Psychologische Testtheorie

Sitzung 3

Testtheoretische Modelle I

Wahr-oder-Falsch Aussagen zur letzten Sitzung

1. Das durch die Zufallsvariable X_i beschriebene doppelte Zufallsexperiment lässt sich in zwei Zufallsexperimente aufteilen, die wiederum durch die Zufallsvariablen $\tau_{i\mathrm{Person}}$ und $X_{iPerson}$ beschrieben werden.

2. Der Erwartungswert E($X_{iPerson}$) entspricht dem wahren Wert der Person auf dem Item i und wird mit $\tau_{iPerson}$ bezeichnet.

3. Die Zufallsvariable τ_i beschreibt den zufälligen wahren Wert, den eine aus der Population gezogene Person auf Item i hat.

4. Jedes Item i verfügt über einen zufälligen wahren Wert τ_i und eine Fehlervariable ε_i.

5. Falls sich der zufällige wahre Wert τ_1 in $\tau_{1Person} = 1.4$ realisiert hat und die Person auf Item 1 den Wert 2 angekreuzt hat, dann hat sich die Fehlervariable ε_1 im Wert 3.4 realisiert.

WH: Axiome der Testtheorie

- D1: Für jedes Item i ist τ_i die Zufallsvariable, deren Realisation der wahre Wert $\tau_{iPerson}$ der Itemantwort der zufällig gezogenen Person ist
- D2: Für jedes Item i ist die Fehlervariable ε_i eine Zufallsvariable, die wie folgt definiert ist: $\varepsilon_i = X_i - \tau_i$

WH: Folgerungen der Testtheorie

- E1: Die Itemantwort einer zufällig gezogenen Person auf ein Item ist die Summe des zufälligen wahren Wertes und der Fehlervariable dieses Items: $\boldsymbol{X_i = \tau_i + \varepsilon_i}$
- E2: Der Erwartungswert der Fehlervariable eines Items ist 0: $\boldsymbol{E(\varepsilon_i) = 0}$
- E3: Der Erwartungswert eines Items entspricht dem Erwartungswert des zufälligen wahren Wertes auf diesem Item: $\boldsymbol{E(X_i) = E(\tau_i)}$
- E4: Die Kovarianz des zufälligen wahren Wertes eines Items mit der Fehlervariable dieses Items ist 0: $\boldsymbol{COV(\tau_i, \varepsilon_i) = 0}$
- E5: Die Kovarianz des zufälligen wahren Wertes eines Items mit der Fehlervariable eines beliebigen anderen Item ist 0: $\boldsymbol{COV\left(\tau_i, \varepsilon_j\right) = 0}$
- E6: Die Varianz eines Items i entspricht der Summe der Varianzen des zufälligen wahren Wertes und der Fehlervariable dieses Items: $\boldsymbol{VAR(X_i) = VAR(\tau_i) + VAR(\varepsilon_i)}$

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	

➢ In der heutigen Vorlesung werden wir uns genauer damit beschäftigen, wie die zufälligen wahren Werte τ_i der Items eines psychologischen Tests im Rahmen von testtheoretischen Modellen mit den uns eigentlich interessierenden latenten Variablen zusammenhängen.

1. Latente Variablen

Latente Variablen - Wiederholung

- Zur Erinnerung: Unter latenten Variablen verstehen wir nicht direkt beobachtbare Variablen (z.B. Intelligenz, Aufmerksamkeit, Extraversion)
- Um diese latenten Variablen durch einen psychologischen Test erfassen zu können, müssen wir sie mithilfe von mathematischen Gleichungen (also testtheoretischen Modellen) mit den manifesten Itemantworten des Tests in Beziehung setzen
- Zunächst müssen wir jedoch noch einige grundlegende Überlegungen anstellen…

