Psychologische Testtheorie

Sitzung 7

Skalierung II

We are happy to share our materials openly:
The content of these Open Educational Resources by Lehrstuhl für Psychologische Methodenlehre und Diagnostik, Ludwig-Maximilians-Universität München is licensed under CC BY-SA 4.0. The CC Attribution-ShareAlike 4.0 International license means that you can reuse or transform the content of our materials for any purpose as long as you cite our original materials and share your derivatives under the same license.

Wahr-oder-Falsch Aussagen zur letzten Sitzung

1. Ein psychologischer Test gilt als skalierbar, wenn die Zuordnung der Messwerte zu den Personen auf der Basis eines theoretisch plausiblen testtheoretischen Modells erfolgt.
2. Auch wenn ein psychologischer Test keinem testtheoretischen Modell folgt, kann der Test dennoch eine zufriedenstellende Genauigkeit (d.h., Reliabilität) aufweisen.
3. Die Annahmen der testtheoretischen Modelle können wir direkt mithilfe statistischer Hypothesentests überprüfen.
4. In einem psychologischen Test, welcher dem τ-*äquivalenten* Modell folgt, müssen alle Items den gleichen Erwartungswert aufweisen.
5. Angenommen es wurde nachgewiesen, dass ein psychologischer Test dem τ-äquivalenten Modell folgt. Um zu prüfen, ob für diesen Test auch das *parallele* Modell gilt, müssen die Varianzen seiner Items untersucht werden.
6. Falls die Kovarianz zweier beliebiger Itempaare eines Tests nicht identisch ist, kann für diesen Test nicht das τ-*kongenerische* Modell gelten.

Überblick

Sitzung	Datum	Thema	Themenblock
1	14.10.24	Einführung	Begriffe, Modellierung von Antwortverhalten durch Zufallsvariablen & mathematische Grundlagen der Testtheorie
2	21.10.24	Wahrscheinlichkeitstheoret. Grundlagen	
3	28.10.24	Testtheoretische Modelle I	Testtheoretische Modelle
4	04.11.24	Testtheoretische Modelle II	
5	11.11.24	Testtheoretische Modelle III	
6	18.11.24	Skalierung I	
7	25.11.24	Skalierung II	

➛ In der heutigen Vorlesung beschäftigen wir uns mit statistischen Methoden zur Überprüfung der Folgerungen aus den Modellannahmen und wenden diese auf einen Beispielfragebogen an. Außerdem befassen wir uns mit der Parameterschätzung für eindimensionale testtheoretische Modelle.

3. Statistische Hypothesen

Wiederholung: Folgerungen aus den Modellannahmen

- In der letzten Vorlesung haben wir gesehen, dass sich aus jedem testtheoretischen Modell Folgerungen für die **Erwartungswerte**, **Varianzen** und **Kovarianzen** der Items ableiten lassen
- Für das **parallele Modell** ergab sich zum Beispiel:

$E(X_i) = E(X_j)$ für alle Items i, j

$VAR(X_i) = VAR(X_j)$ für alle Items i, j

$COV(X_i, X_j) = COV(X_o, X_u)$ für alle Itempaare i, j *und* o, u

- Diese Folgerungen werden wir mithilfe von statistischen Hypothesentests überprüfen

Statistische Hypothesen I

- Es gibt mehrere Möglichkeiten, die statistischen Hypothesen zu formulieren

→ **Möglichkeit 1:**

- Eine Möglichkeit besteht darin, für jede Folgerung ein Hypothesenpaar aus **Nullhypothese** und **Alternativhypothese** aufzustellen
- z.B., im parallelen Modell:

