Lehrstuhl für Psychologische Methodenlehre und Diagnostik der Ludwig-Maximilians-Universität München

Psychologische Testtheorie

Sitzung 6

Skalierung I

Wahr-oder-Falsch Aussagen zur letzten Sitzung

1. Falls die Items eines Tests, für den das 2-dimensionale τ-kongenerische Modell gilt, jeweils nur mit einer der beiden latenten Variablen zusammenhängen, kann der Test in zwei eindimensionale τ-kongenerische Modell zerlegt werden.
2. In einem q-dimensionalen τ-kongenerischen Modell gibt es insgesamt q Itemparameter.
3. Angenommen, ein Test für politische Einstellungen mit 6 Items folgt einem 2-dimensionalen τ-kongenerischen Modell, wobei θ_1 der latenten Variable *Autoritarismus* entspricht und θ_2 der latenten Variable *Konservatismus*. Den Wert von $\beta_{32} = 3.3$ können wir also folgendermaßen interpretieren: „Falls sich die latente Variable Konservatismus um 1 Punkt erhöht, erhöht sich die durchschnittliche Itemantwort auf Item 3 um 3.3 Punkte, falls die latente Variable Autoritarismus konstant bleibt.“
4. Wenn für einen konkreten Test das essentiell τ-äquivalente Modell gilt, dann gilt für diesen automatisch auch das τ-äquivalente Modell.
5. Wenn ein Item eines Tests von verschiedenen Personen unterschiedlich interpretiert wird, kann dies die Annahme der Personenhomogenität gefährden.

Überblick

Sitzung	Datum	Thema	Themenblock
1	13.10.25	Einführung	Begriffe, Modellierung von Antwortverhalten durch Zufallsvariablen & mathematische Grundlagen der Testtheorie
2	20.10.25	Wahrscheinlichkeitstheoret. Grundlagen	
3	27.10.25	Testtheoretische Modelle I	Testtheoretische Modelle
4	03.11.25	Testtheoretische Modelle II	
5	10.11.25	Testtheoretische Modelle III	
6	17.11.25	Skalierung I	Gütekriterien psychologischer Tests
7	24.11.25	Skalierung II	
8	01.12.25	Faktorenanalyse I	
9	08.12.25	Faktorenanalyse II	
10	15.12.25	Reliabilität I	
	22.12.25	Sitzung entfällt wegen Weihnachten!	
	29.12.25	Offizielle Winterpause	
	05.01.26		
11	12.01.26	Reliabilität II	Gütekriterien psychologischer Tests
12	19.01.26	Validität	
13	26.01.26	Einzelfalldiagnostik I	Methoden der Einzelfalldiagnostik
14	02.02.26	Einzelfalldiagnostik II	

Welchem testtheoretischen Modell folgt ein Test?

Überblick

Sitzung	Datum	Thema	Themenblock
1	13.10.25	Einführung	Begriffe, Modellierung von Antwortverhalten durch Zufallsvariablen & mathematische Grundlagen der Testtheorie
2	20.10.25	Wahrscheinlichkeitstheoret. Grundlagen	
3	27.10.25	Testtheoretische Modelle I	Testtheoretische Modelle
4	03.11.25	Testtheoretische Modelle II	
5	10.11.25	Testtheoretische Modelle III	
6	17.11.25	Skalierung I	

➢ In der heutigen Vorlesung werden wir uns damit beschäftigen, was „Skalierung“ genau bedeutet und warum sie wichtig ist. Außerdem werden wir herausfinden, welche empirisch überprüfbaren Folgerungen sich aus den testtheoretischen Modellen ergeben und wie wir diese mithilfe statistischer Methoden überprüfen können.

1. Einführung zur Skalierung

Wiederholung Skalierung

Ein psychologischer Test gilt als **skalierbar**, wenn die Zuordnung der Messwerte zu den Personen auf der Basis eines <u>empirisch nachgewiesenen testtheoretischen Modells</u> geschieht.

