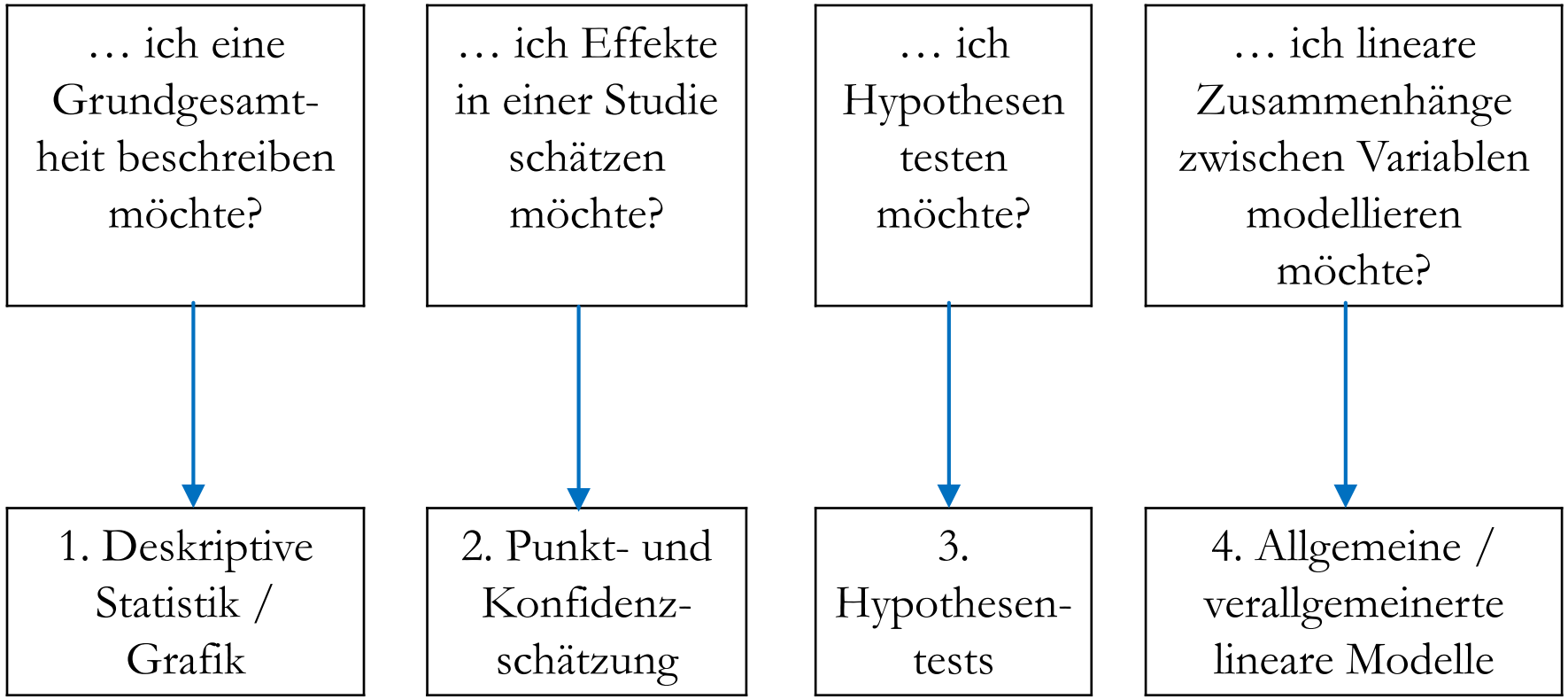
"Statistik von oben"

-

Wiederholung einmal anders

Was mache ich, wenn …

Was mache ich wenn …



1. Deskriptive Statistik

1.1 metrische Variablen

Verteilung prüfen:

- Histogramm
- (empirische Verteilungsfunktion)

1.2 kategoriale Variablen

- statistische Kenngrößen:
  - absolute Häufigkeiten
  - relative Häufigkeiten
- Grafik
  - (ggf. gruppiertes) Balkendiagramm
  - (Kreisdiagramm)

1.1 Metrische Größen

1.1.1 kleine Fallzahl oder extreme Werte (schiefe Verteilung)