$H_0: E(X_i) = E(X_j)$ für alle Items i, j

$H_1: E(X_i) \neq E(X_j)$ für mindestens ein Itempaar i, j

$H_0: VAR(X_i) = VAR(X_j)$ für alle Items i, j

$H_1: VAR(X_i) \neq VAR(X_j)$ für mindestens ein Itempaar i, j

$H_0: COV(X_i, X_j) = COV(X_o, X_u)$ für alle Itempaare i, j *und* o, u

$H_1: COV(X_i, X_j) \neq COV(X_o, X_u)$ für mindestens ein Paar von Itempaaren i, j *und* o, u

- Zur Überprüfung dieser statistischen Hypothesen müssten dann drei Hypothesentests durchgeführt werden

→ Möglichkeit 2:

- Eine weitere Möglichkeit besteht darin, **Omnibushypothesen** aufzustellen, sodass alle Folgerungen eines Modells durch einen einzigen Hypothesentest getestet werden können:

  H_0: Die Folgerungen aus den Modellannahmen sind erfüllt

  H_1: Mindestens eine der Folgerungen ist nicht erfüllt

- Ein „Nachweis“ der Skalierung besteht in diesem Fall einfach aus der Annahme der Nullhypothese für eines der Modelle
- Der Omnibustest weist gegenüber den einzelnen Hypothesenpaaren eine höhere Power auf
- Wir werden daher im Folgenden nur noch den Omnibustest betrachten
- Der Omnibustest wird häufig auch einfach „Modelltest“ genannt

Bemerkung

- Es gibt neben dem Omnibustest noch zahlreiche weitere statistische Methoden zur Überprüfung der Folgerungen aus den Modellannahmen
- Die wichtigsten Methoden werden wir in der nächsten Vorlesung kennenlernen

4. Anwendung

Anwendungsbeispiele

- Wir werden nun das Vorgehen bei der Überprüfung der Skalierung anhand von einem **Fragebogen (NEO-FFI)** und einem **Leistungstest (TAP)** darstellen
- Den R-Code zur Durchführung der Hypothesentests werden wir in dieser Vorlesung nicht besprechen
- Besprochen wird die Durchführung mit R im Seminar Fragebogenentwicklung
- Wir werden hier (und in der Klausur) nur Outputs interpretieren

4. Anwendung

Beispiel 1: NEO-FFI (Extraversion)

Ostendorf & Angleiter (2004)

Extraversions-Skala des NEO-FFI

- Latente Variable: Extraversion
- 12 Items mit einer Antwortskala von 0 (starke Ablehnung) bis 3 (starke Zustimmung), ohne Mittelkategorie (neutral)
- Negativ (d.h., umgekehrt) gepolte Items* sind im Datensatz schon umkodiert

NEO-FFI	Item	NEO-PI-R Facette	Trennschärfe
Item N2	gerne viele Leute um sich	Geselligkeit (E2)	.56
Item N7	leicht zum Lachen zu bringen	Frohsinn (E6)	.45
Item N12	*nicht besonders fröhlich	Frohsinn (E6)	.51
Item N17	gerne mit anderen unterhalten	Herzlichkeit (E1)	.52
Item N22	gern im Mittelpunkt stehen	Erlebnishunger (E5)	.38
Item N27	*vorziehen, Dinge alleine zu tun	Geselligkeit (E2)	.33
Item N32	vor Energie überschäumen	Aktivität (E4)	.41
Item N37	fröhlicher Mensch	Frohsinn (E6)	.62
Item N42	*kein gut gelaunter Optimist	Frohsinn (E6)	.51
Item N47	hektisches Leben	Aktivität (E4)	.15
Item N52	aktiver Mensch sein	Aktivität (E4)	.52
Item N57	*lieber eigene Wege gehen	Durchsetzungsfähigkeit (E3)	.32