1. Latente Variablen

1.1. Zufällige latente Variablen

1.2. Latente Variablen und wahre Werte

Zufällige latente Variable θ

- In der letzten Vorlesung haben wir für die Items eines Tests den Unterschied zwischen dem wahren Wert $\tau_{iPerson}$ einer festen Person und dem zufälligen wahren Wert τ_i eingeführt
- Diese Unterscheidung werden wir jetzt auch in Bezug auf latente Variablen treffen
- Den Wert einer **festen Person** auf der latenten Variable bezeichnen wir mit $\boldsymbol{\theta_{Person}}$. Dies ist der tatsächliche Wert der Person auf der latenten Variable.
  - Beispiel: Falls es sich bei der latenten Variable um Intelligenz handelt, bedeutet $\theta_{Person} = 100$, dass die feste Person eine Intelligenz von 100 aufweist
- Wenn die **Person zufällig gezogen** wird, muss sich der Wert θ_{Person} im Zuge der Ziehung erst einmal realisieren (genau wie $\tau_{iPerson}$). Wir wissen ja z.B. vor der Ziehung nicht, wie intelligent die gezogene Person sein wird.
- Die Zufallsvariable, die für diesen zufälligen Vorgang steht und deren Realisation θ_{Person} ist, nennen wir **zufällige latente Variable $\boldsymbol{\theta}$** (Theta)

Erwartungswert und Varianz von θ I

- Da die zufällige latente Variable θ eine Zufallsvariable ist, hat sie auch seinen **Erwartungswert $\mathbf{E}(\boldsymbol{\theta})$** und eine **Varianz $\mathbf{VAR}(\boldsymbol{\theta})$**
- Wir erinnern uns an die Interpretation des Erwartungswertes:
  Wenn man das der Zufallsvariable θ zugrunde liegende Zufallsexperiment unendlich oft wiederholen würde und den Mittelwert aus allen Realisationen der Zufallsvariable berechnen würde, wäre dies der Erwartungswert
- Das heißt: Wenn wir unendlich viele Personen aus der Population ziehen würden (vereinfacht gesagt: die komplette Population) und bei jeder gezogenen Person den Wert auf der latenten Variable θ_{Person} bestimmen würden (falls dies möglich wäre), dann wäre $E(\theta)$ der Mittelwert dieser Werte

Erwartungswert und Varianz von θ II

- Wir können also $\mathbf{E}(\boldsymbol{\theta})$ als Mittelwert der Werte auf der latenten Variable in der Population interpretieren
  - Beispiel: Falls es sich bei der latenten Variable um Intelligenz handelt, bedeutet $E(\theta) = 100$, dass die durchschnittliche Intelligenz in der Population 100 beträgt
- Analog ist dann $\mathbf{VAR}(\boldsymbol{\theta})$ die Varianz der Werte auf der latenten Variable in der Population, also ein Maß dafür, wie stark sich die Personen in der Population in ihren Werten auf der latenten Variable unterscheiden
  - Beispiel: Falls es sich bei der latenten Variable um Intelligenz handelt, bedeutet $VAR(\theta) = 225$, dass die Varianz der Intelligenz in der Population 225 beträgt

1. Latente Variablen

1.1. Zufällige latente Variablen

1.2. Latente Variablen und wahre Werte

Latente Variablen und wahre Werte I

- Aus der letzten Vorlesung wissen wir, dass wir die Antworten einer festen Person auf ein Item i als Zufallsvariable $X_{iPerson}$ auffassen können
- Diese Zufallsvariable besitzt für jede Person einen Erwartungswert $E(X_{iPerson}) = \tau_{iPerson}$, den wahren Wert der festen Person
- Zudem besitzt sie eine Varianz $VAR(X_{iPerson})$, da die Person je nach Kontext, Stimmung, etc. in ihrer Itemantwort variiert

Latente Variablen und wahre Werte II

- Wie können wir nun die Itemantworten $X_{iPerson}$ mit dem Wert auf der latenten Variable θ_{Person} in Beziehung setzen?
- Wenn wir annehmen, dass die Varianz in der Itemantwort nur durch zufällige Einflüsse wie Kontext, Stimmung etc. zustande kommt, sollte die latente Variable nur einen Einfluss darauf haben, wie eine Person „im Durchschnitt" auf das Item antwortet
- Das heißt: Die latente Variable sollte mit der Itemantwort nur über deren wahren Wert (also dem „durchschnittlichen" Wert) zusammenhängen

