Diagnostic Classification Models

Ae Kyong Jung
November 12, 2025

Contents

CFA vs. IRT vs. DCMs

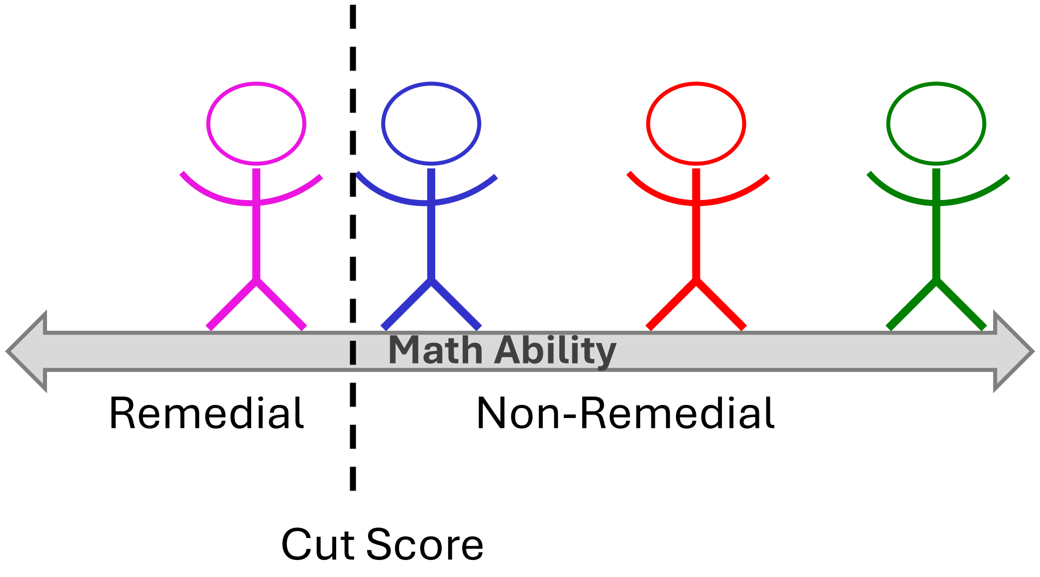
Models	Response	Latent Variable
CFA	Continuous	Continuous
IRT/IFA	Discrete	Continuous
DCMs	Discrete	Discrete

CFA: Confirmatory Factor Analysis; IRT: Item Response Theory; IFA: Item Factor Analysis; DCMs: Diagnostic Classification Models

DCMs
• DCMs is the name of the models used to obtain **classifications/diagnoses**. • DCMs can do a **diagnostic decision** that is being made based on information. • DCMs can decide if the person has **mastered** or **non-mastered** the skill.

Motivation for DCMs

- ❑ Information from **Continuum**:
- Blue has more math ability than Pink
- Red scored in the 70^{th} percentile
- Green scored a 240 on the test
- ❑ Traditional psychometrics cannot answer:
  - → Why the pink is so low?
  - → How much ability is enough to pass?
  - → What math skills have the student mastered?
- ❑ **Diagnosis/Decision** from Cut Score:
- Pink scored below the cut score
- Pink will take the remedial math course



Diagnosis

❑ Diagnoses featured in this lecture:

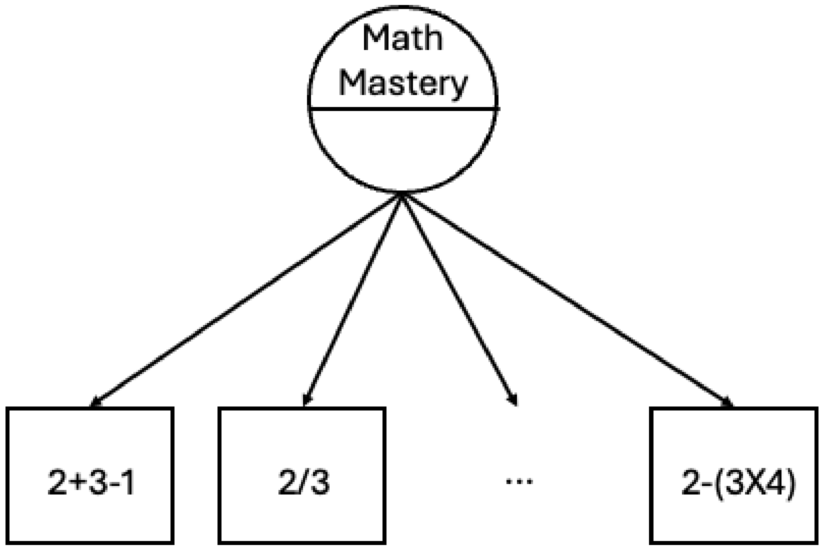
- Educational Measurement
  - The competencies (skills) that a person has or has not mastered
    - Math: convert a whole number to a fraction, separate a whole number from a fraction, simplify before subtracting, find a common denominator, etc. (de la Torre & Douglas, 2004)
    - Reading: synthesize scattered information, recognize relevant information, identify gist of a passage, apply relevant background knowledge, etc. (Buck, Tatsuoka, & Kostic, 1997)