Datensatz Beispiel 1

- *n* = 1449 Personen haben auf die zwölf Items geantwortet
- Ausschnitt aus dem Datensatz:

	n2	n7	n12	n17	n22	n27	n32	n37	n42	n47	n52	n57
1	2	2	2	3	1	3	2	2	2	2	2	1
2	2	2	3	3	2	2	1	2	2	1	3	2
3	3	2	2	3	1	2	1	2	2	1	2	3
4	2	2	2	3	2	2	1	1	0	1	2	2
5	1	1	3	2	1	1	2	1	2	1	2	1
6	2	2	1	3	1	1	1	1	1	1	1	1
7	2	2	1	2	3	1	2	1	1	3	2	1
8	2	2	1	3	2	1	3	2	1	2	2	1
…												

- Die Antwortkategorien waren: 0 = „starke Ablehnung, 1 = „Ablehnung“, 2 = „Zustimmung“, 3 = „starke Zustimmung“
- Abweichend vom Original-Fragebogen, wurden aufgrund einer Forschungsfrage keine Mittelkategorie („neutral“) vorgegeben

Vorgehen

- Aus der Theorie ergibt sich ein **eindimensionales Modell** mit der **latenten Variable „Extraversion“**
- Deshalb überprüfen wir zunächst, ob die Folgerungen des am wenigsten strengen eindimensionalen Modell erfüllt sind: das τ-kongenerische Modell
- Falls dies der Fall ist, könnten wir im Anschluss die restlichen eindimensionalen Modelle überprüfen, denn eventuell kann auch ein strengeres Modell angenommen werden

τ-kongenerisches Modell

- Hypothesen:
  - H_0: Die Folgerungen aus dem **τ-kongenerischen Modell** sind erfüllt
  - H_1: Mindestens eine Folgerung aus dem **τ-kongenerischen Modell** ist nicht erfüllt
- R-Output:

```
Model Test User Model:

  Test statistic                                918.730
  Degrees of freedom                                 54
  P-value (Chi-square)                            0.000
```

- **Testentscheidung**: Auf Basis der Daten entscheiden wir uns **gegen** die Geltung des τ-kongenerischen Modells, da **$p = 0.000 < 0.05 \rightarrow H_1$**

Vorgehen

- Nach der Ablehnung des τ-kongenerischen Modells könnten wir an dieser Stelle eigentlich aufhören und müssten im nächsten Schritt das mehrdimensionale τ-kongenerische Modell prüfen (Vorlesung #08)
- Alle weiteren eindimensionalen Modelle treffen strengere Annahmen als das τ-kongenerische Modell und ziehen zusätzliche (strengere) Folgerungen nach sich (z.B. Varianzgleichheit im essentiell parallelen Modell)
- Wenn schon die Folgerungen aus dem τ-kongenerische Modell nicht zutreffen, dann können auch die Folgerungen der strengeren Modelle nicht zutreffen
- Wir werden die Tests der anderen Modelle jedoch der Vollständigkeit halber noch durchführen

Essentiell τ-äquivalentes Modell

- Hypothesen:
  - H_0: Die Folgerungen aus dem **essentiell τ-äquivalenten Modell** sind erfüllt
  - H_1: Mindestens eine Folgerung aus dem **essentiell τ-äquivalenten Modell** ist nicht erfüllt
- R-Output:

```
Model Test User Model:

  Test statistic                              1444.115
  Degrees of freedom                                65
  P-value (Chi-square)                           0.000
```

- **Testentscheidung**: Auf Basis der Daten entscheiden wir uns **gegen** die Geltung des essentiell τ-äquivalenten Modells, da **$p = 0.000 < 0.05 \rightarrow H_1$**

τ-äquivalentes Modell

- Hypothesen:
  - H_0: Die Folgerungen aus dem **τ-äquivalenten Modell** sind erfüllt
  - H_1: Mindestens eine Folgerung aus dem **τ-äquivalenten Modell** ist nicht erfüllt
- R-Output:

```
Model Test User Model:

  Test statistic                              4738.033
  Degrees of freedom                                76
  P-value (Chi-square)                           0.000
```