Latente Variablen und wahre Werte III

- Formal: Für jedes Item i sollte es eine itemspezifische Gleichung geben, die jeweils den wahren Wert $\tau_{iPerson}$ der Person auf diesem Item mit ihrem Wert θ_{Person} auf der latenten Variablen in Beziehung setzt, z.B.:

$$\tau_{iPerson} = \sigma_i + \beta_i \cdot \theta_{Person}$$

- Wir werden Gleichungen dieser Form von jetzt an gleich auf der Ebene der zufälligen wahren Werte und zufälligen latenten Variablen formulieren.
  Das heißt für das obige Beispiel:

$$\tau_i = \sigma_i + \beta_i \cdot \theta$$

- Inhaltlich gesehen ändert dies nichts an der Beziehung zwischen Item und latenter Variable

Latente Variablen und wahre Werte IV

- Wir erinnern uns nun an die erste Folgerung aus den Axiomen:

$$X_i = \tau_i + \varepsilon_i$$

- Wenn wir die zweite Gleichung der letzten Folie hier einsetzen, erhalten wir

$$X_i = \sigma_i + \beta_i \cdot \theta + \varepsilon_i$$

- **Wir sind also bei einem testtheoretischen Modell angekommen!**
  - Es setzt latente Variablen und manifeste Itemantworten zueinander in Beziehung und enthält eine Fehlervariable (vgl. Vorlesung 1)
  - Da wir aber den „Umweg“ über die zufälligen wahren Werte gegangen sind, können wir alle Folgerungen aus den Axiomen verwenden, was sich als sehr praktisch herausstellen wird

2. Testtheoretische Modelle

- Wir werden nun verschiedene testtheoretische Modelle kennenlernen
- Diese unterscheiden sich in ihren Modellannahmen bezüglich
  - des **Zusammenhangs** der **zufälligen latenten Variable $\boldsymbol{\theta}$** mit den **zufälligen wahren Werten $\boldsymbol{\tau_i}$** der Items des Tests
  - der Eigenschaften der **Fehlervariablen** ε_i
- **Diese Annahmen können – im Gegensatz zu den Axiomen und Folgerungen aus der letzten Vorlesung – falsch sein, d.h. nicht der Realität entsprechen!**
- Die Frage, welches Modell das „richtige“ für einen konkreten psychologischen Test ist, ist gleichbedeutend mit der Frage, für welches Modell alle Annahmen erfüllt sind
- Dies ist also eine empirische Frage, mit der wir uns später im Semester beschäftigen werden
- Sehr wichtig: **Aus theoretischer Sicht ist keines der Modelle per se „besser“ als das andere!**

2. Testtheoretische Modelle

Paralleles Modell – Annahmen

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

- **Rückgriff 1** auf Ausgangsmodell: $\boldsymbol{\tau_i} = \sigma_i + \beta_i \cdot \theta = 0 + 1 \cdot \theta = \boldsymbol{\theta}$
- **Rückgriff 2** auf Ausgangsmodell: $\boldsymbol{X_i} = \sigma_i + \beta_i \cdot \theta + \varepsilon_i = 0 + 1 \cdot \theta = \boldsymbol{\theta + \varepsilon_i}$

- Die inhaltliche Bedeutung dieser Annahmen diskutieren wir auf den nachfolgenden Folien…

