

## **Varianzanalyse:**

### **allg. Sinn und Zweck:**

Untersuchung des Effekts einer oder mehrerer unabhängiger Variablen (Vergleich von Gruppenmittel-werten und Varianzen)

### **Voraussetzungen:**

- Normalverteilung der Fehlerkomponenten in der Grundgesamtheit ( $X^2$ -Test der Güte der Anpassung) / normalverteilte Daten (Intervallskala)
- Varianzhomogenität, d.h. Varianz ist für alle Bedingungen gleich und somit vom Mittelwert unabhängig (F-Test)  
psychologische Bedeutung: unkontrollierte Daten, die die Varianz der Daten erzeugen dürfen nicht an die kontrollierten UV gekoppelt sein
- Unabhängigkeit der Fehlerkomponenten, d.h. Additivität der Treatment- und Fehlereffekte  
Verletzungen der Voraussetzungen evtl. bei genügend großer Stichprobe.

### **Strukturgleichung:**

$$x_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$$

Haupteffekte und Interaktionseffekte (1.Ordnung, ...)

### **fixed versus random:**

fixed: Aussage nur über experimentell realisierte Faktorstufen → diskrete Aussage

random: experimentell realisierte Faktorstufen als Zufallsauswahl aller Faktorstufen, über die man eine Aussage machen möchte → kontinuierliche Aussage

### **Typen:**

fixed Modell: nur feste Faktoren

random Modell: nur zufällige Faktoren

gemischtes Modell: feste und zufällige Faktoren

### **replicated versus repeated measurements:**

replicated: wiederholte Messungen an verschiedenen Personen → unabhängige Werte

repeated: wiederholte Messungen an einer Person → abhängige Werte (keine Varianzhomogenität)

Vorteil: Fehlervarianz ist um die interindividuelle Varianz reduziert; weniger  $V_{pn}$

Nachteil: andere statistische Auswertung

### **Interpretation der Signifikanzen:**

- bei mehr als 2 Stufen nur unspezifische Aussagen (globale Effekte)

- Signifikanz von Haupteffekten:

gemittelt über Gruppen (Normale und Alkoholiker - quantitative Unterscheidung)

gemittelt über Skalen (nicht unbedingt sinnvoll, wenn Skalen unterschiedl. Eigenschaften repräsentieren)

### **Varianzanalyse von Häufigkeiten:**

#### Verteilung:

- Unabhängigkeit der Häufigkeiten + Anzahl der Verhaltensweisen festgelegt und WS für jede Verhaltensweise konstant (sonst: bedingte WS) → Binomialverteilung  $E(x) = np$  und  $VAR(x) = np(1-p)$ , d.h. Varianz ist an den Mittelwert geknüpft und somit keine Varianzhomogenität (Vernachlässigung für Mittelwerte  $.45 < np < .65$ ) → für großes n gegen Normalverteilung
- abhängige Häufigkeiten → hypergeometrische Verteilung

#### Analysemöglichkeiten:

- Varianzanalyse: evtl.
- $X^2$ -Test der Güte der Anpassung → stoch. Unabhängigkeit? (Voraussetzung: Normalverteilung; Kontingenztabelle der Grundgesamtheit)
- $X^2$ -Wert zur Testung der Existenz / Signifikanz des Zusammenhangs bzw. der Interaktion zwischen UV und ihrer Wirkung auf die AV
- mehrdimensionale Kontingenztafeln → Binomialverteilung und log-lineare Modelle: Häufigkeiten werden auf eine Weise ausgedrückt, die der Strukturgleichung der Varianzanalyse entspricht:

Unabhängigkeit:  $p(A_i \cap B_j) = p(A_i) \cdot p(B_j)$

Logarithmus:  $\log p(A_i \cap B_j) = \log p(A_i) + \log p(B_j)$

Varianzanalyse:  $x_{ij} = \mu + \mu_i + \mu_j + e_{ij}$

### **a posteriori Tests:**

nur wenn keine spezifischen Hypothesen formuliert werden konnten!

Scheffé-Test: Überprüfung, welche Einzelvergleiche signifikant sind

- robust gegen Verletzungen der Voraussetzungen
- tendenziell konservative Entscheidung ( $H_0$ )  $\rightarrow$   $\alpha$ -Fehler niedrig

### **Regressionsrechnung:**

#### **Voraussetzungen:**

- Normalverteilung
- Intervallskalenniveau oder dichotome Nominalverteilung

#### **Multiple Regression:**

$$y_i = b_0 + b_1x_{i1} + b_2x_{i2} + \dots + e_i$$

(y = Kriterium; x = Prädiktor; b = Regressionsgewicht)

Klein-z-Transformation:  $\beta$ -Gewichte

#### **Interpretation der Regressionsgewichte:**

- stat. Unabhängigkeit der Prädiktoren: Anteile der Variablen  $x_i$  am Kriterium y  
(zusätzlich: gleiche Fehlervarianz der Prädiktoren)  
aber: ein Prädiktor läßt sich in diesem Fall nicht durch die anderen vorhersagen
- stat. Abhängigkeit der Prädiktoren: Korrelation des Prädiktors mit dem Kriterium nach Auspartialisierung der restlichen Prädiktoren

#### **Suppressoreffekte (konfundierende Variable):**

Korrelation mit der konfundierenden Variable separat erfassen und Effekt auspartialisieren:

$$r_{12.3} = r_{12} \rightarrow x_3 \text{ ohne Bedeutung}$$

$$|r_{12.3}| > |r_{12}| \rightarrow x_3 \text{ ist ein Suppressor}$$

$$|r_{12.3}| < |r_{12}| \rightarrow x_3 \text{ täuscht einen Effekt vor}$$

$\Rightarrow$  Prädiktor ohne Kriteriumskorrelation kann moderierend auf den Einfluß eines anderen Prädiktors wirken.