- **Testentscheidung**: Auf Basis der Daten entscheiden wir uns **gegen** die Geltung des τ-äquivalenten Modells, da **$p = 0.000 < 0.05 \rightarrow H_1$**

Essentiell paralleles Modell

- Hypothesen:
  - H_0: Die Folgerungen aus dem **essentiell parallelen Modell** sind erfüllt
  - H_1: Mindestens eine Folgerung aus dem **essentiell parallelen Modell** ist nicht erfüllt
- R-Output:

```
Model Test User Model:

  Test statistic                              1610.789
  Degrees of freedom                                76
  P-value (Chi-square)                           0.000
```

- **Testentscheidung**: Auf Basis der Daten entscheiden wir uns **gegen** die Geltung des essentiell parallelen Modells, da **$p = 0.000 < 0.05 \rightarrow H_1$**

Paralleles Modell

- Hypothesen:
  - H_0: Die Folgerungen aus dem **parallelen Modell** sind erfüllt
  - H_1: Mindestens eine Folgerung aus dem **parallelen Modell** ist nicht erfüllt
- R-Output:

```
Model Test User Model:

  Test statistic                                5043.530
  Degrees of freedom                                  87
  P-value (Chi-square)                             0.000
```

- **Testentscheidung**: Auf Basis der Daten entscheiden wir uns **gegen** die Geltung des parallelen Modells, da **$p = 0.000 < 0.05 \rightarrow H_1$**

Zusammenfassung Beispiel 1

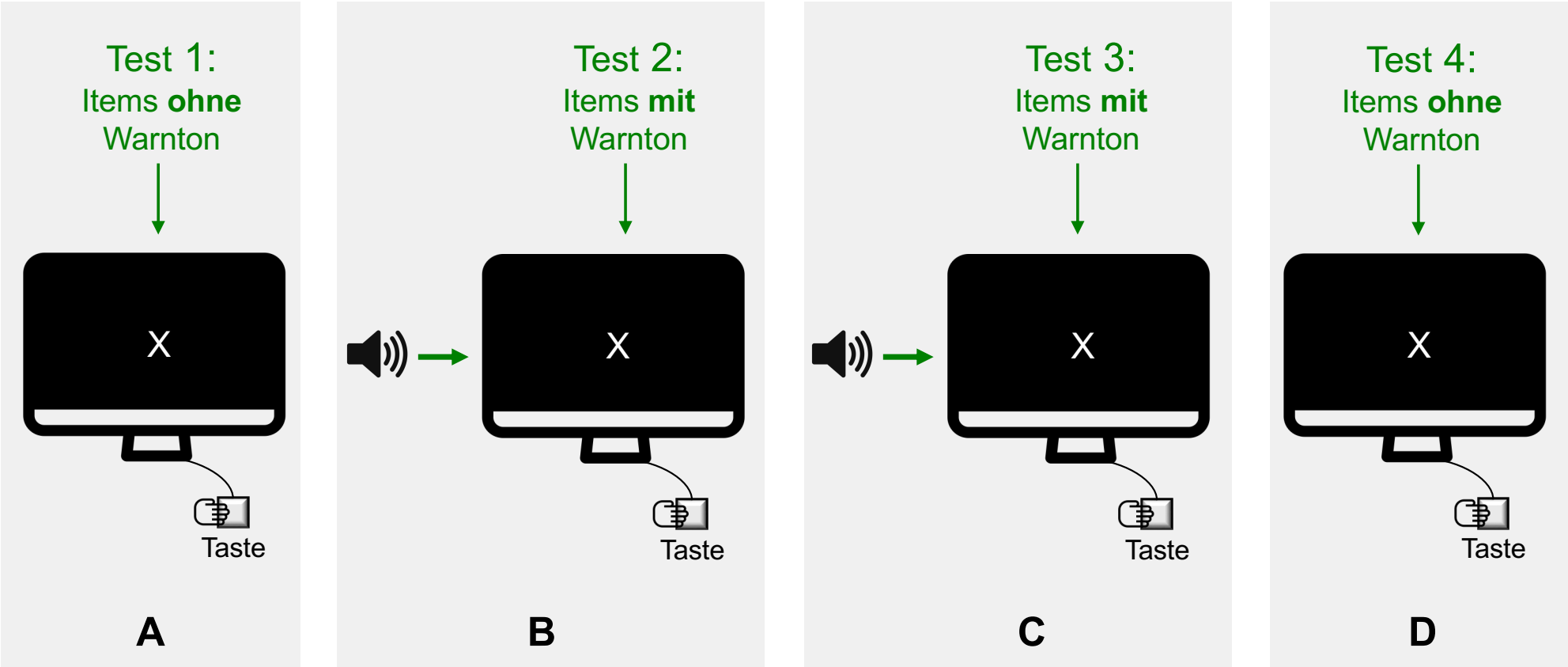
- Alle eindimensionalen Modelle wurden abgelehnt
- Wir müssen somit nach einem Modell suchen, das weniger streng ist, z.B. das mehrdimensionale τ-kongenerische Modell (Vorlesung #08)
- Wir können auf der Basis der Daten davon ausgehen, dass nicht eine einzelne latente Variable das Antwortverhalten der Personen in diesem Fragebogen bestimmt
- Für die Dimension „Extraversion“ des NEO-FFI kann also das Gütekriterium der Skalierung zunächst <u>nicht</u> angenommen werden
- Aber: Die Stichprobe ist mit $n = 1449$ sehr groß und der Modelltest reagiert sehr sensitiv auf kleine Abweichungen von einem perfekten Modell