Paralleles Modell – Erste Annahme I

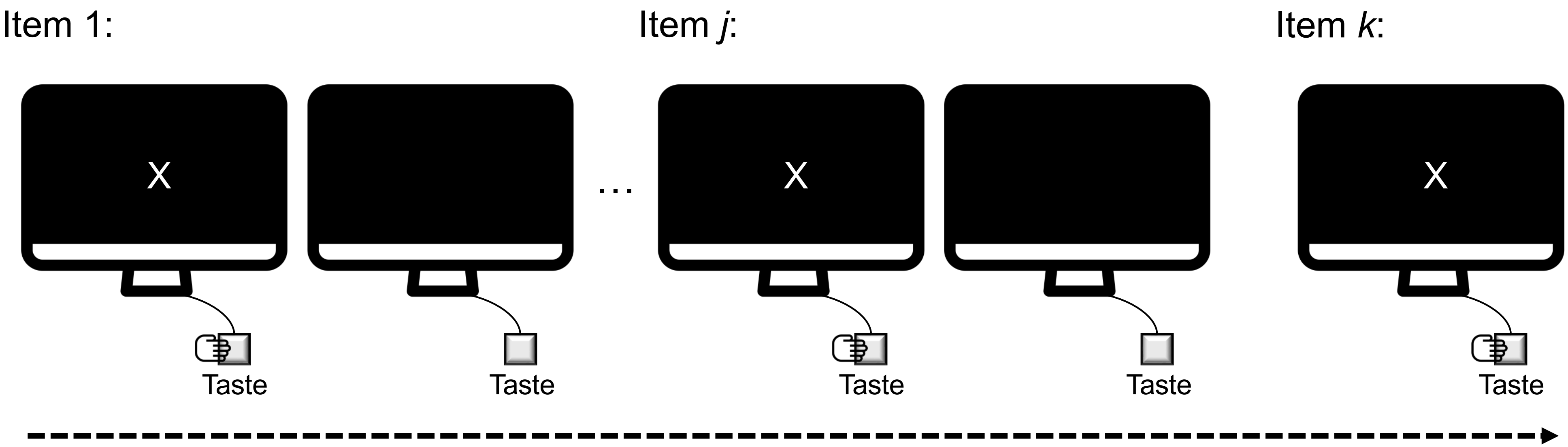
$\tau_i = \theta$ und somit $X_i = \theta + \varepsilon_i$ für alle Items i

- Die erste Annahme bezieht sich auf den Zusammenhang der zufälligen latenten Variable θ mit den zufälligen wahren Werten τ_i der Items des Tests
- In Worten: „Die zufällige latente Variable entspricht dem zufälligen wahren Wert auf jedem Item i."
- Auf Ebene einer festen Person heißt das: $\tau_{iPerson} = \theta_{Person}$ für alle Items i
- Das Modell nimmt also an, dass der wahre Wert auf jedem Item i identisch mit dem Wert auf der latenten Variable ist
- Da diese Gleichung für alle Items des Tests angenommen wird, muss im parallelen Modell wegen $\tau_{iPerson} = \theta_{Person} = \tau_{jPerson}$ eine Person denselben wahren Wert auf allen Items haben!

Beispiel I

Beispiel zur Illustration der Annahmen:

- Reaktionszeit-Test mit drei Items der TAP:

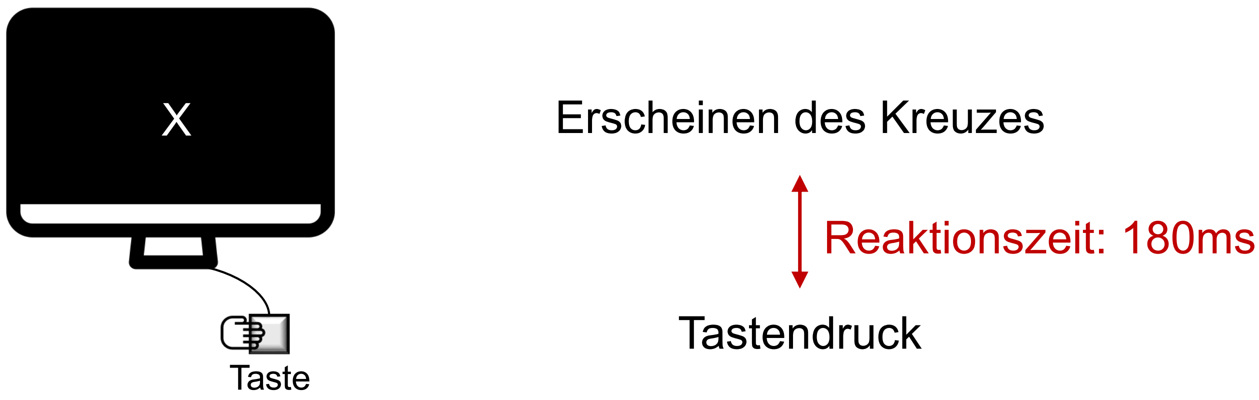


- Die Testpersonen müssen jeweils auf einen Knopf drücken, wenn der Stimulus erscheint
- Die Zeit zwischen Erscheinen des Stimulus und Drücken des Knopfes entspricht jeweils der Itemantwort

