Item Response Theorie

N. Rose

9. Vorlesung (16.01.2020)

Erweiterungen ein- und zweiparametrischer IRT Modelle

- Mehrgruppen IRT
- Längsschnittliche IRT Modelle
- Wichtig: **Invarianz des Messmodells** über Gruppen und Messzeitpunkte

Invarianz des Messmodells

Invarianz des Messmodells

- **Mehrgruppen-IRT Modelle** werden verwendet, …
  - um Verteilungsunterschiede in den latenten Variablen über Subpopulationen zu analysieren
  - um Unterschiede in den Zusammenhängen zwischen latenten Variablen (mehrdimensionale Mehrgruppen IRT Modelle!) über Gruppen zu analysieren
  - um Unterschiede in den Zusammenhängen zwischen Drittvariablen, d. h. Prädiktoren oder abhängigen Variablen der latenten Variablen des IRT Modells über Gruppen zu analysieren ➔ Moderationseffekte in IRT Messmodellen kombiniert mit Regressionsmodellen

Invarianz des Messmodells

- **Längsschnittliche IRT Modelle** werden verwendet, …
  - um Verteilungsunterschiede in den latenten Variablen über die Zeit zu analysieren
  - um intraindividuelle Veränderungen hinsichtlich der latenten Variablen zu analysieren
  - um intraindividuelle Veränderungen anhand von Drittvariablen, d. h. Prädiktoren zu erklären (IRT Messmodellen kombiniert mit Regressionsmodellen)

Invarianz des Messmodells

- Um Verteilungen von latenten Variablen über Subpopulationen hinweg vergleichen zu können, muss das **Messmodell invariant über die Subpopulation** sein ➔ die latenten Variablen müssen die **gleiche Metrik** in den Subpopulationen aufweisen.

- Um Verteilungen von latenten Variablen über Messzeitpunke hinweg vergleichen zu können, muss das **Messmodell invariant über die Messzeitpunkte** sein ➔ die latenten Variablen müssen die **gleiche Metrik** zu allen Messzeitpunkten aufweisen

Invarianz des Messmodells

- Aus der faktorenanalytischen Tradition wird zwischen **konfigureller, schwacher, starker** und **strikter Invarianz** des Messmodells unterschieden.

- Bei **Konfigureller Invarianz** des Messmodells ist die **Dimensionalität** der latenten Variable und das Ladungsmuster über die Subpopulationen bzw. über die Zeit gleich, aber nicht notwendigerweise die Faktorladungen, die Intercepts, und Messfehlervarianzen (SEM) bzw. die Itemdiskriminationen und Itemschwierigkeiten (IRT)

- **Schwache Invarianz** des Messmodells = **konfigurelle Invarianz** + **Gleichheit der Faktorladungen** (SEM) bzw. **Itemdiskriminationen** (IRT) über die Subpopulationen bzw. Messzeitpunkte

Invarianz des Messmodells

- **Starke Invarianz** der Messmodells = **konfigurelle + schwache Invarianz + Gleichheit der Intercepts** (SEM) bzw. **Itemschwierigkeiten** (IRT) über die Subpopulationen bzw. Messzeitpunkte

- **Strikte Invarianz** des Messmodells = **konfigurelle + schwache + starke Invarianz** + Gleichheit der Messfehlervarianzen (nur bei SEM) über die Subpopulationen bzw. Messzeitpunkte

Invarianz des Messmodells

- **Implikationen** der verschiedenen Stufen der Invarianz des Messmodelles:

- Bei **Konfigureller Invarianz** des Messmodells sind lediglich Korrelationen zwischen latenten Variablen über Subgruppen/Messzeitpunkte hinweg vergleichbar.
- Bei **schwacher Invarianz** des Messmodells sind lediglich Korrelationen und Kovarianzen zwischen latenten Variablen sowie die Varianzen der latenten Variablen über Subgruppen/Messzeitpunkte vergleichbar.
- Bei **starker Invarianz** des Messmodells sind Korrelationen und Kovarianzen zwischen latenten Variablen sowie die Varianzen und Erwartungswerte der latenten Variablen über Subgruppen vergleichbar.
- Bei **strikter Invarianz** folgt, dass alle Verteilungsunterschiede in den manifesten Variablen über Subpopulationen vergleichbar sind ➔ wichtig bei metrischen Variablen (SEM)!

Invarianz des Messmodells

- Die **starke Invarianz des Messmodells** gewährleistet die **gleiche Metrik der latenten Variablen** über Subpopulationen/Messzeitpunkte, und somit die sinnvolle Vergleichbarkeit von Mittelwerten, Varianzen und Kovarianzen über diese Subpopulation.

- Die **Invarianz des Messmodells** über Subpopulationen oder Messzeitpunkte ist eine **prüfbare Hypothese** (siehe Wald-Test und Andersen Likelihood Ratio Test zur Prüfung der Gleichheit der Itemschwierigkeiten im Rasch-Modell)

Mehrgruppen IRT Modelle

Mehrgruppen IRT Modelle

- **In Anwendung von Mehrgruppen-IRT Modellen** werden die Itemparameter und ggf. die Verteilungsparameter der latenten Variablen für unterschiedliche Gruppen in einem gemeinsamen Modell geschätzt

- Dabei können die Itemparameter über die Gruppen gleich gesetzt werden (sog. **Gleichheitsrestriktionen von Itemparametern**, *equality constraints*), um eine gemeinsame Metrik der latenten Variablen (starke Invarianz) sicherzustellen.