4. Anwendung

Beispiel 2: Leistungstest TAP

Zimmermann & Fimm (2009)

TAP


Test 1:
Items **ohne** Warnton
X
Taste
A
Test 2:
Items **mit** Warnton
X
Taste
B
Test 3:
Items **mit** Warnton
X
Taste
C
Test 4:
Items **ohne** Warnton
X
Taste
D

Datensatz Beispiel 2

- *n* = 136 Personen haben auf die Items reagiert (Angaben in Millisekunden)
- Ausschnitt aus dem Datensatz:

	almd1	almd2	almd3	almd4
1	200.0	217.0	217.0	202.0
2	281.0	246.5	258.0	243.0
3	226.0	225.5	195.5	220.0
4	295.0	327.5	309.5	338.0
5	294.0	275.0	239.0	249.0
6	289.5	350.0	335.0	306.0
7	387.0	320.5	301.0	324.0
8	226.5	277.0	241.0	305.0
9	229.0	218.0	246.0	246.0
10	230.0	220.0	218.0	232.0
11	285.0	238.0	283.0	332.5
12	352.0	266.0	305.0	230.5
13	259.0	273.5	259.5	268.0

…

Vorgehen

- Auch hier ergibt sich aus der Theorie ein **eindimensionales Modell**, diesmal mit der **latenten Variable „Alertness“**
- Deshalb prüfen wir wieder zunächst, ob die Annahmen des am wenigsten strengen eindimensionalen Modells erfüllt sind: das τ-kongenerische Modell

τ-kongenerisches Modell

- Hypothesen:
  - H_0: Die Folgerungen aus dem **τ-kongenerischen Modell** sind erfüllt
  - H_1: Mindestens eine Folgerung aus dem **τ-kongenerischen Modell** ist nicht erfüllt
- R-Output:

```
Model Test User Model:

  Test statistic                                24.062
  Degrees of freedom                                 2
  P-value (Chi-square)                           0.000
```

- **Testentscheidung**: Auf Basis der Daten entscheiden wir uns **gegen** die Geltung des τ-kongenerischen Modells, da **$p = 0.000 < 0.05 \rightarrow H_1$**