Beispiel II

Beispiel zur Illustration der Annahmen:

- Reaktionszeit-Test mit drei Items der TAP:



- Wenn also z.B. eine zufällig gezogene Person nach Erscheinen des ersten Stimulus (Kreuz) 180ms gebraucht hat, um auf den Knopf zu drücken, hat sich die Itemantwort X_1 in dem Wert $x_1 = 180$ realisiert
- Wir wollen mit diesem Test die latente Variable „Alertness" erfassen

Paralleles Modell – Erste Annahme II

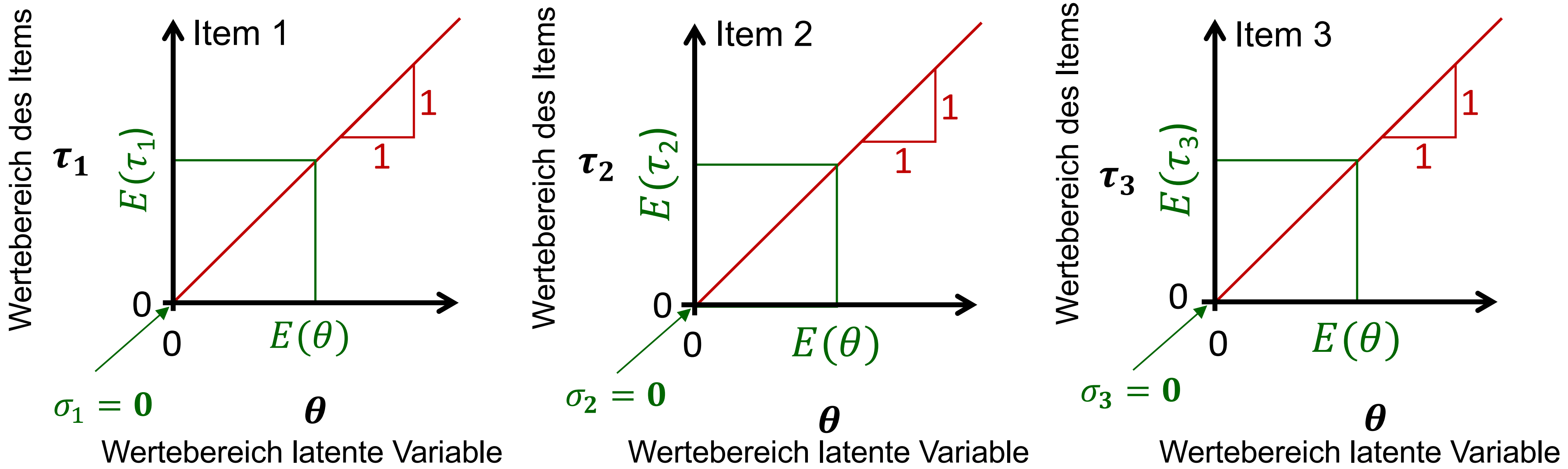
$\tau_i = \theta$ und somit $X_i = \theta + \varepsilon_i$ für alle Items i

- Bezogen auf unser Reaktionszeit-Beispiel:
  - Das Modell nimmt an, dass die durchschnittliche Reaktionszeit einer Person auf alle Items des Tests bei unendlicher Wiederholung des Tests exakt gleich ist und exakt ihrer Ausprägung auf der latenten Variable *Alertness* entspricht
  - Wenn z.B. der wahre Wert (also die „wahre Reaktionszeit") einer Person auf dem ersten Stimulus 180(ms) ist, dann muss der wahre Wert (also die „wahre Reaktionszeit") auf dem zweiten Stimulus auch 180(ms) betragen, und die Person hätte in diesem Fall einen Alertnesswert von 180
  - Wenn der wahre Wert (also die „wahre Reaktionszeit") einer anderen Person auf dem ersten Stimulus 200(ms) ist und auf dem zweiten Stimulus auch 200(ms), dann hätte die Person einen Alertnesswert von 200 (Person 2 hat einen höheren Alertnesswert als Person 1 → Person 2 ist im Mittel langsamer)