- Die Gleichheit der der Itemparameter kann inferenzstatistisch anhand verschiedener Modelle und Teststatistiken getestet werden.

Mehrgruppen IRT Modelle

- Das Phänomen ungleicher Itemparameter über Subpopulationen wird in der IRT als **differential item functioning (DIF)** bezeichnet

- Items die DIF aufweisen werden meist aus dem Test oder Itempool ausgeschlossen

- Alternativ kann die Gleichheitsrestriktion der Itemparameter für einzelnen aufgehoben werden ➔ Voraussetzung ist, dass genügend Items vorhanden sind die kein DIF aufweisen, um eine gemeinsame Metrik sicherzustellen (**Partielle Invarianz des Messmodells**)

Mehrgruppen IRT Modelle

- Mehrgruppen IRT Modelle können anhand unterschiedlicher Pakete berechnet werden.
  1. Ein- und zweiparametrische Mehrgruppen-IRT Modelle **mit Logit-Linkfunktion** können anhand des R-Paketes **TAM** berechnet werden

  2. Ein- und zweiparametrische Mehrgruppen IRT Modelle können unter Verwendung von **Strukturgleichungsmodellen für dichotome und ordinale Variablen** geschätzt werden. ➔ Umrechnung der Faktorladungen, Intercepts und Schwellenparameter in Itemdiskriminationen und Itemschwierigkeiten ist anhand der bekannten Tranformationen möglich.

Mehrgruppen IRT Modelle

- **Modellidentifikation** unter Annahme in einem Mehrgruppen IRT Modell mit *G* Gruppen der **starken Invarianz des Messmodells**:
  - **Mehrgruppen-IRT Modelle mit Logit-Linkfunktion**
    - Fixieren einer Itemdiskrimination und einer Itemschwierigkeit (z. B. $\alpha_i = 1$, $\beta_i = 0$) in allen Gruppen und Gleichsetzung (Gleichheitsrestriktion) der übrigen Itemdiskriminationen α_i und Itemschwierigkeiten β_i über die Gruppen. Die Verteilungsparameter $Var(\xi \mid X = x)$ und $E(\xi \mid X = x)$ können dann in allen Gruppen $x = 1, \ldots, G$ frei geschätzt werden

    oder ...

    - Fixieren der Verteilungsparameter $Var(\xi \mid X = x)$ und $E(\xi \mid X = x)$ der latenten Variablen in einer Gruppe (z. B. $Var(\xi \mid X = 1) = 1$ und $E(\xi \mid X = 1) = 0$). Die Itemdiskriminationen α_i und Itemschwierigkeiten β_i werden über die Gruppen gleichgesetzt (Gleichheitsrestriktion) aber frei geschätzt. Die Verteilungsparameter $Var(\xi \mid X = x)$ und $E(\xi \mid X = x)$ in den übrigen Gruppe sind dann frei schätzbar.

Mehrgruppen IRT Modelle

- **Modellidentifikation** unter Annahme in einem Mehrguppen IRT Modell mit *G* Gruppen der **starken Invarianz des Messmodells**:
  - **Mehrgruppen-SEM für dichotome/ordinale Variablen**
    - Fixieren einer Faktorladung und eines Schwellenparameters (z. B. $\lambda_i = 1$, $\delta_i = 0$). Die Verteilungsparameter $Var(\xi \mid X = x)$ und $E(\xi \mid X = x)$ können dann in allen Gruppen $x = 1, \ldots, G$ frei geschätzt werden

    oder ...

    - Fixieren der Verteilungsparameter $Var(\xi \mid X = x)$ und $E(\xi \mid X = x)$ der latenten Variablen in einer Gruppe (z. B. $Var(\xi \mid X = 1) = 1$ und $E(\xi \mid X = 1) = 0$). Die Faktorladungen λ_i und Schwellenparameter δ_i werden über die Gruppen gleichgesetzt (Gleichheitsrestriktion) aber frei geschätzt. Die Verteilungsparameter $Var(\xi \mid X = x)$ und $E(\xi \mid X = x)$ in den übrigen Gruppe sind dann frei schätzbar.