Zusammenfassung Beispiel 2

- Mit dem τ-kongenerischen Modell wurde das am wenigsten strenge eindimensionale Modell abgelehnt, sodass auch alle strengeren eindimensionalen Modelle nicht angenommen werden können
- Wir können somit nicht davon ausgehen, dass eine (einzige) latente Variable „Alertness“ hinter dem Antwortverhalten der Personen steht
- Weiteres Vorgehen: Wir werden in der nächsten Vorlesung mithilfe der Faktorenanalyse untersuchen, ob wir für den TAP Test ein psychologisch interpretierbares mehrdimensionales Modell annehmen können:
- Vielleicht stehen hinter den Itemantworten statt einer latenten Variable „Alertness“ mehrere „Alertness-Komponenten“, z.B.:
  - Alertness ohne Warnton
  - Alertness mit Warnton
- Das würde auch die Theorie hinter dem Test wiederspiegeln
- Tatsächlich werden wir in der nächsten VL sehen, dass die Daten doch für das **essentiell τ-äquivalente Messmodell** im TAP-Test Alertness sprechen

5. Anmerkungen zu den Modelltests

Anmerkungen zu den Modelltests I

- Was tun, wenn mehrere Modelle auf Basis der Modelltests angenommen wurden und weiterverwendet werden könnten?
- Zwei Möglichkeiten:
  - Wahl des am wenigsten strengen nicht abgelehnten Modells
  - Wahl des strengsten nicht abgelehnten Modells
- Hier gibt es keine allgemeingültige Antwort, aber zwei wichtige Überlegungen:
  - Bei der Überprüfung der zusätzlichen Modellannahmen für jedes strengere Modelle besteht immer das Risiko für **Fehlentscheidungen**
  - Für strengere Modelle müssen weniger Parameter geschätzt werden, d.h., sie haben eine niedrigere **Komplexität** (nach Occam‘s Razor wünschenswert)
- Es gibt statistische Methoden, welche die Komplexität der Modelle bei der Beurteilung ihrer Modellpassung berücksichtigen (Vorlesung #08)
- In der Praxis kann die Wahl zwischen mehreren angenommenen Modellen vom Anwendungsfall (z.B. der Stichprobengröße) abhängen

Anmerkungen zu den Modelltests II

- Wir haben für den NEO-FFI mehrere Hypothesentests durchführt (einen Hypothesentest pro Modell)
- Damit haben wir ein multiples Testproblem und müssten eigentlich durch eine Korrektur der p-Werte oder des α-Niveaus die Wahrscheinlichkeit für den Fehler 1. Art bezogen auf die gesamte Testfamilie kontrollieren
- Aber: Da bei den Modelltests im Gegensatz zu den meisten aus Statistik I und II bekannten Hypothesentests die Nullhypothese die „Wunschhypothese“ darstellt (sie besagt ja, dass das Modell gilt), ist der hier eigentlich entscheidende und zu kontrollierende Fehler der **Fehler 2. Art**
- Es müsste also eigentlich **die Power kontrolliert werden** (im Sinne einer höheren Power für die einzelnen Tests). Dies kann jedoch nur durch **größere Stichproben** geschehen
- Eine strengere Kontrolle des Fehlers 1. Art würde zwar die Wahrscheinlichkeit für mindestens einen Fehler 1. Art verringern, die Wahrscheinlichkeit für mindestens einen Fehler 2. Art jedoch erhöhen. Daher sollten die p-Werte im Rahmen der Überprüfung der Modellgeltung eher nicht korrigiert werden!