- Median
- Quartile

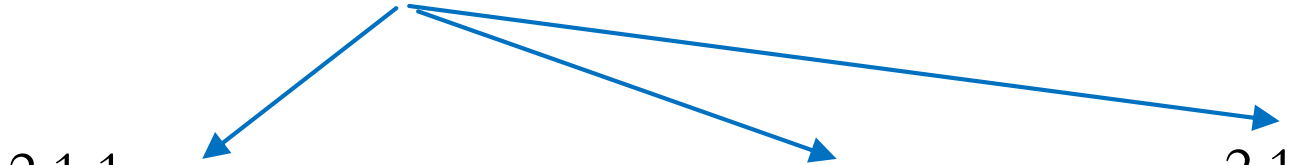
Grafik:

- Boxplot

1.1.2 moderate / große Fallzahl, ~ symmetrische Verteilung

- Mittelwert
- Std.-Abweichung

2.1 Effekte einer Variablen schätzen



2.1.1
Zählvariablen
(nominal / ordinal

2.1.2
Ordinale Variablen

2.1.3
Metrische (kontinuierliche)
Variablen

geschätzte Wsk. $\hat{p}$
= relative Häufigkeit

mit $(1-\alpha)$ $C.I.$

Mittelwert:
$\hat{\mu} = \bar{x}$

mit $(1-\alpha)$ $C.I.$

2.2 Gruppenunterschiede schätzen

2.1.1 Zählvariablen (nominal / ordinal)

- Differenz relativer Häufigkeiten

mit $(1-\alpha)\ C.I.$

- Odds Ratio

mit $(1-\alpha)\ C.I.$

2.1.2 Ordinale Variablen

- mediane Differenz

mit $(1-\alpha)\ C.I.$

- A (stochastic superiority)

2.1.3 Metrische (kontinuierliche) Variablen

- Mittelwertdifferenz:

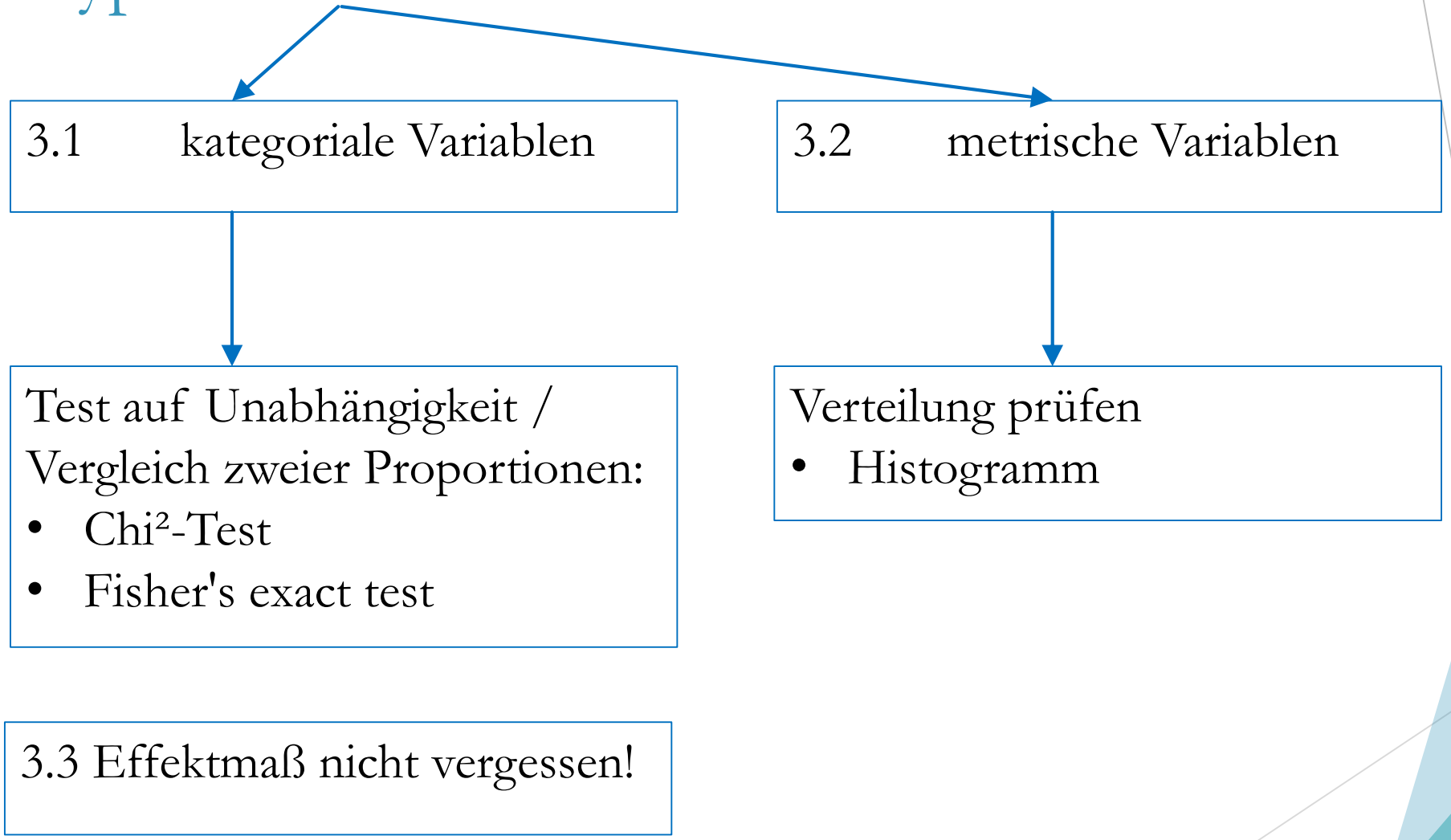
$$\Delta\hat{\mu} = \overline{x_1} - \overline{x_2}$$