Mehrgruppen IRT Modelle

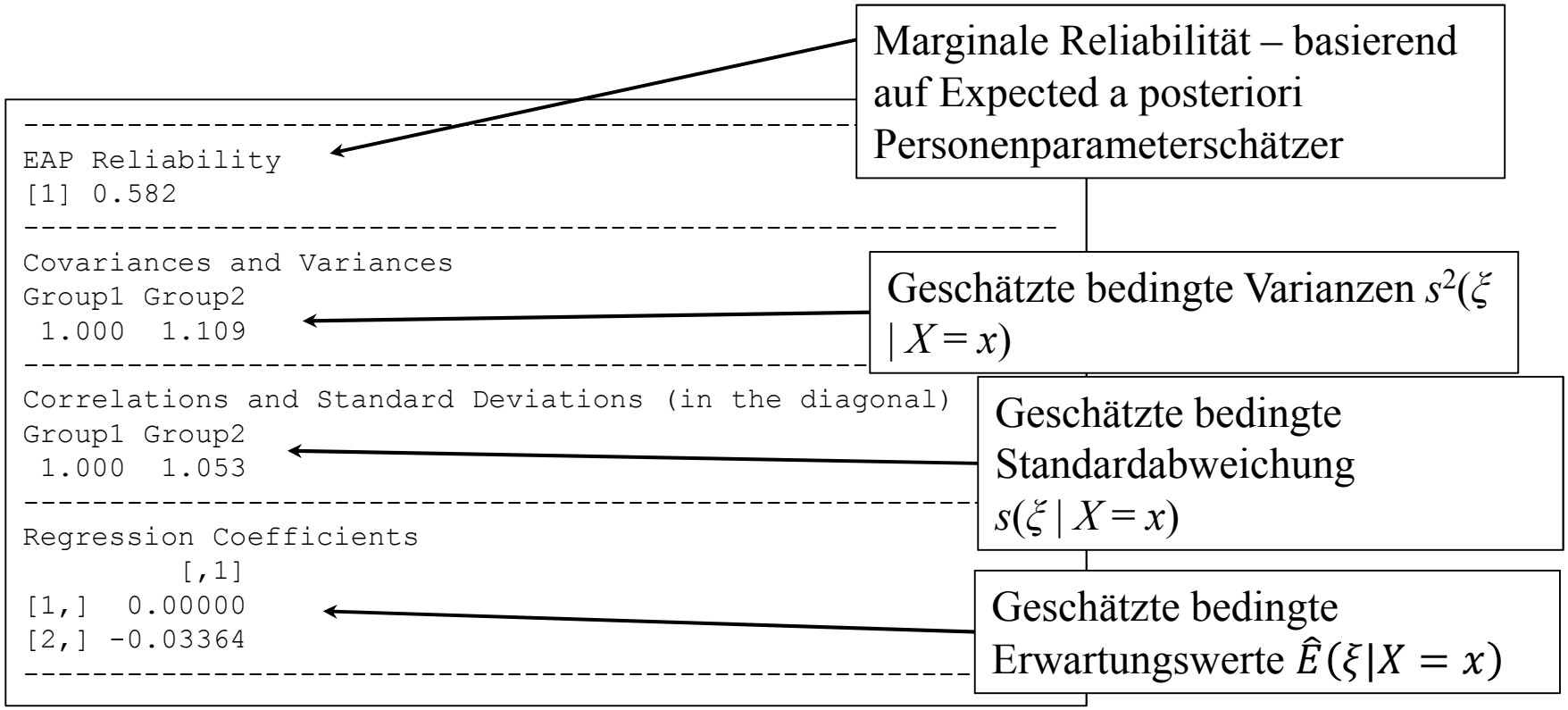
- **Fragestellung**: Gibt es Verteilungsunterschiede hinsichtlich des induktiven Denkens zwischen Männern und Frauen?

- Die folgenden Analysen werden unter Verwendung des Datensatzes von Wichert et al. (2012) durchgeführt!
- **2PL-Zweigruppen-IRT Modell mit Logit-Linkfunktion ➔ Paket TAM**

```
# Mehrgruppen IRT Modell mit Logit Link ########################################

# Beachte: Fälle mit fehlenden Werten hinsichtlich der Gruppierungvariable
# müssen ausgeschlossen werden
tmp.dat <- na.omit(dat.1[,c(paste0("rav",6:20),"sex")])

# 2-Gruppen-Birnbaum-Modell
BBM.sex <- tam.mml.2pl(resp=tmp.dat[,1:15], irtmodel="2PL", group=tmp.dat$sex)

# Ergbenisse
summary(BBM.sex)
```

Mehrgruppen IRT Modelle

- **2PL-Zweigruppen-IRT Modell mit Logit-Linkfunktion ➔ Paket TAM - Output**



Mehrgruppen IRT Modelle

- **2PL-Zweigruppen-IRT Modell mit Logit-Linkfunktion ➔ Paket TAM - Output**

```
Item Parameters in IRT parameterization
    item alpha  beta
1    rav6 1.217 -3.079
2    rav7 1.593 -2.192
3    rav8 1.164 -2.444
4    rav9 1.250 -2.878
5   rav10 1.609 -2.077
6   rav11 1.946 -2.514
7   rav12 1.008 -2.442
8   rav13 0.883 -0.968
9   rav14 0.993 -2.405
10  rav15 0.766 -2.666
11  rav16 0.996 -1.806
12  rav17 0.505 -2.459
13  rav18 0.924 -1.213
14  rav19 0.889 -1.222
15  rav20 0.712 -1.349
```

Geschätzte Itemdiskriminationen

Geschätzte Itemschwierigkeiten

Mehrgruppen IRT Modelle

- **Zweigruppen-IRT Modell mit Mehrgruppen-SEM für dichotome/ordinale Variablen ➔ R-Paket lavaan**

```
bbm <-  'xi =~ NA*rav6 + rav7 + rav8 + rav9 + rav10 + rav11 + rav12 +
               rav13 + rav14 + rav15 + rav16 + rav17 + rav18 + rav19 +
               rav20
         # Varianz fixieren
         xi ~~ c(NA,1)*xi
         xi ~ c(NA,0)*1
         '
fit <- sem(model = bbm, ordered = paste0("rav",6:20),estimator="WLSMV",
           group="sex",
           group.equal = c("loadings", "thresholds", "intercepts"),
           data  = dat.1)

# Output
summary(fit,fit.measures=T)
```