Zwei Aspekte:

- Empirischer „Nachweis“ , dass ein bestimmtes testtheoretisches Modell gilt, also das „richtige“ ist (aktuelle Vorlesung)
- Auf der Basis des „nachgewiesenen“ testtheoretischen Modells einzelnen Personen Werte auf den latenten Variablen zuweisen (Vorlesungen #13 und #14)

Wiederholung Modellannahmen

- Was bedeutet die Aussage „Ein testtheoretisches Modell gilt / ist richtig"?
- Testtheoretische Modelle treffen **Modellannahmen** bezüglich
  - des Zusammenhangs der zufälligen latenten Variable θ mit den zufälligen wahren Werten τ_i der Items des Tests
  - der Eigenschaften der Fehlervariablen ε_i
- Diese Annahmen können falsch sein, d.h. nicht der Realität entsprechen
- Die Frage, welches Modell gilt bzw. das „richtige" für einen konkreten psychologischen Test ist, ist gleichbedeutend mit der Frage, für welches Modell **alle Annahmen** erfüllt sind

→ Dies ist eine **empirische Frage:**

- Wir können die Annahmen der Modelle anhand einer Stichprobe überprüfen (ähnlich wie wir z.B. in Statistik II die Normalverteilungsannahme und die Homoskedastizitätsannahme der Regressionsmodelle überprüfen konnten)

Bedeutung Skalierung I

Der Nachweis der Skalierung ist die Voraussetzung für das komplette weitere Vorgehen!

- Zwei mögliche Situationen:
  1) Skalierung ist <u>nicht</u> gegeben – das heißt kein Modell gilt:
     - Wenn man nicht zeigen kann, dass eine latente Variable (oder mehrere) hinter den Items eines Tests steht, ergibt es auch keinen Sinn, nach der Genauigkeit des Tests zu fragen (Reliabilität) oder sich damit zu beschäftigen, ob die latente Variable tatsächlich die gewünschte psychologische Interpretation hat (Validität)

Der Nachweis der Skalierung ist die Voraussetzung für das komplette weitere Vorgehen!

- Zwei mögliche Situationen:
  2) Skalierung ist gegeben – das heißt eines der Modelle gilt:
     – Wenn ein „passendes“ Modell gefunden wurde, unterscheidet sich das weitere Vorgehen, je nachdem, welches Modell das „passende“ ist:
       → Die Bestimmung der Reliabilität unterscheidet sich je nach Modell
       → Die Schätzung der Werte der Personen auf der latenten Variable unterscheidet sich je nach Modell

Bedeutung Skalierung III

- Ausnahme:
  - Psychologische Tests, die gar nicht den Anspruch vertreten, latente Variablen zu messen, sondern nur, bestimmte Kriterien vorherzusagen (z.B. Studienerfolg, Rückfallwahrscheinlichkeit bei psychischen Störungen etc.)
  - In diesem Fall sind weder Skalierung noch Reliabilität des Tests relevant, sondern ausschließlich die Kriteriumsvalidität (und evtl. Nebengütekriterien wie Fairness)
- Die meisten psychologischen Tests werden jedoch zur Messung von latenten Variablen eingesetzt

2. Folgerungen aus dem Modellannahmen

Allgemeine Probleme bei der Überprüfung der Annahmen testtheoretischer Modelle

- Problem: die Annahmen testtheoretischer Modelle treffen Aussagen über Zufallsvariablen, deren Realisationen wir nicht beobachten können:
  - θ
  - τ_i
  - ε_i
- Die Annahmen der testtheoretischen Modelle sind also **nicht direkt überprüfbar**
- Wir können jedoch aus den Annahmen jeweils Folgerungen in Bezug auf beobachtbare Größen ableiten, die erfüllt sein müssen, falls die Annahmen gelten
- Diese Folgerungen sind dann statistische Hypothesen, die wir mithilfe von Stichproben und statistischen Hypothesentests überprüfen können
- Auf diesem Weg sind die Annahmen der testtheoretischen Modelle **indirekt überprüfbar**

2.1. Paralleles Modell

Paralleles Modell

Im parallelen Modell werden folgende Annahmen getroffen:

$\tau_i = \theta$ und somit $X_i = \theta + \varepsilon_i$ für alle Items i

$VAR(\varepsilon_i) = VAR(\varepsilon_j)$ für alle Itempaare i, j

$COV(\varepsilon_i, \varepsilon_j) = 0$ für alle Itempaare i, j

- Keine dieser Annahmen ist direkt überprüfbar
- Beobachten können wir nur die Realisationen der Zufallsvariablen X_i , also die Itemantworten der zufällig gezogenen Personen auf alle Items $i = 1, 2, \dots, k$
- Wir werden also versuchen, aus den Annahmen des parallelen Modells Folgerungen bezüglich der beobachtbaren Zufallsvariablen X_i abzuleiten

