



Human Centered Multimedia
Institute of Computer Science

UNA Universität
Augsburg
University

Auswahl statistischer Tests

Ilhan Aslan, Chi Tai Dang, Björn Bittner,
Katrín Janowski, Elisabeth André



Human Centered Multimedia
Institute of Computer Science
Augsburg University
Universitätsstr. 6a
86159 Augsburg, Germany

■ Ziel:

- Wir wollen untersuchen, ob ein Muster A eher vom Nutzer erkannt wird als ein Muster B.

■ Methode 1:

- Zeige 100 Leuten Muster A und Muster B und bitte sie, anzugeben, welches Muster ihrer Meinung nach schneller erkannt wird.
- Ergebnis: Daten mit drei Häufigkeiten:
 - Anzahl der Leute, die angeben, dass sie Muster A schneller erkennen
 - Anzahl der Leute, die angeben, dass sie Muster B schneller erkennen
 - Anzahl der Leute, die angeben, dass es keinen Unterschied gibt

■ Datenauswertung:

- Anwendung eines χ^2 -Tests zur Überprüfung, ob sich die beobachteten Häufigkeiten von der erwarteten Gleichverteilung unterscheiden.
- Ist das Ergebnis statistisch signifikant und sagen die meisten Leute, dass eines der Muster besser erkannt wird, dann war der Test informativ.
- ABER:
 - Keine Information über Ausmaß der Verbesserung
 - Keine Information über die Homogenität der Leute, die angeben, dass eines der Muster besser erkannt wird.

■ Ziel:

- Wir wollen untersuchen, ob ein Muster A eher vom Nutzer erkannt wird als ein Muster B

■ Methode 2:

- Zeige 100 Leuten Videos mit den beiden Mustern und bitte sie, jedes Muster auf einer 7-Punkte-Skala zu bewerten.
- Mit dieser Methode erhalten wir für jede Versuchsperson zwei Bewertungen:
 - Bewertung für Muster A
 - Bewertung für Muster B

■ Datenauswertung:

- Anwendung eines t-Tests zur Überprüfung, ob die Bewertungen für die beiden Muster signifikant unterschiedlich sind.
- Mit einer kleinen Änderung im experimentellen Design erhalten wir nun wesentlich mehr Information
- Beispiel: Ausmaß der Verbesserung durch deskriptive Statistiken

Mögliche Abwandlungen:

- Bisher rein subjektive Äußerungen
- Alternative:

Messung der Zeit bis die unterschiedlichen Muster erkannt werden.

- Sofern sich Erkennbarkeit anhand von Reaktionszeiten messen lässt, erhalten wir damit eine direktere und objektivere Messung.
- Darüber hinaus haben wir nun die Bedingungen für einen parametrischen Test erfüllt.
 - repeated-measures t-Test für die Reaktionszeiten

Weitere Abwandlungen:

- Sofern wir eine ähnliche Anzahl von Männern und Frauen haben, können wir das Geschlecht als weitere unabhängige Variable hinzunehmen.
 - Gibt es Unterschiede zwischen Männern und Frauen bei der Erkennung von Mustern?
- Sofern wir eine genügend große Anzahl von Personen unterschiedlichen Alters haben, können wir die Altersgruppe als weitere unabhängige Variable hinzunehmen.
 - Gibt es Unterschiede zwischen Personen unterschiedlicher Altersgruppen bei der Erkennung von Mustern?

- Beide Tests untersuchen eine Beziehung zwischen zwei Variablen, aber sie tun es auf unterschiedliche Weise.
- Beispiel: Untersuche Zusammenhang zwischen Alter und Furcht
- Frage: Welche Art von Daten liefern die Teilnehmer?
 - Eine Anzahl?
 - Ein Paar von Bewertungen?
- Anzahl: Teilnehmer werden Kategorien zugeordnet (alt oder jung, furchtlos oder ängstlich) => zweidimensionale Chi²-Test
- Bewertungen: Teilnehmer liefern sowohl für ihr Alter als auch für ihre Furcht eine Bewertung. => Korrelation

- Annahme:
 - Wir haben Paare von Bewertungen.
- Frage:
 - Sollten wir eine Korrelationsanalyse oder einen t-Test durchführen?
- Dies hängt davon ab, welche Fragestellung wir untersuchen wollen.
 - Frage: Existiert ein **Zusammenhang** zwischen den Bewertungen?
=> Korrelationsanalyse
 - Frage: **Unterscheiden** sich die Bewertungen?
=> t-Test

Häufigkeiten oder Bewertungen?

Häufigkeiten

Bewertungen

Anzahl der unabhängigen Variablen

Experimentelles oder korrelatives Design?

