\infty}^x f(v) dv$$

## Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Erwartungswert und Varianz von Zufallsvariablen

## Erwartungswert

bei diskreten Zufallsvariablen

$$EX = \sum x_i \cdot f(x_i)$$

bei stetigen Zufallsvariablen

$$EX = \int_{-\infty}^{+\infty} x \cdot f(x) dx$$

## Varianz

bei diskreten Zufallsvariablen

$$VX = \sum_i (x_i - EX)^2 \cdot f(x_i)$$

bei stetigen Zufallsvariablen

$$VX = \int_{-\infty}^{+\infty} (x - EX)^2 \cdot f(x) dx$$

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Diskrete Verteilungen

Verteilung	Binomial	Hypergeometrisch
<b>Merkmalstyp</b>	Diskret	Diskret
<b>Parameter</b>	n Wiederholungen p Eintrittswahrscheinlichkeit	n Wiederholungen N Gesamtzahl M Anzahl mit gewünschter Eigenschaft in der Grundgesamtheit
<b>Wahrscheinlichkeitsfunktion</b>	$W(X = x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$	$W(X = x) = \frac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}}$
<b>Fall</b>	„mit“ Zurücklegen Unabhängigkeit	„ohne“ Zurücklegen Abhängigkeit
<b>Erwartungswert</b>	$E(X) = n \cdot p$	$E(X) = n \cdot \left( \frac{M}{N} \right)$
<b>Varianz</b>	$V(X) = n \cdot p \cdot (1-p)$	$V(X) = n \cdot \left( \frac{M}{N} \right) \cdot \left( 1 - \frac{M}{N} \right) \cdot \frac{N-n}{N-1}$

## Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Stetige Verteilungen - Normalverteilung (I)

- stetige Zufallsvariable
- mit Parametern  $\mu$  und  $\sigma$

Dichtefunktion

$$f(x|\mu; \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Verteilungsfunktion

$$F(x|\mu; \sigma) = \int_{-\infty}^x f(v|\mu; \sigma) dv$$

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Stetige Verteilungen - Normalverteilung (II)

- Eine normalverteilte Zufallsvariable mit  $\mu = 0$  und  $\sigma = 1$  wird als Standardnormalverteilung bezeichnet
- Wert  $x$  einer beliebig-normalverteilten Zufallsvariablen  $X$  mit Mittelwert  $\mu$  und Standardabweichung  $\sigma$  kann durch Transformation in den Wert  $z$  einer standardnormalverteilten Zufallsvariablen  $Z$  transformiert werden.

$$Z = \frac{X - \mu}{\sigma}$$

## Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

# Zentraler Grenzwertsatz

- Liefert die Erklärung für das häufige Auftreten von näherungsweise normalverteilten Zufallsvariablen
- $X_1, X_2, \dots, X_n$  seien unabhängige Zufallsvariablen ...
- Die Summe dieser Zufallsvariablen ist selbst eine Zufallsvariable

$$U = X_1 + X_2 + \dots + X_n$$

- Für U gilt der zentrale Grenzwertsatz -> für große n ist U annähernd normalverteilt
- Im Falle einer annähernden Normalverteilung von U, gilt für die Zufallsvariable:

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Approximationen der Binomial- und hypergeometrischen Verteilung durch die Normalverteilung

- Ziel: Anwendung der Normalverteilung (Rechtfertigung durch Grenzwertsatz)
- Zuvor Überprüfung der Voraussetzungen, bei deren Erfüllung die festgestellte Wahrscheinlichkeitsverteilung durch die Normalverteilung ersetzt (= approximiert) werden kann
- Verbesserung der Approximation durch Stetigkeitskorrektur

### Binomialverteilung -> Normalverteilung

$$np(1 - p) > 9$$

### Hypergeometrischen Verteilung -> Normalverteilung

$$np(1 - p) > 9$$

$$\frac{n}{N} \leq 0,05$$

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Stetige Verteilungen - Chi-Quadrat-Verteilung