Paralleles Modell – Erste Annahme III

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

- Grafische Veranschaulichung:



$\tau_1 = 0 + 1 \cdot \theta = \theta$ $\tau_2 = 0 + 1 \cdot \theta = \theta$ $\tau_3 = 0 + 1 \cdot \theta = \theta$

da $\beta_1 = \beta_2 = \beta_3 = 1$ und $\sigma_1 = \sigma_2 = \sigma_3 = 0$

Paralleles Modell – Zweite Annahme I

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

- Die zweite Annahme bezieht sich auf die Varianzen der Fehlervariablen ε_i der Items des Tests
- In Worten: „Die Varianzen der Fehlervariablen sind für alle Items gleich groß"
- Zur Erinnerung: Die Fehlervariable auf einem Item i ist definiert als die Abweichung zwischen zufälligem wahrem Wert τ_i und der Itemantwort X_i auf dem Item i
- Bemerkung: Die Fehlervariablen ε_i selbst müssen nicht gleich sein, z.B. $\varepsilon_1 = 4$, $\varepsilon_2 = -3$, $\varepsilon_j = 1$ und $\varepsilon_k = -2$, nur Ihre Varianzen müssen gleich sein, z.B. $VAR(\varepsilon_1) = 3, VAR(\varepsilon_2) = 3, VAR(\varepsilon_j) = 3$ und $VAR(\varepsilon_k) = 3$

Paralleles Modell – Zweite Annahme II

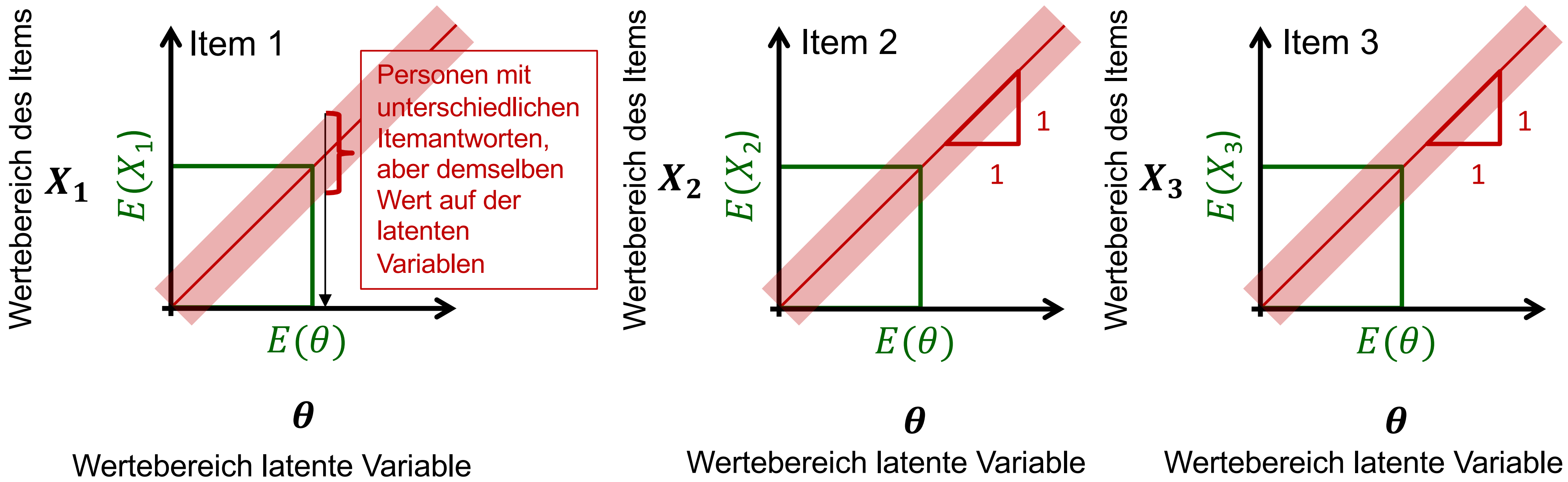
$VAR(\varepsilon_i) = VAR(\varepsilon_j)$ für alle Itempaare i, j