(evtl. auch Mischung aus Längsschnitt und Querschnitt bei Kohorteneffekten)

#### **Kreuzvalidierung:**

2 Teilstichproben (even-odd)  $\rightarrow$  2 multiple Regressionsgleichungen

- gegenseitige Kriteriumsvorhersage
  - Korrelation zwischen der vorhergesagten Kriterien und Kriterien der Eichstichprobe
- $\Rightarrow$  Überprüfung der Stabilität der Regressionsgewichte (evtl. praktischer Einsatz)

### **Faktorenanalyse:**

#### **Allgemeiner Sinn und Zweck:**

Extraktion von unabhängigen Bewertungsdimensionen (Faktoren), die in minimaler Anzahl einen maximalen Varianzanteil der Ergebnisse ausklären.

Im Prinzip multiple Regression, wobei die Prädiktoren extrahiert werden = latente Variablen (miteinander unkorreliert).

$\rightarrow$  Datenreduktion

$\rightarrow$  heuristisches, hypothesengenerierendes Verfahren

$\rightarrow$  Überprüfung der Dimensionalität komplexer Merkmale

#### **Voraussetzungen:**

- Normalverteilung
- Intervallskala

### **Strukturgleichung:**

$$x_{mi} = a_{i1} \cdot f_{m1} + a_{i2} \cdot f_{m2} + \dots + a_{iq} \cdot f_{mq} + e_{mi}$$

$x_{mi}$  = Leistung der Person m beim Test i

$a_{ij}$  = Bedeutung des j-ten Faktors für den Test i = Faktorladung

$f_{mj}$  = Ausstattung der Person m mit dem Faktor j = Faktorwert

q = Anzahl der Faktoren

i = Laufindex der n Tests

j = Laufindex der q Faktoren

m = Laufindex der Personen

Psychologische Bedeutung der Unabhängigkeit von  $f_{mj}$  und  $a_{ij}$  :

die j-te Dimension ist für jede Person im Ausmaß  $a_{ij}$  im i-ten Test enthalten.

### **Bedeutung von Faktorladungen und Faktorwerten:**

Faktorladung: Bedeutung des Faktors j für eine Aufgabe i ( $a_{ij}$ )

Faktorwert: Ausstattung einer Person m mit dem Faktor j ( $f_{mj}$ )

### **Personenraum und Testraum:**

graphische Darstellung:

Personen bzw. Test als Achsen eines Koordinatensystems → Punktwolke (Scatterplot)

Personenraum:

bei Kenntnis der Grunddimensionen können Personen aufgrund ihrer Scores im Faktorraum repräsentiert werden.

Def. des Faktorraumes als euklidischer Vektorraum: Ähnlichkeiten zwischen Personen als Differenz der Vektoren.

### **Eindeutigkeit der Lösung:**

- theoretisch läßt sich die Anzahl der Faktoren so weit erhöhen, bis die Varianz aller Variablen vollständig erklärt sind.

Allgemein wird vorher rel. willkürlich abgebrochen!

(unbedeutende Varianzanteile können auch auf fehlerhafte unsystematische Effekte zurückgehen)

- nur wieviele (nicht welche) Faktoren, d.h. die Benennung erfolgt relativ willkürlich!

- die Bedeutung der einzelnen Variablen ist abhängig von der Rotation!

### **Erklärung / Beschreibung:**

Die Faktorenanalyse ist eine Korrelationstechnik:

- keine Überprüfung von inhaltlichen Hypothesen

- keine kausalen Schlußfolgerungen

→ hypothesengenerierendes Verfahren

### **Kanonische Korrelation:**

Zusammenhang zwischen mehreren Prädiktorvariablen und mehreren Kriteriumsvariablen:

- getrennte Faktorisierung der Prädiktor- und Kriteriumsvariablen

- Rotation des jeweils 1. Faktors, so daß die Interkorrelation (kanonische Korrelation) maximal wird usw.

⇒ max. Kovarianzaufklärung zwischen Kriterium und Prädiktor

### **Allgemein:**

#### **Mittelung von Messungen:**

→ evtl. qualitativ anderes Ergebnis

Mittelung beinhaltet eine starke Homogenitätsannahme

- bei Profilen (implizit bei Berechnung von Korrelationen)

Sinnvoll:

- Annäherung an  $\mu$

- homogene Kurvenverläufe

Nomothetischer Ansatz in der Therapieforschung:

- Vorhersage
- Vergleich mit Normen

**Längsschnitt / Querschnitt:**

- große systematische Unterschiede (Bsp.: Gedächtnis)
- Längsschnitt: Therapieerfolg; Interventionskontrolle, ....
- Querschnitt: evtl. Kohorteneffekte

**Methoden:**

- Mittelwertsvergleiche (Varianzanalyse)
- Häufigkeitsvergleiche
- korrelative Studien (Faktorenanalyse)

**X<sup>2</sup>-Verteilung:**

Summe stochastisch unabhängiger normalverteilter Werte

**F-Bruch:**

Quotient aus 2 X<sup>2</sup>-Verteilungen