- $v$  unabhängige standardnormalverteilte Zufallsvariable

$$Z_1, Z_2, \dots, Z_v$$

- $Z_1^2 + Z_2^2 + \dots + Z_v^2$

-> Zufallsvariable, deren Verteilung Chi-Quadrat-Verteilung mit  $v$  Freiheitsgraden genannt wird

- Für große  $v$  ( $v \geq 100$ ) kann eine mit  $v$  Freiheitsgraden chi-quadrat-verteilte Zufallsvariable als näherungsweise normalverteilt betrachtet werden
- Die Werte werden mit  $\chi^2$  bezeichnet



### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- **Theoretische Verteilungen**
- Schluss von d. Stichprobe auf d. Grundgesamtheit

## Stetige Verteilungen – Studentverteilung (t-Verteilung)

Z -> standardnormaverteilte Zufallsvariable

Y -> chi-quadrat-verteilte Zufallsvariable mit  $\nu$  Freiheitsgraden

$$T = \frac{Z}{\sqrt{\frac{Y}{\nu}}}$$

T -> neue Zufallsvariable, die studentverteilt mit  $\nu$  Freiheitsgraden ist

- Bei  $\nu \geq 30$  kann die Studentverteilung durch die Standardnormalverteilung ersetzt werden

# STATISTIK

LB Sabine Gondrom

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

# **Stichprobenverfahren**

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Zufällige Stichprobenverfahren**

- jedes Element der Grundgesamtheit hat eine bekannte bzw. berechenbare Wahrscheinlichkeit, Element der Stichprobe zu werden
- der Stichprobenfehler lässt sich mathematisch berechnen
- Beispiele für Methoden:
  - Originalverfahren
  - Schlussziffernverfahren
  - Buchstabenverfahren
  - Auswahl nach dem Geburtsdatum
  - Systematische Auswahl nach dem Zufallsstart

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Zufällige Stichprobenverfahren**

<b>Vorteile</b>	<b>Nachteile</b>
Verfahren sind breit einsetzbar.	Relativ hohe Kosten
Mathematisch-statistische Fehlerberechnungen sind ebenso möglich wie die auf dem Wahrscheinlichkeitsprinzip beruhenden Auswertungs- und Testverfahren. Der so genannte Stichprobenfehler lässt sich berechnen.	Probleme der Stichprobenausschöpfung (Erreichbarkeit, Verweigerung, Abbruch). Es sollte auf keinen Fall bei geringer Ausschöpfung versucht werden, die einmal gezogene Stichprobe zu verändern, insbesondere zu vergrößern. Dieses hätte mit dem Zufallsprinzip nichts mehr zu tun.
Kenntnis über die Verteilung relevanter Merkmale in der Grundgesamtheit ist nur soweit von Nöten, als dass sinnvolle Schichten oder Klumpen gebildet werden können.	Ein Verzeichnis über die Grundgesamtheit, das vollständig ist (Aktualitätsproblem) und die Grundgesamtheit abgrenzt, ist notwendig.

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Zufällige Stichprobenverfahren:** **Einfache Zufallsauswahl**

- jedes Element der Grundgesamtheit hat die gleiche Wahrscheinlichkeit, Element der Stichprobe zu werden
- es werden keine Kenntnisse über die Merkmalsstruktur der Grundgesamtheit benötigt
- eine einfache Zufallsauswahl ist immer möglich, aber nicht immer optimal

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Zufällige Stichprobenverfahren:** **Geschichtete Zufallsauswahl (I)**