- Bezogen auf unser Reaktionszeit-Beispiel:
  - Das Modell nimmt an, dass die Varianzen der Abweichungen zwischen den durchschnittlichen Reaktionszeiten einer Person und den tatsächlichen Reaktionszeiten bei unendlicher Wiederholung des Tests auf beiden Items exakt gleich sind
  - Wenn z.B. der wahre Wert (also die „wahre Reaktionszeit“) einer Person auf dem ersten Stimulus 180(ms) ist und auf dem zweiten Stimulus auch 180(ms), dann sollten die tatsächlichen Reaktionszeiten der Person auf diese beiden Stimuli bei unendlicher Wiederholung jeweils in gleichem Maße um diese beiden wahren Werte streuen

Paralleles Modell – Zweite Annahme III

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

- Grafische Veranschaulichung:



Rote Balken gleich breit → gleiche Fehlervarianz

Paralleles Modell – Dritte Annahme I

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

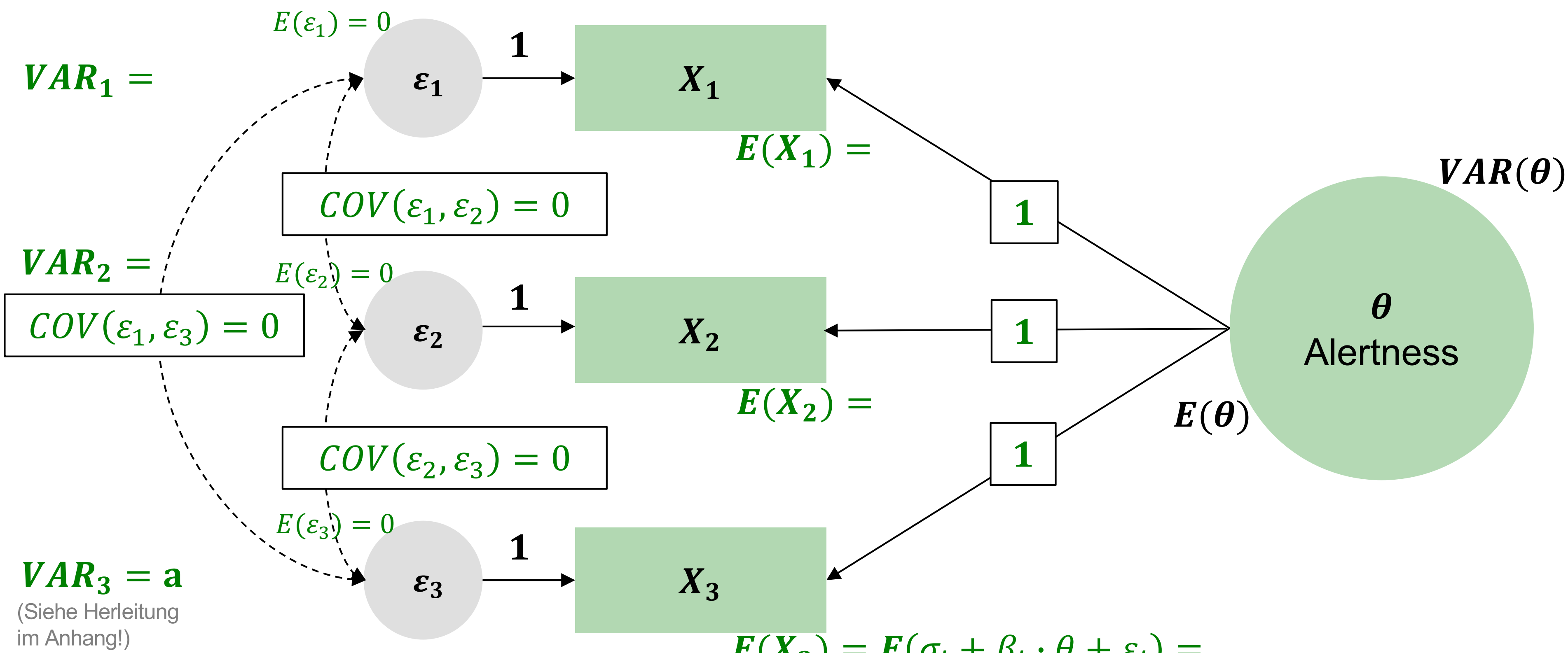
- Die dritte Annahme bezieht sich auf den Zusammenhang der Fehlervariablen zweier beliebiger Items des Tests
- In Worten: „Die Fehlervariablen zweier beliebiger Items des Tests kovariieren nicht"
- Die Annahme bedeutet, dass es keinen Zusammenhang zwischen der Abweichung des tatsächlichen Wertes von dem wahren Wert auf einem Item und der Abweichung des tatsächlichen Wertes von dem wahren Wert auf einem anderen Item geben darf
- Beispiele für Verletzungen der Annahme: Inhaltliche Abhängigkeit der Items, Übungseffekte, Frustration während der Testbeantwortung, etc.