Modellidentifikation: $E(\xi|X=1)=0$ und $Var(\xi|X=1)=1$

Spezifizierung der Gruppierungsvariable („*sex*“) uns Gleichheitsrestriktion der Itemparameter über die Gruppen

Mehrgruppen IRT Modelle

- **Nullhypothese des Allgemeinen Modellgeltungstests:** Gleichheit der wahren und wahren modellimplizierten tetrachorischen Korrelationsmatrizen, und Gleichheit der wahren und wahren modellimplizierten Schwellenstruktur **in allen Gruppen.**

Estimator		DWLS	**Robust**	
Model Fit Test Statistic		138.835	190.226	
Degrees of freedom		193	193	
P-value (Chi-square)		0.999	0.543	
Scaling correction factor			1.281	
Shift parameter for each group:				
2			59.184	
1			22.666	
for simple second-order correction (Mplus variant)				
Chi-square for each group:				
2		59.353	105.516	
1		79.481	84.710	
User model versus baseline model:				
Comparative Fit Index (CFI)		1.000	1.000	
Tucker-Lewis Index (TLI)		1.043	1.005	
Root Mean Square Error of Approximation:				
RMSEA		0.000	0.000	
90 Percent Confidence Interval	0.000	0.000	0.000	0.026
P-value RMSEA <= 0.05		1.000	1.000	

Allgemeiner Modelltest

Fit Indices

Mehrgruppen IRT Modelle

- **Faktorladungen (Gruppe 1: *sex* = 2)**

Beachte: Output nach Gruppen getrennt

```
Group 1 [2]:

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)
  xi =~
    rav6     (.p1.)    0.757    0.228    3.314    0.001
    rav7     (.p2.)    0.743    0.233    3.185    0.001
    rav8     (.p3.)    0.711    0.221    3.224    0.001
    rav9     (.p4.)    0.648    0.226    2.874    0.004
    rav10    (.p5.)    0.769    0.245    3.136    0.002
    rav11    (.p6.)    0.931    0.307    3.036    0.002
    rav12    (.p7.)    0.614    0.198    3.100    0.002
    rav13    (.p8.)    0.478    0.188    2.537    0.011
    rav14    (.p9.)    0.573    0.182    3.140    0.002
    rav15    (.10.)    0.493    0.151    3.268    0.001
    rav16    (.11.)    0.517    0.170    3.035    0.002
    rav17    (.12.)    0.329    0.115    2.858    0.004
    rav18    (.13.)    0.528    0.189    2.799    0.005
    rav19    (.14.)    0.523    0.182    2.868    0.004
    rav20    (.15.)    0.396    0.168    2.355    0.019
```

Mehrgruppen IRT Modelle

- **Schwellenparameter (Gruppe 1: *sex* = 2)**

Thresholds:

	Estimate	Std.Err	z-value	P(>|z|)
rav6|t1 (.18.)	-1.920	0.472	-4.072	0.000
rav7|t1 (.19.)	-1.509	0.454	-3.326	0.001
rav8|t1 (.20.)	-1.535	0.437	-3.514	0.000
rav9|t1 (.21.)	-1.646	0.391	-4.213	0.000
rv10|t1 (.22.)	-1.514	0.467	-3.241	0.001
rv11|t1 (.23.)	-1.949	0.572	-3.405	0.001
rv12|t1 (.24.)	-1.298	0.378	-3.434	0.001
rv13|t1 (.25.)	-0.548	0.279	-1.964	0.050
rv14|t1 (.26.)	-1.241	0.348	-3.571	0.000
rv15|t1 (.27.)	-1.135	0.303	-3.745	0.000
rv16|t1 (.28.)	-0.978	0.315	-3.106	0.002
rv17|t1 (.29.)	-0.746	0.205	-3.648	0.000
rv18|t1 (.30.)	-0.716	0.312	-2.293	0.022
rv19|t1 (.31.)	-0.709	0.310	-2.285	0.022
rv20|t1 (.32.)	-0.576	0.239	-2.416	0.016

Mehrgruppen IRT Modelle

- **Verteilungsparameter latente Variable (Gruppe 1: *sex* = 2)**

```
Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)
    xi               -0.202    0.534   -0.377    0.706

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
    xi                0.881    0.582    1.514    0.130
```

- Beachte, Achsenabschnitte sind gleich den Erwartungswerten, wenn die die Variablen nicht als abhängige Variablen in einer Regression spezifiziert wurden!
- Die Nullhypothese, der zufolge sich Männer und Frauen im induktiven Denken nicht unterscheiden, kann beibehalten werden.

Mehrgruppen IRT Modelle

- **Skalierungsfaktoren (engl. *scale factors*)**

(Gruppe 1: *sex* = 2)

Werte sind auf den Wert 1 fixiert!