❑ Also possible:

- Psychiatric Assessment
  - The Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria that a person meets
    - Leads to a broader diagnosis of a disorder
- Personnel selection
  - Matching traits of individual to needs of jobs

Single Dimension

DCMs with Single Latent Variable (Single Attribute)



Path Diagram

	Math Mastery
Item 1	1
Item 2	1
...	...
Item N	1

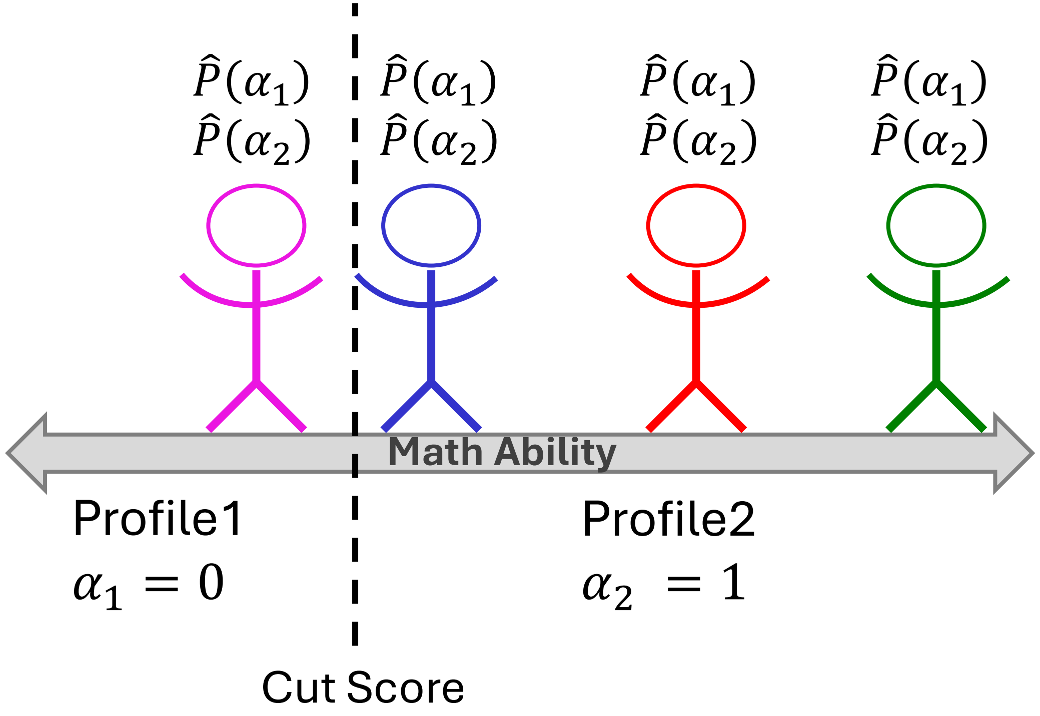
Q-matrix

Attribute is typically synonymous with the terms 'latent trait' and 'latent characteristic'. Attribute has a history of use in the literature on factor analysis (e.g., McDonald, 1999)

The specification of which attributes are constitutive of the response process for each item is done numerically in a table with a particular structure called a ***Q-matrix*** (Tatsuoka, 1983) $(q_1 = 1, q_2 = 1, \dots, q_n = 1)$

Respondent Profiles (Classes / Attribute Patterns /α_c)

Profiles	Math Mastery (α)
Profile 1 ($c = 1$)	0
Profile 2 ($c = 2$)	1



- Respondents are characterized by attribute profiles specifying which attributes have been mastered
- Respondent profile estimates are in the form of probabilities of master $\hat{P}(\alpha_c)$

CFA vs. IRT vs. DCMs (Single Latent Variable)