mit $(1-\alpha)\ C.I.$

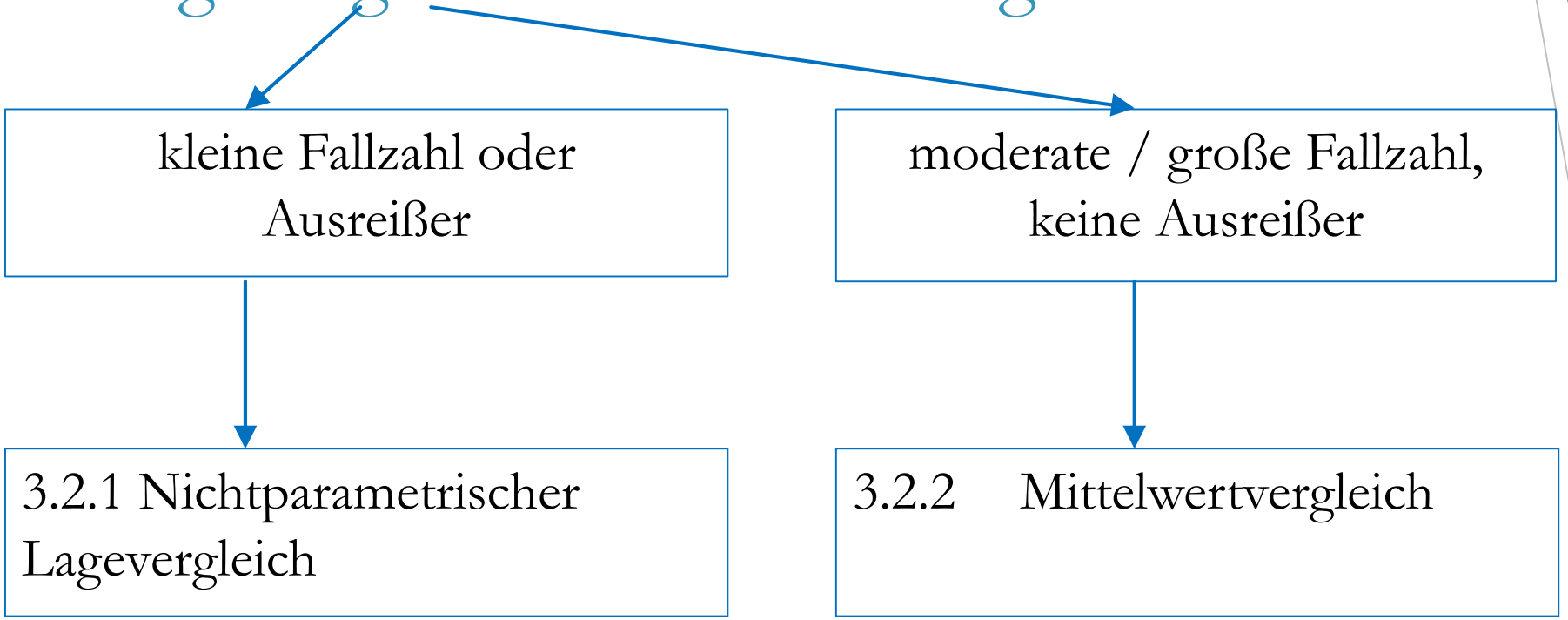
- SMD

(standardised mean difference)