Scales y*:

	Estimate	Std.Err	z-value	P(>|z|)
rav6	1.000			
rav7	1.000			
rav8	1.000			
rav9	1.000			
rav10	1.000			
rav11	1.000			
rav12	1.000			
rav13	1.000			
rav14	1.000			
rav15	1.000			
rav16	1.000			
rav17	1.000			
rav18	1.000			
rav19	1.000			
rav20	1.000			

(Gruppe 2: *sex* = 1)

Scales y*:

	Estimate	Std.Err	z-value	P(>|z|)
rav6	0.789	0.207	3.808	0.000
rav7	1.081	0.330	3.274	0.001
rav8	0.766	0.223	3.438	0.001
rav9	1.102	0.386	2.858	0.004
rav10	0.909	0.283	3.215	0.001
rav11	0.917	0.288	3.187	0.001
rav12	1.004	0.309	3.252	0.001
rav13	0.771	0.334	2.308	0.021
rav14	1.077	0.307	3.505	0.000
rav15	1.055	0.298	3.537	0.000
rav16	0.987	0.327	3.014	0.003
rav17	1.060	0.338	3.134	0.002
rav18	0.665	0.264	2.516	0.012
rav19	0.690	0.271	2.545	0.011
rav20	1.045	0.446	2.341	0.019

Werte sind frei geschätzt!

Mehrgruppen IRT Modelle

- **Skalierungsfaktoren (engl. *scale factors*)**
  - Die Skalierungfaktoren sf_i sind definiert als Kehrwert der Standard-abweichung der latent response variables Y_i^*

$$sf_i = \frac{1}{Std\left(Y_i^*\right)}$$

  - Skalierungsfaktoren werden benötigt, um in einzelnen Gruppen Varianzen $Var(Y_i^* \mid X = \mathrm{x})$ der latent response variables abweichend von Eins zulassen zu können.
  - Unterschiede der Varianzen der latent response variables $Var(Y_i^* \mid X = g) \neq Var(Y_i^* \mid X = h)$ zwischen zwei Gruppen g und h folgen, wenn bei invarianten Itemparametern Varianzunterschiede $Var(\xi \mid X = g) \neq Var(\xi \mid X = h)$ hinsichtlich der latenten Variable zwischen den Gruppen g und h bestehen.

Mehrgruppen IRT Modelle

- **Skalierungsfaktoren (engl. *scale factors*)**
  - In Anwendung von Mehrgruppen-SEM sind die Skalierungsfaktoren sf_i nur in einer Gruppe fixiert ($sf_{i \mid X=x} = 1$) und werden in den anderen Gruppen frei geschätzt, um mögliche Varianzunterschiede in der latenten Variablen ξ zwischen Gruppen angemessen zu berücksichtigen.

    ➔ Ist default Einstellung in lavaan und braucht nicht vom Anwender extra spezifiziert werden!

Mehrgruppen IRT Modelle

- **Residualvarianzen der latent response variables** werden nicht extra geschätzt sondern werden aus Faktorladungen, der Varianz der latenten Variable und den Skalierungsfaktoren berechnet.

(Gruppe 1: *sex* = 2)

$$Var\left(\varepsilon_i^* \mid X = x\right)$$

```
Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .rav6              1.032
   .rav7              0.303
   .rav8              1.199
   .rav9              0.403
   .rav10             0.619
   .rav11             0.322
   .rav12             0.614
   .rav13             1.454
   .rav14             0.534
   .rav15             0.655
   .rav16             0.759
   .rav17             0.781
   .rav18             1.985
   .rav19             1.829
   .rav20             0.760
```

(Gruppe 2: *sex* = 1)

```
Variances:
                   Estimate  Std.Err  z-value  P(>|z|)
   .rav6              0.495
   .rav7              0.514
   .rav8              0.555
   .rav9              0.630
   .rav10             0.480
   .rav11             0.237
   .rav12             0.668
   .rav13             0.799
   .rav14             0.711
   .rav15             0.786
   .rav16             0.764
   .rav17             0.905
   .rav18             0.754
   .rav19             0.759
   .rav20             0.862
```

Mehrgruppen IRT Modelle

- **Residualvarianzen der latent response variables** werden nicht extra geschätzt sondern werden aus Faktorladungen, der Varianz der latenten Variable und den Skalierungsfaktoren berechnet.

$$Var\left(\varepsilon_i^* \mid X = x\right) = \left(\frac{1}{sf_{i|X=x}}\right)^2 - \lambda_i^2 Var\left(\xi \mid X = x\right)$$

- Dabei ist $sf_{i \mid X=x}$ der Skalierungsfaktor von Item i in Gruppe $X = x$.