Model	Response	Latent Variable	Equation
CFA	Continuous	Continuous	$\mathrm{y_{ip}} = \mu_i + \lambda_i \theta_p + \epsilon, \qquad \epsilon \sim \mathrm{N}(0, \sigma_\epsilon^2)$
2-PL IRT	Discrete	Continuous	$ln\left(\frac{P(y_{ip} = 1 \| \theta_p)}{P(y_{ip} = 0 \| \theta_p)}\right) = \mu_i + \lambda_i \theta_p$
LDCM	Discrete	Discrete	$ln\left(\frac{P(y_{ic} = 1 \| \alpha_c)}{P(y_{ic} = 0 \| \alpha_c)}\right) = \mu_i + \lambda_i \alpha_c, q_i$

*The log-linear cognitive diagnostic model (LCDM) is the general form of DCMs which subsumes all other latent-variable DCMs. Equation represent the conditional probability that respondent p with attribute profile $\alpha_p = \alpha_c$ provides a correct response to item i.

- i: item
- p: person
- q_i : q-matrix of item i (q-element here because we only have single attribute)

- θ_p : math ability of person p $(\theta \sim N(0,1))$
- α_c : mastery/non-mastery of math ability of class c $(\alpha = 0, 1)$
- c: class (c=1,2)
- Non-mastery class's attribute vector: $\alpha_{c=1} = 0$,
- Mastery class's attribute vector : $\alpha_{c=2} = 1$

Multiple Dimensions

Psychometrics Multiple Dimensions

- The set of skills represent the multiple dimensions of elementary mathematics ability (e.g., $+, -, \times, \div$)

- Other psychometrics approaches have been developed for multiple dimensions
  - Classical Test Theory – Scale Subscores
  - CFA with more than two factors
  - Multidimensional Item Response Theory (MIRT)

- Yet, issues in application have remained:
  - Reliability of estimates is often poor for most practical test lengths
  - Dimensions are often very highly correlated
  - Large samples are needed to calibrate item parameters in MIRT

Note: From Lecture#3 PSQF 6249 (slide 70-71) by Templin, J., Hoffman, L, 2025. https://docs.google.com/viewer?url=https://raw.githubusercontent.com/jonathantemplin/Structural-Equation-Modeling-Fall-2025/main/_lectures/03_Confirmatory_Factor_Analysis/psqf6249f2025_lecture03.pdf

DCMs with Multiple Dimensions

- DCMs provide respondents valuable information with fewer data demands than other multidimensional models (Templin and Bradshaw (2013))
  - Higher reliability than comparable IRT/MIRT models
  - Complex item structures possible
  - 18 items diagnostic test for 0.8 reliability
    → diagnose mastery of 3 attributes
  - 34 items traditional test for 0.8 reliability
    → scale unidimensional ability

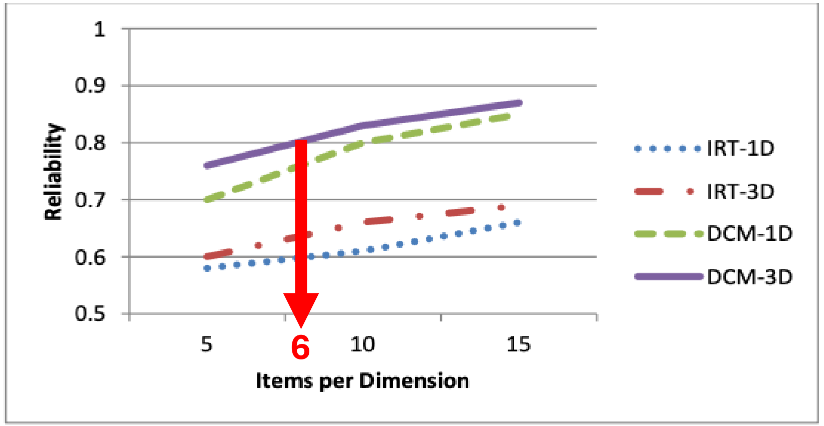
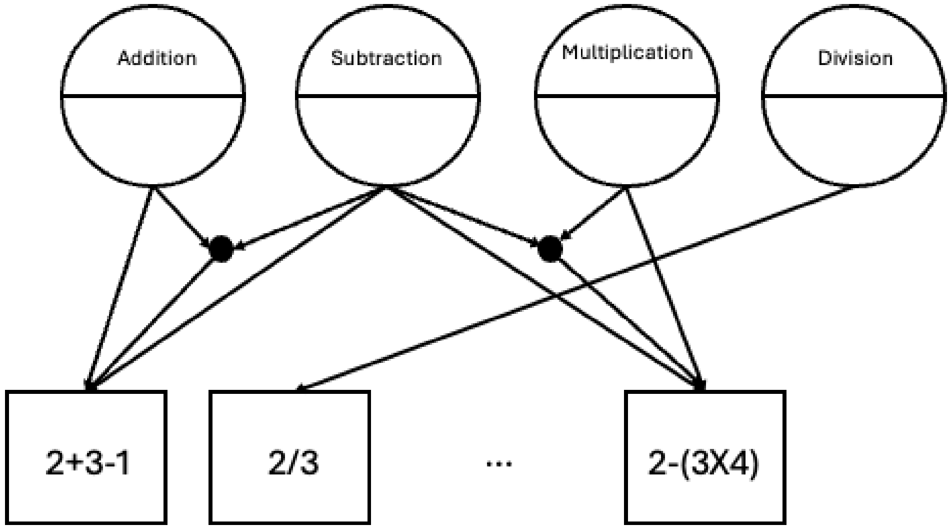