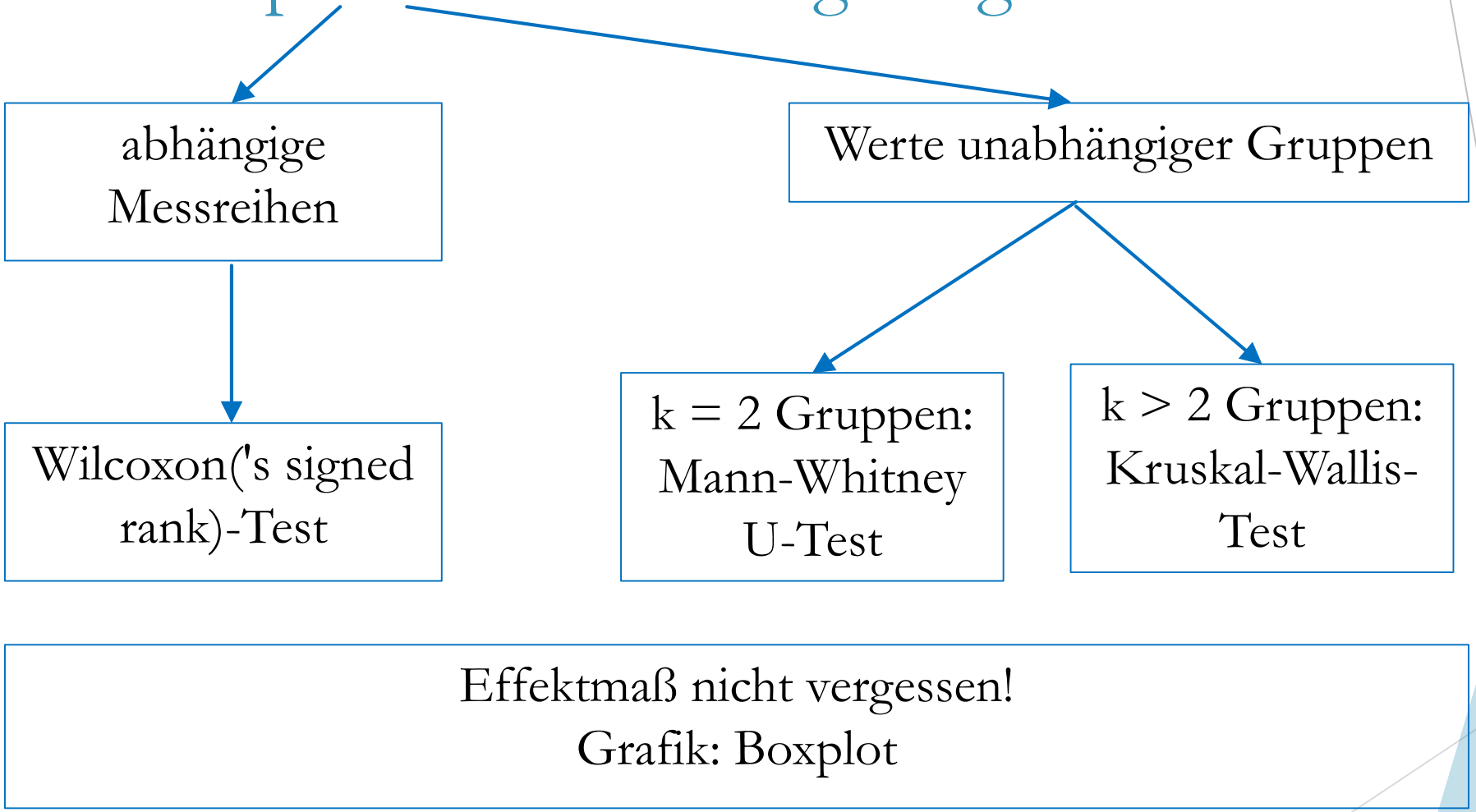
3. Hypothesen testen



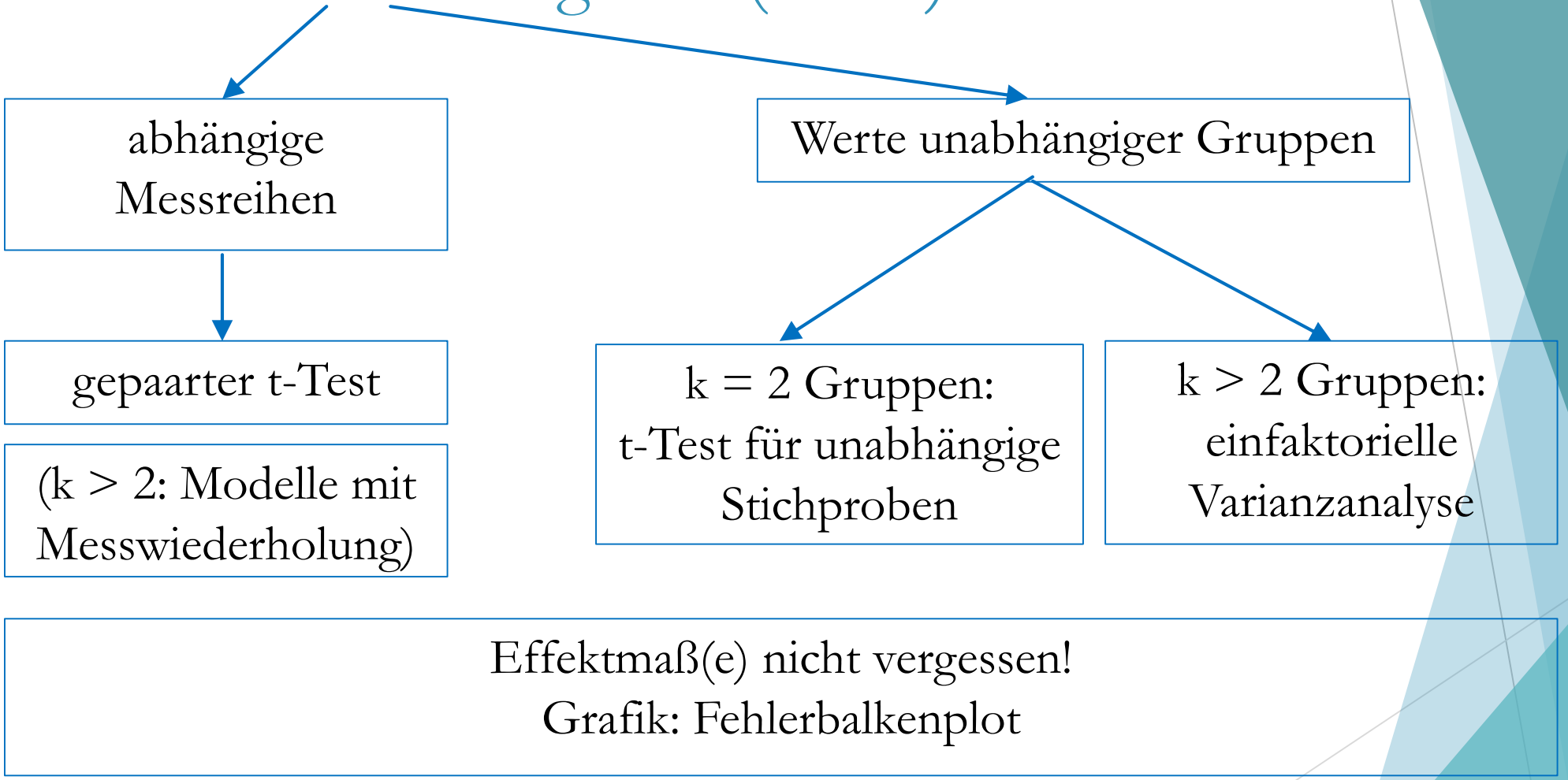
3.2 Lagevergleich von Verteilungen



3.2.1 Nichtparametrischer Lagevergleich



3.2.2 Mittelwertvergleich (k = 2)



3.3 Effektmaße (Auswahl)

Metrisch (symmetrische Verteilung)	Metrisch (mit Ausreißern, kleine Stichproben)	Kategorial (nominal / ordinal)	
▶ Mittelwert-differenz	Mediane Differenz (Hodges-Lehmann)	Differenz relativer Häufigkeiten	absolute Effektmaße
▶ SMD (standardis. Mittelw.-diff.)	A (stochastic superiority)	Odds Ratio Risikoverhältnis	relative Effektmaße