Anmerkungen zu den Modelltests III

- Da die Annahmen der testtheoretischen Modelle nicht direkt getestet werden können, sondern nur Folgerungen aus ihnen, können die Modelle nicht „nachgewiesen“, sondern **lediglich falsifiziert** werden
- Falls wir uns in einem der Modelltests für die H_1 entscheiden (Folgerungen sind nicht erfüllt), können wir (mit einer gewissen Fehlerwahrscheinlichkeit) daraus schließen, dass das Modell nicht gilt
- **Falls wir uns jedoch für die H_0 entscheiden (Folgerungen sind erfüllt), können wir daraus <u>nie</u> schließen, dass das Modell gilt**
- Aus den Annahmen können nämlich noch weitere Folgerungen außer den hier besprochenen abgeleitet werden, die auch bei Geltung dieser noch verletzt sein können
- Beispielsweise könnte trotz Geltung der Folgerungen bezüglich der Erwartungswerte, Varianzen und Kovarianzen der Items immer noch die „Personenhomogenität“ verletzt sein

Anmerkungen zu den Modelltests IV

- Was tun, wenn keines der Modelle gilt?
- Hinter den Testmodellen steht inhaltlich gesehen eine zusammengesetzte Hypothese:
  - (A) Die latente Variable existiert und
  - (B) kann durch die Items des vorliegenden Tests erfasst werden und
  - (C) der Zusammenhang zwischen Items und latenter Variable ist wie im Modell spezifiziert.
- Wenn alle Testmodelle durch die Hypothesentests abgelehnt werden, kann (A), (B) und / oder (C) falsch sein.

→ Wir wissen zunächst nicht, woran es liegt.

Anmerkungen zu den Modelltests V

- Was tun, wenn keines der Modelle gilt?
- Zusammengesetzte Hypothese:
  - (A) Die latente Variable existiert und
  - (B) kann durch die Items des vorliegenden Tests erfasst werden und
  - (C) Zusammenhang zwischen Items und latenter Variable ist wie im Modell.
- Vorgehen:
  - Überprüfen, ob es an (C) liegt: Komplexere (z.B. nonlineare) Modelle ausprobieren (Achtung: teilweise extrem große Stichproben nötig)
  - Überprüfen, ob es an (B) liegt: eventuell geben die Modelltests Aufschluss über Verbesserungsmöglichkeiten in Bezug auf den Test (z.B. könnte sich zeigen, dass einzelne Items Probleme verursachen)
  - Falls weder (C) noch (B) der Grund für die Ablehnung des Modells sind: Revision der Theorie bezüglich (A)
  - Wenn keine Auflösung möglich: Das Modell nehmen, das am besten passt (Vorlesung #08)

6. Parameterschätzung

Parameterschätzung

- Im Zuge der Überprüfung der Modellgeltung haben wir uns (im besten Fall) für eines der testtheoretischen Modelle entschieden
- Um mit diesem Modell weiterarbeiten zu können, müssen wir jedoch noch die in ihm vorkommenden unbekannten Parameter schätzen
- Zu schätzende Parameter sind <u>je</u> <u>nach</u> <u>Modell</u>:
  - Die **Varianzen $\boldsymbol{VAR(\varepsilon_i)}$ der Fehlervariablen** in allen Modellen
  - Die **Itemparameter $\boldsymbol{\sigma_i}$** im essentiell parallelen und essentiell τ-äquivalenten Modell und in den τ - kongenerischen Modellen
  - Die **Steigungsparameter $\boldsymbol{\beta_i}$** in den τ - kongenerischen Modellen
  - **Erwartungswert $\boldsymbol{E(\theta)}$** und **Varianz $\boldsymbol{VAR(\theta)}$** der latenten Variable(n)
- Die Schätzung der Parameter ist abhängig davon, welches Modell gilt!
- Wir werden zunächst nur die Schätzung der Parameter in den eindimensionalen Modellen besprechen. Die mehrdimensionalen Modelle folgen in Vorlesung #08.