Längsschnittliche IRT Modelle

Längsschnittliche IRT Modelle

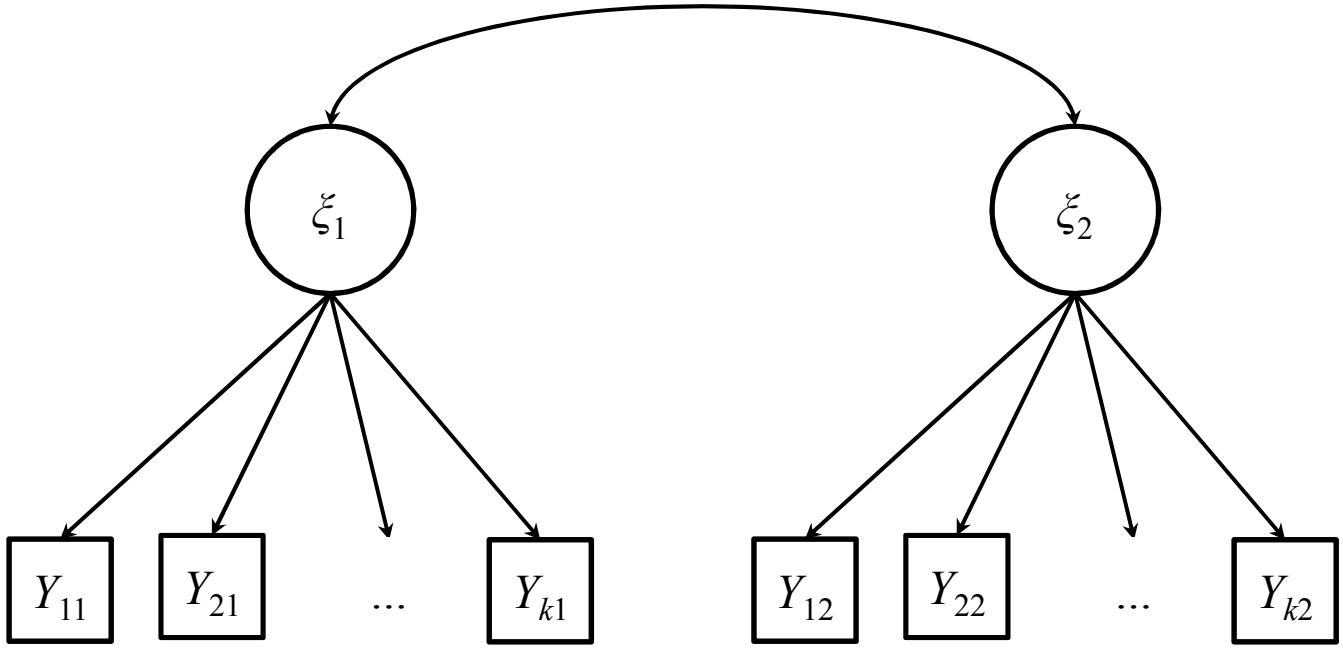
- **Längsschnittliche IRT Modelle sind Speziallfälle von mehrdimensionalen IRT Modellen**, wobei die Itemparameter und die Verteilungsparameter der latenten Variablen für unterschiedliche Messzeitpunkte in einem gemeinsamen Modell geschätzt werden.

- Dabei können die Itemparameter über die Messzeitpunkte hinweg gleich gesetzt werden (sog. **Gleichheitsrestriktionen von Itemparametern**, *equality constraints*), um eine gemeinsame Metrik der latenten Variablen (starke Invarianz) sicherzustellen.

- Die Gleichheit der der Itemparameter kann inferenzstatistisch anhand verschiedener Modelle und Teststatistiken getestet werden.

Längsschnittliche IRT Modelle

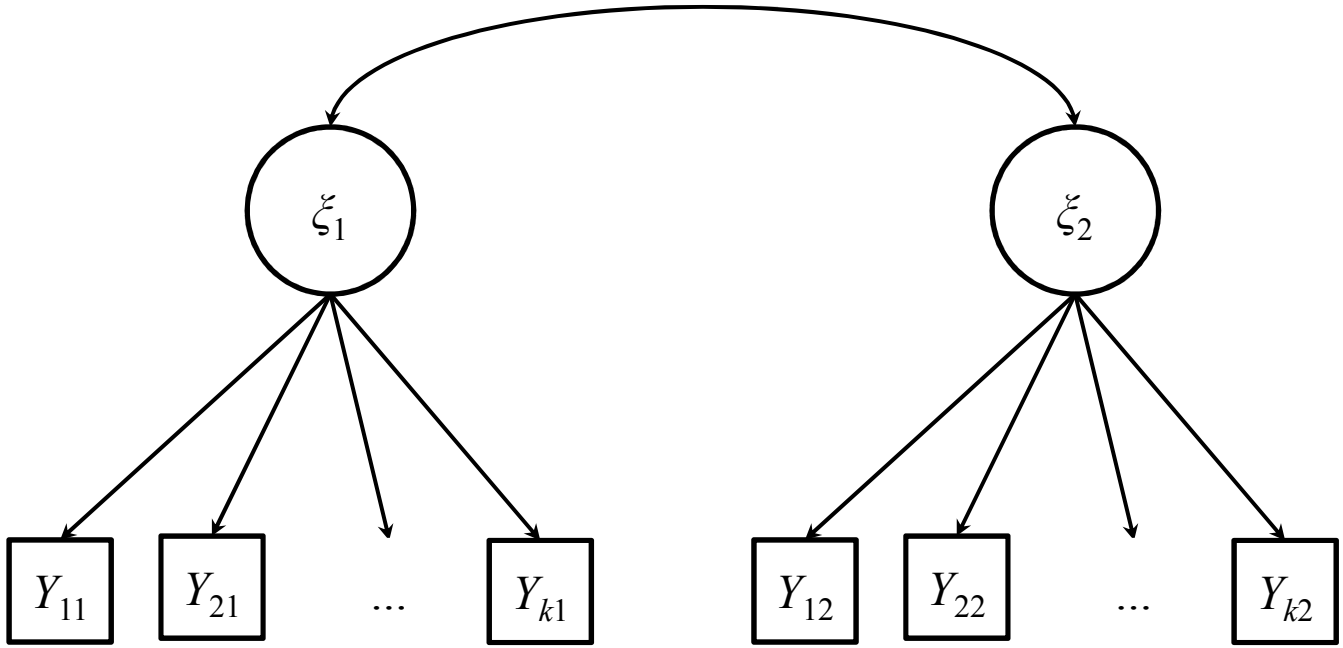
- Beachte Indizes der Variablen und Parameter!