Paralleles Modell: Folgerung aus der ersten Annahme I

- Zunächst leiten wir aus der ersten Annahme eine **Folgerung für die Erwartungswerte** $E(X_i)$ der Items ab:

$$E(X_i) = E(\theta + \varepsilon_i) = E(\theta) + E(\varepsilon_i) = E(\theta)$$

Erste Annahme: $X_i = \theta + \varepsilon_i$

A5: Erwartungswert-Rechenregel

E2: Folgerung aus den Axiomen $E(\varepsilon_i) = 0$ $E(\theta) + \mathbf{0} = E(\theta)$

- Wir wissen zwar nicht, welchen Wert $E(\theta)$ hat, aber da die Gleichung für alle Items i gilt, folgt:

$$E(X_i) = E(\theta) = E(X_j) \text{ für alle Items } i, j$$

Paralleles Modell: Folgerung aus der ersten Annahme II

- Das heißt: Falls die erste Annahme des parallelen Modells erfüllt ist, müssen **alle Items den gleichen Erwartungswert** (≈ den gleichen Mittelwert in der Population) haben
- Umkehrschluss: Falls die Items nicht alle den gleichen Erwartungswert haben, kann auch die erste Annahme des parallelen Modells nicht erfüllt sein, das parallele Modell kann also nicht gelten

Paralleles Modell: Folgerung aus der ersten Annahme III

- Die Aussage

$$\mathbf{E(X_i) = E(X_j)}\ \textbf{für alle Items}\ \mathbf{i, j}$$

ist eine empirisch überprüfbare statistische Hypothese.

- Ideen für Hypothesentests?

Paralleles Modell: Folgerung aus der zweiten Annahme I

- Nun leiten wir aus der zweiten (und ersten) Annahme eine **Folgerung für die Varianzen** $VAR(X_i)$ der Items ab:

$$VAR(X_i) = VAR(\tau_i) + VAR(\varepsilon_i) = VAR(\theta) + VAR(\varepsilon_i)$$

E6: Folgerung aus den Axiomen

Erste Annahme: $\tau_i = \theta$

- Wir wissen zwar nicht, welche Werte $VAR(\theta)$ und $VAR(\varepsilon_i)$ haben, aber da die Gleichung für alle Items i gilt, folgt:

$$VAR(X_i) = VAR(\theta) + VAR(\varepsilon_i) = VAR(\theta) + VAR(\varepsilon_j) = VAR(X_j)$$

für alle Items i, j

Zweite Annahme:
$VAR(\varepsilon_i) = VAR(\varepsilon_j)$ für alle Itempaare

Paralleles Modell: Folgerung aus der zweiten Annahme II

- Das heißt: Falls die erste und zweite Annahme des parallelen Modells erfüllt ist, müssen **alle Items die gleiche Varianz** (≈ die gleiche Varianz in der Population) haben
- Umkehrschluss: Falls die Items nicht alle die gleiche Varianz haben, ist entweder die erste oder die zweite Annahme des parallelen Modells nicht erfüllt, das parallele Modell kann also nicht gelten

Paralleles Modell: Folgerung aus der zweiten Annahme III

- Die Aussage

$$\boldsymbol{VAR(X_i) = VAR(X_j)} \textbf{ für alle Items } \boldsymbol{i, j}$$

ist eine empirisch überprüfbare statistische Hypothese.

Paralleles Modell: Folgerung aus der dritten Annahme I

- Zuletzt leiten wir aus der dritten (und ersten) Annahme eine **Folgerung für die Kovarianzen** $COV(X_i, X_j)$ zweier Items i und j ab:

$$COV(X_i, X_j) = COV(\tau_i + \varepsilon_i, \tau_j + \varepsilon_j) =$$

E1: Folgerung aus den Axiomen: $X_i = \tau_i + \varepsilon_i$

C7: Kovarianz-Rechenregel

$$COV(\tau_i, \tau_j) + COV(\tau_i, \varepsilon_j) + COV(\varepsilon_i, \tau_j) + COV(\varepsilon_i, \varepsilon_j) =$$