Schätzung der Varianzen der Fehlervariablen

- Wir besprechen im Folgenden den Output für das TAP-Beispiel für das **essentiell τ-äquivalenten Modell**
- Die genaue Berechnung der Schätzwerte für die Varianzen der Fehlervariablen ist etwas kompliziert und daher beschränken wir uns auf den R-Output
- Die Schätzwerte für die Varianzen (`Variances`) der Fehlervariablen (und die übrigen Parameter) werden zusammen mit dem Output des Modelltests ausgegeben

```
Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    f1             3132.875  407.514    7.688    0.000    1.000    1.000
   .almd1          2575.903  340.542    7.564    0.000 2575.903    0.451
   .almd2           664.165  123.319    5.386    0.000  664.165    0.175
   .almd3           522.041  111.204    4.694    0.000  522.041    0.143
   .almd4          1487.491  211.206    7.043    0.000 1487.491    0.322
```

➛ Die Varianzen der Tests mit Warnton sind geringer als ohne Warnton

Schätzung der Schwierigkeitsparameter

- Die Itemparameter σ_i entsprechen in allen Modellen, in denen sie vorkommen, wegen der Normierung $E(\theta) = 0$ den Erwartungswerten $E(X_i)$ der Items:

$$\sigma_i = E(X_i)$$

- Aus der zweiten Vorlesung wissen wir, dass wir die Erwartungswerte der Items durch die Mittelwerte der Items in der Stichprobe schätzen können
- Wir können daher auch die Itemparameter σ_i (`Intercepts`) durch die Mittelwerte der Items in der Stichprobe schätzen

```
Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    f1                0.000                               0.000    0.000
   .almd1           288.691    6.479   44.559    0.000  288.691    3.821
   .almd2           269.739    5.284   51.049    0.000  269.739    4.377
   .almd3           266.092    5.184   51.329    0.000  266.092    4.401
   .almd4           287.835    5.829   49.383    0.000  287.835    4.235
```

➤ Die Mittelwerte auf die Tests mit Warnton sind in der Stichprobe geringer

Schätzung der Steigungsparameter

- Da wir in unserem Beispiel von einem essentiell τ-äquivalenten Modell, ist die Schätzung der Steigungsparameter (z.B. `f1 =~ almd1`) nicht nötig, da sie auf den Wert „1“ festgelegt werden → Konstanten
- R-Output für das TAP-Beispiel:

```
Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  f1 =~
    almd1             1.000                               55.972    0.741
    almd2             1.000                               55.972    0.908
    almd3             1.000                               55.972    0.926
    almd4             1.000                               55.972    0.823
```

Schätzung der Erwartungswerte und Varianz der latenten Variable

- Auch die Schätzwerte für den Erwartungswert und die Varianz der latenten Variable finden sich im R-Output zum jeweiligen Modelltest
- Da wir in unserem Beispiel jedoch von einem essentiell τ-äquivalenten Modell ausgehen, ist ein Parameter per Normierung auf $E(\theta) = 0$ festgelegt, während der Parameter $\boldsymbol{VAR(\theta)}$ mit `3132.875` geschätzt wird

```
Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    f1             3132.875  407.514    7.688    0.000    1.000    1.000

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    f1                0.000                               0.000    0.000
```

Das war‘s schon fast für heute…

- *Ausblick:* In der nächsten Vorlesung beschäftigen wir uns mit der Schätzung der Parameter und der Bestimmung der Dimensionalität im mehrdimensionalen τ-kongenerischen Modell mittels Faktorenanalyse.

- *Aber zuerst:*
  - **Gibt es offene Fragen zur heutigen Vorlesung?**
  - Zur Vertiefung:
    - Übungsblatt 5 zur Skalierung auf Moodle
    - R Code zu den Beispielen der heutigen VL im Zusatzmaterial auf Moodle
    - R Code wird in den nächsten Wochen in den UKs vertieft thematisiert

Quellen

- Ostendorf, F., & Angleitner, A. (2004). *Neo-Persönlichkeitsinventar nach Costa und McCrae: Neo-PI-R.* Hogrefe.
- Zimmermann, P., & Fimm, B. (2009). *Testbatterie zur Aufmerksamkeitsprüfung (TAP).* Psytest.