Paralleles Modell – Dritte Annahme I

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

- Bezogen auf unser Reaktionszeit-Beispiel:
  - Das Modell nimmt an, dass die Kovarianz der Abweichungen der tatsächlichen Reaktionszeiten von den durchschnittlichen Reaktionszeiten einer Person bei unendlicher Wiederholung des Tests zwischen beiden Items exakt Null ist
  - Wenn z.B. der wahre Wert (also die „wahre Reaktionszeit“) einer Person auf beiden Stimuli 180(ms) ist und sie schon nach 150(ms) auf den ersten Stimuli reagiert (d.h. negative Abweichung von ihrem wahren Wert), dann sollte dies keinen Einfluss darauf haben, ob sie auf den zweiten Stimuli langsamer oder schneller als ihr wahrer Wert (=180) reagiert

Grafische Veranschaulichung des parallelen Modells



$E(X_3) = E(\sigma_i + \beta_i \cdot \theta + \varepsilon_i) =$
$E(0 + 1 \cdot \theta + \varepsilon_i) = E(\theta)$

Fixierte Parameter in dunkelgrün!

Paralleles Modell – Relevanz?

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

Alle Items messen die latente Variable im selben Wertebereich, gleich gut und genau!

→ Für welche Arten von Items ist dies ein realistisches Modell?

- am ehesten: Reaktionszeitaufgaben mit identischen Stimuli
- aber: In der Praxis gilt dieses Modell so gut wie nie!

Das war‘s schon fast für heute…

- *Ausblick:* In der nächsten Vorlesung lernen wir weitere testtheoretische Modelle kennen.

- *Aber zuerst:*
  - **Gibt es offene Fragen zur heutigen Vorlesung?**
  - Zur Vertiefung:
    - Übungsblatt 3 auf Moodle
    - Bühner (2021, Kapitel 4.4, S. 119-152)

Anhang

- Warum folgt aus dem parallelen Modell, dass bei Gleichheit der Fehlervarianzen auch alle Varianzen gleich sind?
- Wir wissen von einem parallelen Modell:
  $\boldsymbol{\tau_i = \theta}$ und $\boldsymbol{X_i = \theta + \varepsilon_i}$ und $\boldsymbol{VAR(\varepsilon_i) = VAR(\varepsilon_j)}$
- Beispiel mit drei Items:

$$\boldsymbol{VAR(X_1) = VAR(\theta + \varepsilon_1) = VAR(\theta) + VAR(\varepsilon_1) = VAR(\theta) + VAR(\varepsilon)}$$

$$\boldsymbol{VAR(X_2) = VAR(\theta + \varepsilon_2) = VAR(\theta) + VAR(\varepsilon_2) = VAR(\theta) + VAR(\varepsilon)}$$

$$\boldsymbol{VAR(X_3) = VAR(\theta + \varepsilon_3) = VAR(\theta) + VAR(\varepsilon_3) = VAR(\theta) + VAR(\varepsilon)}$$

denn:

- Rechenregel B6: falls $COV(X, Y) = 0$ gilt, dass $VAR(X + Y) = VAR(X) + VAR(Y)$
- Folgerung E4: $COV(\tau_i, \varepsilon_i) = 0$