- Grundgesamtheit wird in Schichten (Teil-Grundgesamtheiten) aufgeteilt bevor Stichprobe gezogen wird
- dafür sind Kenntnisse über die Merkmalsstruktur der Grundgesamtheit Voraussetzung
- Stichprobenziehung wird dann in jeder Schicht mittels eines reinen Zufallsverfahrens realisiert
- demnach hat jedes Element einer Schicht wie jedes andere Element dieser Schicht die gleiche Wahrscheinlichkeit, Element der Stichprobe zu werden
- die Wahrscheinlichkeit eines Elementes der Grundgesamtheit, Element der Stichprobe zu werden, variieren allerdings von Schicht zu Schicht
- Nach der Stichprobenziehung ergibt sich für jede Schicht ein Ergebnis
- diese Ergebnisse der unterschiedlichen Schichten werden vor der Hochrechnung zum Gesamtergebnis gewichtet

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Zufällige Stichprobenverfahren:** **Geschichtete Zufallsauswahl (II)**

- Vorteile der geschichteten Zufallsauswahl
  - Verringerung des Stichprobenfehlers
  - Schichten sind für das Untersuchungsziel von Interesse
  - Sicherstellen relevanter Untergruppen in der Stichprobe (statistische Relevanz)
  - Möglichkeit der Anwendung unterschiedlicher Untersuchungsdesigns pro Schicht



### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Zufällige Stichprobenverfahren:** **Klumpenverfahren**

- Sonderform der Stichprobenziehung
- die Grundgesamtheit wird in Klumpen (= natürliche Konglomerate von Untersuchungseinheiten) aufgeteilt
- per Zufall wird eine bestimmte Anzahl von Klumpen ausgewählt und mit allen ihren Elementen in die Stichprobe aufgenommen
- somit verkörpern nicht einzelne Elemente der Grundgesamtheit die Stichprobe, sondern ganze Gruppen von Elementen
- Vorteil: Verringerung des Erhebungsaufwandes
- Nachteil: Gefahr eines Klumpeneffekts

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Zufällige Stichprobenverfahren:** **Mehrstufige Zufallsauswahl**

- hierbei wird eine Reihe von Auswahlverfahren auf mehreren Stufen hintereinander durchgeführt
- aus der Grundgesamtheit wird in einem ersten Schritt eine Zwischenstichprobe gezogen, deren enthaltene Einheiten wieder die Grundgesamtheit für die nächste Auswahlstufe darstellen
- Vorteil: geringerer organisatorischer, finanzieller und zeitlicher Aufwand

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Nicht Zufällige Stichprobenverfahren**

- bei fehlenden Voraussetzungen für eine Zufallsauswahl, dies bedeutet, dass das Prinzip der Chancengleichheit und des Zufalls nicht möglich ist
- behilft man sich mit einer entscheidungsorientierten, bewussten Auswahl
- es wird trotzdem versucht, die Verteilung der Merkmalsausprägungen so zu gestalten, dass sie sich der Grundgesamtheit annähert und demnach repräsentativ ist

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Nicht Zufällige Stichprobenverfahren**

<b>Vorteile</b>	<b>Nachteile</b>
Verfahren haben sich in der Praxis bewährt und verursachen nur einen geringen Aufwand (/Kosten, Zeit).	Ausgeprägteres Fälscherproblem und Interviewereinfluss bei der Stichprobenziehung.
Probleme der Stichprobenausschöpfung (Erreichbarkeit, Verweigerung, Abbruch) sind nicht gegeben, da keine von vornherein festgelegte Stichprobe vorliegt. Es wird so lange fortgefahren, bis die gewünschte Anzahl der Stichprobenelemente vorliegt.	Mathematisch-statistische Fehlerberechnungen sind streng genommen nicht möglich. Der so genannte Stichprobenfehler lässt sich nicht berechnen.
Ein vollständiges Verzeichnis über die Grundgesamtheit ist nicht notwendig.	Die Kenntnis über die Verteilung relevanter Merkmale in der Grundgesamtheit ist gegebenenfalls notwendig, um die Stichprobe entsprechend der Merkmalsverteilung bilden zu können.