$$P(Y_{11}=1|\xi_1)=\frac{\exp\left[\alpha_{11}(\xi_1-\beta_{11})\right]}{1+\exp\left[\alpha_{11}(\xi_1-\beta_{11})\right]}$$

$$P(Y_{12}=1|\xi_2)=\frac{\exp\left[\alpha_{12}(\xi_2-\beta_{12})\right]}{1+\exp\left[\alpha_{12}(\xi_2-\beta_{12})\right]}$$

Längsschnittliche IRT Modelle



- Bei **starker Invarianz der Messmodells** über die Zeit gilt

$$\alpha_{i1} = \alpha_{i2} = \ldots = \alpha_{iT}$$
$$\beta_{i1} = \beta_{i2} = \ldots = \beta_{iT}$$

Längsschnittliche IRT Modelle

- Das Phänomen ungleicher Itemparameter über Messzeitpunkte wird in der IRT als **Item-Drift** bezeichnet

- Items die Item-Drift aufweisen werden ggf. aus dem Test oder Itempool ausgeschlossen

- Alternativ kann die Gleichheitsrestriktion der Itemparameter für einzelnen aufgehoben werden ➔ Voraussetzung ist, dass genügend Items vorhanden sind die kein Drift aufweisen, um eine gemeinsame Metrik sicherzustellen (**Partielle Invarianz des Messmodells**)

Längsschnittliche IRT Modelle

- Längsschnittliche IRT Modelle können anhand unterschiedlicher Pakete für mehrdimensionale IRT Modelle berechnet werden.
  1. Ein- und zweiparametrische längsschnittliche IRT Modelle **mit Logit-Linkfunktion** können anhand des R-Paketes **TAM** und **mirt** berechnet werden

  2. Ein- und zweiparametrische längsschnittliche IRT Modelle können unter Verwendung von **Strukturgleichungsmodellen für dichotome und ordinale Variablen** geschätzt werden. ➔ Umrechnung der Faktorladungen, Intercepts und Schwellenparameter in Itemdiskriminationen und Itemschwierigkeiten ist anhand der bekannten Transformationen möglich.

Längsschnittliche IRT Modelle

- **Modellidentifikation** in einem längsschnittlichen IRT Modell mit *T* Messzeitpunkten unter Annahme der **starken Invarianz des Messmodells**:
  - **Längsschnittliche IRT Modelle mit Logit-Linkfunktion**
    - Fixieren einer Itemdiskrimination und einer Itemschwierigkeit (z. B. $\alpha_{i1} = 1$, $\beta_{i1} = 0$) zu allen Messzeitpunkten und Gleichsetzung (Gleichheitsrestriktion) der übrigen Itemdiskriminationen α_{it} und Itemschwierigkeiten β_{it} über die Messzeitpunkte. Die Verteilungsparameter $Var(\xi_t)$ und $E(\xi_t)$ können dann zu allen Messzeitpunkten $t = 1, \ldots, T$ frei geschätzt werden

    oder ...

    - Fixieren der Verteilungsparameter $Var(\xi_t)$ und $E(\xi_t)$ der latenten Variablen zu einem Messzeitpunkt (z. B. $Var(\xi_1) = 1$ und $E(\xi_1) = 0$). Die Itemdiskriminationen α_{it} und Itemschwierigkeiten β_{it} werden über die Messzeitpunkte gleichgesetzt (Gleichheitsrestriktion) aber frei geschätzt. Die Verteilungsparameter $Var(\xi_{t \neq 1})$ und $E(\xi_{t \neq 1})$ zu den andern Messzeitpunkten sind dann frei schätzbar.

Längsschnittliche IRT Modelle

- **Modellidentifikation** in einem längsschnittlichen IRT Modell mit T Messzeitpunkten unter Annahme der **starken Invarianz des Messmodells**:
  - **Mehrgruppen-SEM für dichotome/ordinale Variablen**
    - Fixieren einer Faktorladung und eines Schwellenparameters (z. B. $\lambda_{it} = 1$, $\delta_{it} = 0$). Die Verteilungsparameter $Var(\xi_t)$ und $E(\xi_t)$ können dann zu allen Messzeitpunkten $t = 1, \ldots, T$ frei geschätzt werden.

    oder ...

    - Fixieren der Verteilungsparameter $Var(\xi_t)$ und $E(\xi_t)$ der latenten Variablen zu einem Messzeitpunkt (z. B. $Var(\xi_1) = 1$ und $E(\xi_1) = 0$). Die Faktorladungen λ_{it} und Schwellenparameter δ_{it} werden über die messzeitpunkte gleichgesetzt (Gleichheitsrestriktion) aber frei geschätzt. Die Verteilungsparameter $Var(\xi_{t \neq 1})$ und $E(\xi_{t \neq 1})$ zu den andern Messzeitpunkten sind dann frei schätzbar.