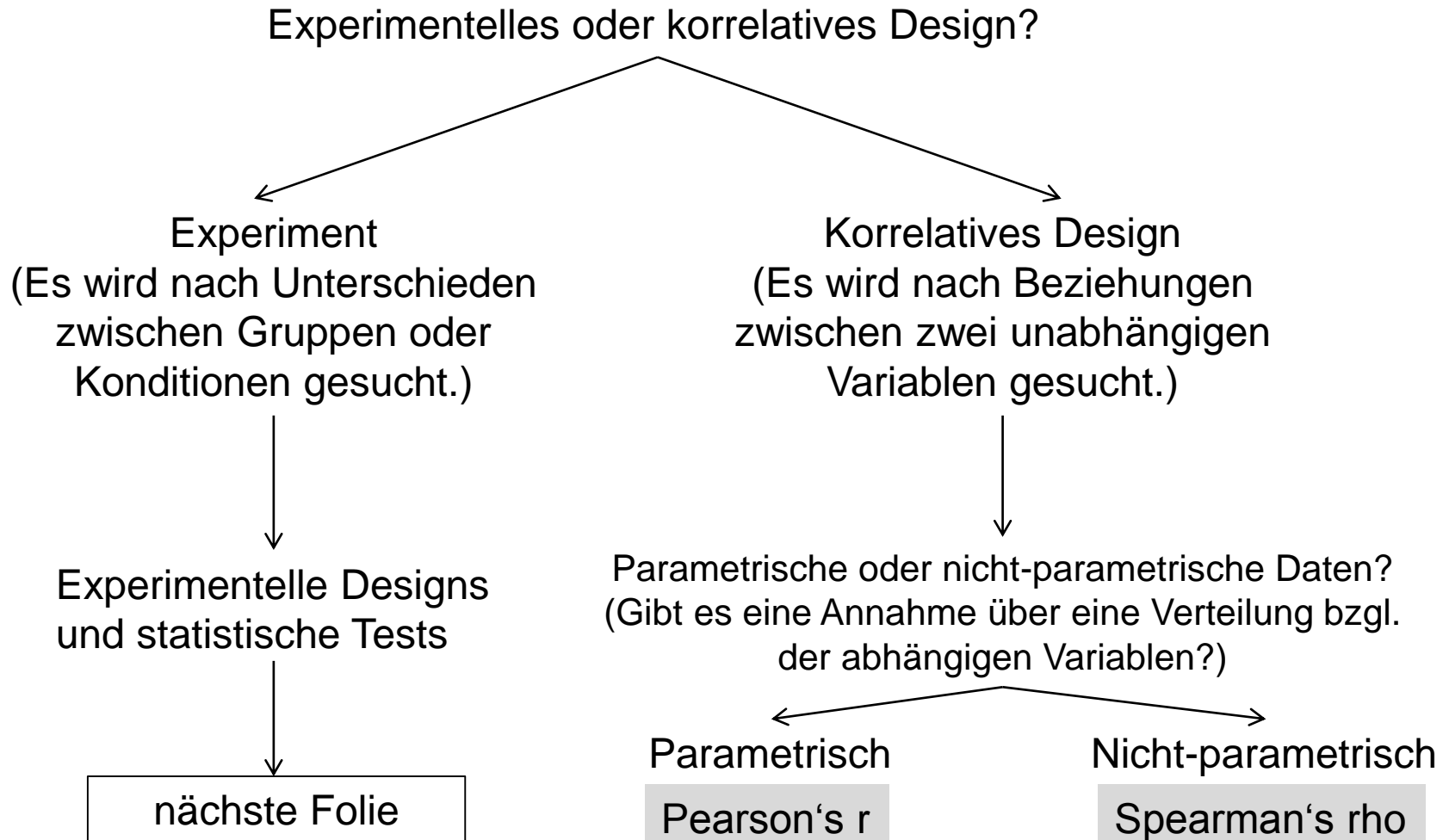
eine

zwei

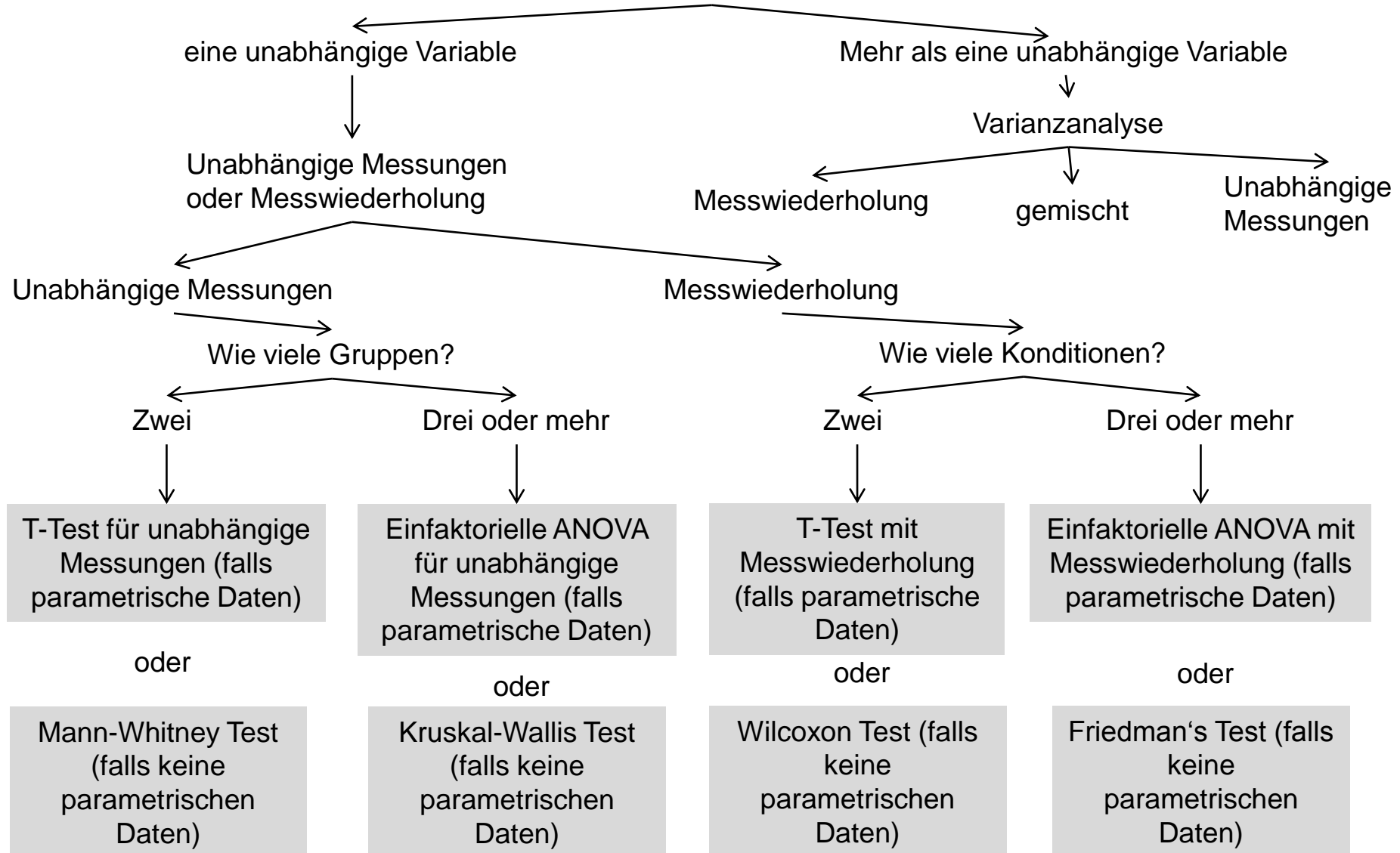
Eindimensionaler
Chi²-Test

Zweidimensionaler
Chi²-Test

nächste Folie



Experimentelle Designs und statistische Tests (Fortsetzung letzte Folie)



- Ein HCI-ler möchte den **Zusammenhang** herausfinden zwischen dem **Studienfach** und der **Präferenz für ein graphisches versus textuelles Interface**.
- Er befragt dazu **50 Studierende der Informatik** und **50 Studierende der Physik**, welches von zwei Interfaces sie bevorzugen.
 - Geben Sie die UV/AV an!
 - Beschreiben Sie die Daten der AV/UV (z.B. Handelt es sich um nominal-, ordinal- oder intervallskalierte Daten?)
 - Was wird untersucht? Häufigkeiten oder Bewertungen?
 - Um was für ein Studiendesign handelt es sich und warum? (z.B. korrelatives, experimentelles?)
 - Bestimmen Sie für die Auswertung der Daten die geeigneten statistischen Mittel!

- Ein HCI-ler möchte den **Zusammenhang** herausfinden zwischen dem **Alter** und der **Reaktionszeit beim Drücken eines Buttons**.
- Er erfasst dazu das **Alter von 100 Leuten** und lässt diese **mit dem Interface interagieren**.
 - Geben Sie die UV/AV an!
 - Beschreiben Sie die Daten der AV/UV (z.B. Handelt es sich um nominal-, ordinal- oder intervallskalierte Daten?)
 - Werden Häufigkeiten oder Bewertungen untersucht?
 - Um was für ein Studiendesign handelt es sich und warum? (z.B. korrelatives, experimentelles?)
 - Bestimmen Sie für die Auswertung der Daten die geeigneten statistischen Mittel!

- Die **Qualitätssicherung der Uni Augsburg** möchte herausfinden, in welchen von **zwei Vorlesungen** Studierende **besser abschneiden**.
- Sie erfasst dazu die Noten von Studierenden in den beiden Vorlesungen.
 - Geben Sie die UV/AV an!
 - Beschreiben Sie die Daten der AV/UV (z.B. Handelt es sich um nominal-, ordinal- oder intervallskalierte Daten?)
 - Werden Häufigkeiten oder Bewertungen untersucht?
 - Um was für ein Studiendesign handelt es sich und warum? (z.B. korrelatives, experimentelles?)
 - Bestimmen Sie für die Auswertung der Daten die geeigneten statistischen Mittel!