E5: Folgerung aus den Axiomen: $Cov(\tau_i, \varepsilon_j) = 0$

Dritte Annahme: $COV(\varepsilon_i, \varepsilon_j) = \mathbf{0}$

$$COV(\tau_i, \tau_j) + 0 + 0 + 0 = COV(\theta, \theta) = VAR(\theta)$$

Erste Annahme: $\tau_i = \theta$

C3: Kovarianz-Rechenregel

- Wir wissen zwar nicht, welchen Wert $VAR(\theta)$ hat, aber da die Gleichung für alle Itempaare i, j und o, u gilt, folgt

$$COV(X_i, X_j) = VAR(\theta) = COV(X_o, X_u) \text{ für alle Itempaare } i, j \text{ und } o, u$$

Paralleles Modell: Folgerung aus der dritten Annahme II

- Das heißt: Falls die erste und dritte Annahme des parallelen Modells erfüllt ist, müssen **alle Itempaare die gleiche Kovarianz** (≈ die gleiche Kovarianz in der Population) haben
- Umkehrschluss: Falls die Itempaare nicht alle die gleiche Kovarianz haben, ist entweder die erste oder die dritte Annahme des parallelen Modells nicht erfüllt, das parallele Modell kann also nicht gelten

Paralleles Modell: Folgerung aus der dritten Annahme III

- Die Aussage

$$\boldsymbol{COV(X_i, X_j) = COV(X_o, X_u)} \textbf{ für alle Itempaare } \boldsymbol{i, j \text{ und } o, u}$$

ist eine empirisch überprüfbare statistische Hypothese.

Paralleles Modell: Folgerungen aus den Annahmen

- Zusammenfassend ergeben sich also aus den Annahmen des **parallelen Modells** folgende überprüfbare Aussagen bzw. statistische Hypothesen:

$$E(X_i) = E(X_j) \text{ für alle Items } i, j$$

$$VAR(X_i) = VAR(X_j) \text{ für alle Items } i, j$$

$$COV(X_i, X_j) = COV(X_o, X_u) \text{ für alle Itempaare } i, j \textit{ und } o, u$$

- Für die weiteren Modelle werden wir nur noch die Folgerungen ohne Beweise angeben. Die Beweise sind jedoch entweder analog oder zumindest sehr ähnlich wie im parallelen Modell.

2.2. Essentiell paralleles Modell

Essentiell paralleles Modell: Folgerungen aus den Annahmen

- Zusammenfassend ergeben sich aus den Annahmen des **essentiell parallelen Modells** folgende überprüfbare Aussagen bzw. statistische Hypothesen:

$$VAR(X_i) = VAR(X_j) \text{ für alle Items } i, j$$

$$COV(X_i, X_j) = COV(X_o, X_u) \text{ für alle Itempaare } i, j \text{ und } o, u$$

- Im essentiell parallelen Modell können wir aufgrund der Itemparameter im Gegensatz zum parallelen Modell **keine** allgemeine Gleichheit der Erwartungswerte der Items herleiten.
- Falls $\sigma_i \neq \sigma_j$ für zwei Items i und j gilt, ist für diese Items

$$E(X_i) = \sigma_i \neq \sigma_j = E(X_j)$$

2.3. τ-äquivalentes Modell

τ-äquivalentes Modell: Folgerungen aus den Annahmen

- Zusammenfassend ergeben sich also aus den Annahmen des **τ-äquivalenten Modells** folgende überprüfbare Aussagen bzw. statistische Hypothesen:

$$E(X_i) = E(X_j) \text{ für alle Items } i, j$$

$$COV(X_i, X_j) = COV(X_o, X_u) \text{ für alle Itempaare } i, j \text{ und } o, u$$

- Im τ-äquivalenten Modell können wir aufgrund der eventuell ungleichen Varianzen der Fehlervariablen im Gegensatz zu den parallelen Modellen **keine** allgemeine Gleichheit der Varianzen der Items herleiten.
- Falls $\boldsymbol{VAR(\varepsilon_i)} \neq \boldsymbol{VAR(\varepsilon_j)}$ für zwei Items i und j gilt, ist für diese Items

$$VAR(X_i) = VAR(\theta) + \boldsymbol{VAR(\varepsilon_i)} \neq VAR(\theta) + \boldsymbol{VAR(\varepsilon_j)} = VAR(X_j)$$