4 Modellierung von Zusammenhängen

4.1 abhängige Variable **binär**

4.2 abhängige Variable: **metrisch**

Unabhängige Variablen

- kategorial: Faktoren
- metrisch: Kovariaten

4 Modellierung von Zusammenhängen

4.1 Logistische Regression

Faktoren:

- als "kategorial" deklarieren
- Kontrast / Referenz-kategorie festlegen

4.2 Allgemeines lineares Modell

Faktoren:

- ANOVA: --
- Regression: Indikator-variablen bilden

4 Modellierung von Zusammenhängen

4.1 Logistische Regression	4.2 Allgemeines lineares Modell
Parameter (Effektschätzer) • als Odds Ratio interpretieren	Parameter (Effektschätzer) • Konstante (Ref.-Kategorie) • Anstieg
• kategorial: OR gegenüber Referenzgruppe • metrisch: OR pro 1 Einheit	• kategorial: Mittelwert-differenz geg. Referenz • metrisch: Anstieg pro 1 Einheit

Statistische Prinzipien 1

Folgende Prinzipien der Statistik haben wir kennengelernt:

1. Wissenschaft (auch die math. Statistik) arbeitet mit Modellen, also Abstraktionen / Vereinfachungen der Wirklichkeit
2. Statistik modelliert Zusammenhänge (Assoziationen). Wie diese erzeugt werden (Kausalität), darüber gibt Statistik keine Auskunft.
3. Um aus Assoziationen auf kausale Aussagen zu schließen, müssen (z. B. durchs Studiendesign) alle möglichen Erklärungen für den Effekt (Confounder) ausgeschlossen werden.

Statistische Prinzipien 2

Folgende Prinzipien der Statistik haben wir kennengelernt:

4. Für wachsende n strebt die relative Häufigkeit gegen die Wahrscheinlichkeit eines Ereignisses
5. Ist n die Fallzahl der Stichprobe, so verändert sich die Genauigkeit des Mittelwertschätzers mit dem Faktor $1/\sqrt{n}$.

Folgerungen

- Extreme Ergebnisse sind vor allem in kleinen Stichproben häufig.
- Jeder noch so kleine Effekt kann "signifikant gemacht" werden, wenn die Fallzahl nur groß genug ist.

Statistische Prinzipien 3

Folgende Prinzipien der Statistik haben wir kennengelernt:

6. Hypothesentests arbeiten mit dem Prinzip der Falsifizierung. H_0 und H_A sind nicht symmetrisch, insbesondere ist "keine Signifikanz" nicht unbedingt ein Indiz für die Gültigkeit von H_0.
7. Statistische Signifikanz und fachwissenschaftliche Relevanz sind zwei (u. U. sehr unterschiedliche) Dinge, deshalb
   - außer dem Testergebnis (p-Wert) immer auch einen Schätzer für den Effekt angeben,
   - prüfen, ob der Effekt fachwissenschaftlich auch interessant ist.

Statistische Prinzipien 4

Folgende Prinzipien der Statistik haben wir kennengelernt:

8. Beim Hypothesentest wird man im "long-run" die Fehler 1. und 2. Art α und β beobachten.
   - Im konkreten Test trifft man eine Entscheidung ohne sicher zu wissen, was die Wahrheit ist.
   - Welcher Fehler der schlimmere ist, hängt von der Situation ab.
   - Je mehr Tests durchgeführt werden, desto mehr akkumulieren sich die Fehler
     - Korrektur für multiples Testen
     - Offenlegen, was alles getestet wurde (kein Data dredging)

Statistische Prinzipien 5

Zufallsprozesse können wir ausnutzen

- um z. B. mit Bootstrap Kenngrößen (z. B. C.I.) zu schätzen
- um Situationen zu simulieren, die zu komplex für Formeln sind
- um Probanden (nahezu) zufällig Gruppen zuzuordnen (Randomisation)
- …

Viel Erfolg bei der Prüfung!