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Nicht Zufällige Stichprobenverfahren:** **Willkürliche Auswahl**

- z. B. bei Interviews an einer Strasse, denen dem Passanten befragt werden
- dabei kein Zufall mit Chancengleichheit vorhanden, da nur die Personen in die Stichprobe gelangen können, welche die Strasse passieren
- oft in der explorativen bzw. psychologischen Marktforschung angewandt

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Nicht Zufällige Stichprobenverfahren:** **Quotenauswahl**

- Verteilungen von wesentlichen Merkmalen in der Grundgesamtheit werden dem Interviewer als Quote vorgegeben
- innerhalb dieser Quote kann Interviewer die Auskunftsperson frei wählen

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Nicht Zufällige Stichprobenverfahren:** **Monographische Auswahl/ typische Auswahl**

- aus der Gesamtmenge der möglichen Elemente wird eine Anzahl herausgegriffen, von der man annimmt, sie sei besonders typisch (angenommene Repräsentativität)
- von diesem Ergebnis wird auf die Gesamtmenge geschlossen
- Beispiel: Testmärkte

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Nicht Zufällige Stichprobenverfahren:** **Konzentrationsprinzip/ Abschneideverfahren**

- die Elemente der Grundgesamtheit werden vor der Stichprobenziehung nach ihrer Wichtigkeit für den Untersuchungszweck sortiert
- es werden die Elemente aussortiert, von denen erwartet wird, dass sie nur wenig zum Ergebnis beitragen



### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Nicht Zufällige Stichprobenverfahren:** **Staffelungsmethode**

- alle Elemente der Grundgesamtheit werden nach einem Merkmal, z. B. Umsatz, geordnet
- als Stichprobe wird ein Segment aus dieser Ordnung verwendet

### **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Nicht Zufällige Stichprobenverfahren:** **Schneeballsystem**

- wenn Grundgesamtheit relativ klein, aber nur schwer abgegrenzt werden kann
- Ausgangspunkt: kleine Stichprobe
- jedes Element wird daraufhin überprüft, ob Beziehungen zu anderen Elementen bestehen, die ebenfalls zur Grundgesamtheit gehören könnten
- auf diese Weise wächst die Stichprobe wie ein Schneeball an

## **Deskriptive Statistik**

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

## **Wahrscheinlichkeitsrechnung und schließende Statistik**

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## **Stichprobenausschöpfung (I)**

- Ausfälle von Elementen in der Stichprobe
  - qualitätsneutrale Ausfälle:  
Haben auf die Güte der Stichprobe keinen Einfluss, da sie bei der Stichprobenanlage bereits ausgeschlossen werden, da z.B.:
    - Haushalt nicht existiert
    - Wohnung unbewohnt ist
    - Unternehmen nicht existiert
  - qualitätsrelevante Ausfälle:  
Alle Ausfälle, bei denen keine Befragung durchgeführt werden kann, obwohl es sich um potenzielle Stichprobenelemente handelt, da z. B.:
    - Person die Teilnahme am Interview verweigert
    - Person Interview abbricht
    - Zielperson trotz mehrerer Versuche nicht angetroffen werden kann

### Deskriptive Statistik

- Typisierung und Darstellung von Daten
- Mittelwerte
- Streuungsmaße
- Indexzahlen
- Korrelation und Regression
- Elemente d. Zeitreihenanalyse

### Wahrscheinlichkeitsrechnung und schließende Statistik

- Kombinatorik
- Wahrscheinlichkeitsrechnung
- Theoretische Verteilungen
- **Schluss von d. Stichprobe auf d. Grundgesamtheit**

## Stichprobenausschöpfung (II)

- Berechnung der Stichprobenausschöpfung

	Anzahl der Fälle	In %
Brutostichprobe	900	100
- Qualitätsneutrale Ausfälle	45	5
= Nettostichprobe/ Bereinigte Stichprobe	855	100
- Qualitätsrelevante Ausfälle	171	20
= Durchgeführte Interviews	684	80
- nicht ausgewertete Interviews	17	2,5
= Ausgewertete Interviews/ Ausschöpfungsquote	667	77,5