2.4. essentiell τ-äquivalentes Modell

Essentiell τ-äquivalentes Modell: Folgerungen aus den Annahmen

- Zusammenfassend ergeben sich aus den Annahmen des **essentiell τ-äquivalenten Modells** folgende überprüfbare Aussagen bzw. statistische Hypothesen:

$$\boxed{COV(X_i, X_j) = COV(X_o, X_u) \text{ für alle Itempaare } i, j \text{ } und \text{ } o, u}$$

- Im essentiell τ-äquivalenten Modell können wir aufgrund der Itemparameter und der eventuell ungleichen Varianzen der Fehlervariablen **weder** eine Gleichheit der Erwartungswerte **noch** der Varianzen der Items herleiten

2.5. τ-kongenerisches Modell

τ-kongenerisches Modell: Folgerungen aus den Annahmen

- Zusammenfassend ergeben sich aus den Annahmen des **τ-kongenerischen Modells** folgende überprüfbare Aussagen bzw. statistische Hypothesen:

$$COV\left(X_i, X_j\right) = \beta_i \cdot \beta_j \text{ für alle Itempaare } i, j$$

- Hier müssen die Kovarianzen im Vergleich zu den vorherigen Modellen also nicht für alle Itempaare gleich sein, sondern nur eine bestimmte, von den Steigungsparematern abhängige, Struktur aufweisen:
  - Die Kovarianz zweier Items muss dem Produkt ihrer Steigungsparameter entsprechen

2.6. Mehrdimensionales τ-kongenerisches Modell

Mehrdimensionales τ-kongenerisches Modell: Folgerungen aus den Annahmen

- Zusammenfassend ergeben sich aus den Annahmen des **mehrdimensionalen τ-kongenerischen Modells** folgende überprüfbare Aussagen bzw. statistische Hypothesen:

$$COV(X_i, X_j) = \sum_{l=1}^{q} \beta_{il} \cdot \beta_{jl} \quad \text{für alle Itempaare } i, j$$

- Auch hier müssen die Kovarianzen nicht für alle Itempaare gleich sein, sondern nur eine bestimmte, von den Steigungsparametern abhängige, Struktur aufweisen:
  - Die Kovarianz zweier Items muss der Summe der Produkte ihrer Steigungsparameter entsprechen

Übersicht über die Folgerungen aus den Modellannahmen

Modell	**gleiche Erwartungswerte aller Items**	**gleiche Varianzen aller Items**	**gleiche Kovarianzen aller Itempaare**
parallel	ja	ja	ja
essentiell parallel	nein	ja	ja
τ-äquivalent	ja	nein	ja
essentiell τ-äquivalent	nein	nein	ja
τ-kongenerisch	nein	nein	nein, aber bestimmte Struktur
mehrdimensional τ-kongenerisch	nein	nein	nein, aber bestimmte Struktur

3. Statistische Hypothesen

Folgerungen aus den Modellannahmen

- Eben haben wir gesehen, dass sich aus jedem testtheoretischen Modell Folgerungen für die **Erwartungswerte**, **Varianzen** und **Kovarianzen** der Items ableiten lassen
- Diese Folgerungen können wir mithilfe von statistischen Hypothesentests überprüfen

- 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 dem Beibehalten 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 und ist Teil der konfirmatorischen Faktorenanalyse (siehe Vorlesung #08)

Bemerkung

- Nächste Woche wenden wir den Omnibustest auf einen Beispielfragebogen an.
- Es gibt neben dem Omnibustest noch zahlreiche weitere statistische Methoden zur Überprüfung der Folgerungen aus den Modellannahmen
- Die wichtigsten Methoden werden wir in Vorlesung #08 zur Faktorenanalyse kennenlernen

Das war‘s schon fast für heute…

- *Ausblick:* In der nächsten Vorlesung wenden wir den Hypothesentest zur Überprüfung der Folgerungen aus den Modellannahmen auf einen Beispielfragebogen an und schauen uns die Parameterschätzung an.

- *Aber zuerst:*
  - **Gibt es offene Fragen zur heutigen Vorlesung?**
  - Zur Vertiefung:
    - Zusatzmaterial zur Ableitung der Folgerungen auf Moodle
    - Aufgaben 1 bis 5 im Übungsblatt 5 zur Skalierung auf Moodle