Längsschnittliche IRT Modelle

- **Fragestellung**: Gibt es im Mittel eine Zunahme in den Ausprägungen der latenten Variable?

- Die folgenden Analysen werden unter Verwendung des simulierten Datensatzes **MathTest_t1t2.dat** durchgeführt!
  - $N = 730$ Personen
  - Zwei Messzeitpunkte

Längsschnittliche IRT Modelle

- Spezifikation eines zweidimensionalen SEM für dichotome Variablen, dass dem zweiparametrischen Modell (Birnbaum) entspricht!

```
# SEM reapeated Measurement (equivalent to two-paramter IRT Model)
twoTimes <- '
  xi1 =~ NA*y1t1 + la1*y1t1 + la2*y2t1 + la3*y3t1 + la4*y4t1 + la5*y5t1 + la6*y6t1 + la7*y7t1 + la8*y8t1 + la9*y9t1 + la10*y10t1
  xi2 =~ NA*y1t2 + la1*y1t2 + la2*y2t2 + la3*y3t2 + la4*y4t2 + la5*y5t2 + la6*y6t2 + la7*y7t2 + la8*y8t2 + la9*y9t2 + la10*y10t2
  # latent variables - variances, covariances, means
  xi1 ~~ 1*xi1
  xi2 ~~ xi2
  xi1 ~~ xi2
  xi1 ~ 0*1
  xi2 ~ 1
  # Equality constraints - thresholds
  y1t1  | thr1*t1
  y1t2  | thr1*t1
  y2t1  | thr2*t1
  y2t2  | thr2*t1
  y3t1  | thr3*t1
  y3t2  | thr3*t1
  y4t1  | thr4*t1
  y4t2  | thr4*t1
  y5t1  | thr5*t1
  y5t2  | thr5*t1
  y6t1  | thr6*t1
  y6t2  | thr6*t1
  y7t1  | thr7*t1
  y7t2  | thr7*t1
  y8t1  | thr8*t1
  y8t2  | thr8*t1
  y9t1  | thr9*t1
  y9t2  | thr9*t1
  y10t1 | thr10*t1
  y10t2 | int10*t1
```

Gleichsetzung der Faktorladungen der Items über beide Messzeitpunkte

Modellidentifikation: $E(\xi_1) = 0$ und $Var(\xi_1) = 1$

Gleichsetzung der Schwellen-parametern über die Zeitpunkte unter Verwendung von *parameter labels* *int1 – int10*

Allgemein werden Schwellenparameter in lavaan durch eine senkrechten Strich | in Kombination mit den Bezeichnern **`tk`** angesprochen!

Längsschnittliche IRT Modelle

- Frei Schätzung der Varianzen $Var(Y_{it}^{*})$ bzw. Fehlervarianzen $Var(\varepsilon_{it}^{*})$ der latent response variables über alle Items zu einem (hier dem 2.) Messzeitpunkt (➔ Fortsetzung des Inputs)

```
  # free scaling factors of all items at time t2
  y1t2  ~*~ NA*y1t2
  y2t2  ~*~ NA*y2t2
  y3t2  ~*~ NA*y3t2
  y4t2  ~*~ NA*y4t2
  y5t2  ~*~ NA*y5t2
  y6t2  ~*~ NA*y6t2
  y7t2  ~*~ NA*y7t2
  y8t2  ~*~ NA*y8t2
  y9t2  ~*~ NA*y9t2
  y10t2 ~*~ NA*y10t2
'
```

- Residualvarianzen $Var(\varepsilon_{it}^{*})$ der latent response variables werden durch ~*~ angesprochen

Längsschnittliche IRT Modelle

- **Nullhypothese des Allgemeinen Modellgeltungstests:** Gleichheit der wahren und wahren modellimplizierten tetrachorischen Korrelationsmatrizen, und Gleichheit der wahren und wahren modellimplizierten Schwellenstruktur.

Estimator		DWLS	Robust	
Model Fit Test Statistic		144.840	173.225	
Degrees of freedom		177	177	
P-value (Chi-square)		0.963	0.566	
Scaling correction factor			0.983	
Shift parameter			25.818	
for simple second-order correction (Mplus variant)				
User model versus baseline model:				
Comparative Fit Index (CFI)		1.000	1.000	
Tucker-Lewis Index (TLI)		1.014	1.002	
Root Mean Square Error of Approximation:				
RMSEA		0.000	0.000	
90 Percent Confidence Interval	0.000	0.000	0.000	0.015
P-value RMSEA <= 0.05		1.000	1.000	
Standardized Root Mean Square Residual:				
SRMR		0.051	0.051	

Allgemeiner Modelltest

Fit Indices

Längsschnittliche IRT Modelle

- **Weitere Ergebnisse**

```
Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)   S
  xi1 ~~
    xi2               0.708    0.076    9.289    0.000

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|)   S
    xi1               0.000                              0.000    0.000
    xi2               0.499    0.070    7.140    0.000    0.446    0.446

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
    xi1               1.000                               1.000    1.000
    xi2               1.250    0.197    6.331    0.000    1.000    1.000
```

Es gibt einen signifikanten durchschnittlichen Zuwachs hinsichtlich der latenten Variable vom ersten zum zweiten Messzeitpunkt

Deskriptiv zeigt sich auch eine Varianzzunahme in den Ausprägungen der latenten Variable