Figure 2. Simulation Study Results: Reliability of DCM and IRT Models with 100 Generated Tests

Note. From Templin & Bradshaw (2013). Measuring the Reliability of Diagnostic Classification Model Examinee Estimates. Journal of classification (30) 251-275.

- Paradox of DCMs
  - Sacrifice fine-grained measurement of a latent trait for only several categories
  - Increased capacity to measure ability multidimensionally

DCMs with Multiple Latent Variables



Path Diagram

	Add.	**Sub.**	**Mult.**	**Div.**
Item 1	1	1	0	0
Item 2	0	0	0	1
...				
Item N	0	1	1	0

Q-matrix

- This is the factor pattern matrix that assigns the loadings in confirmatory factor analysis.
- A Q-matrix is used to indicate the attributes measured by each item. (e.g., $q_1 = c(1, 1, 0, 0)$)

Respondent Profiles (Classes/Attribute Patterns/$\boldsymbol{\alpha}_c$)

The total # profiles = 2^A

Q-matrix

	Add.	**Sub.**	**Mult.**	**Div.**
Item1	1	1	0	0
Item2	0	0	0	1
Item3	0	1	0	1

Respondent Mastery

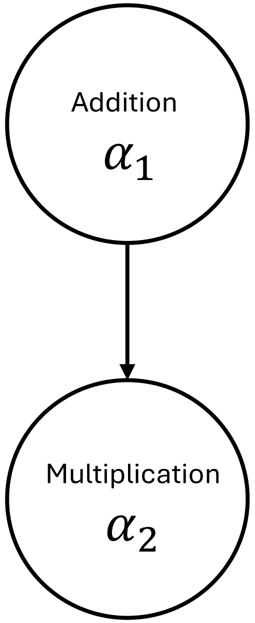
	Add.	**Sub.**	**Mult.**	**Div.**
Respondent1	1	1	1	0

Think. Choose one of: Low/ Medium/ High
1. Respondent1 has a _____ probability to correctly answering item1.
2. Respondent1 has a _____ probability to correctly answering item2.
3. Respondent1 has a _____ probability to correctly answering item3.

What do you think about #3?
The answer depends on your assumption, which affect the interaction terms of the latent variables. Stay tuned!

Attribute Hierarchies

- Attribute hierarchies: specifications of the attribute dependencies in the population of respondents
- Reduce structural parameters to estimate (2^A)

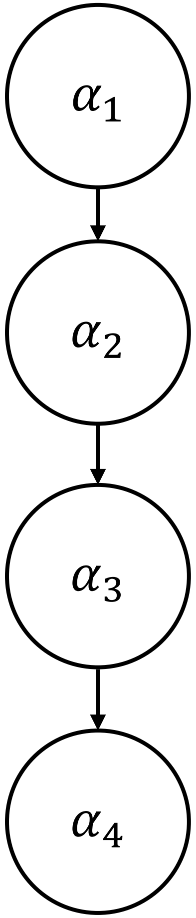


Profiles		
	Add.	**Mult.**
Profile1	0	0
Profile2	1	0
~~Profile3~~	~~0~~	~~1~~
Profile4	1	1

- Suppose that mastery of attribute 1 is prerequisite to mastery of attribute 2
- Attribute profiles where the second but not the first attribute is mastered must logically be empty in the population

Prototypical Attribute Hierarchies I

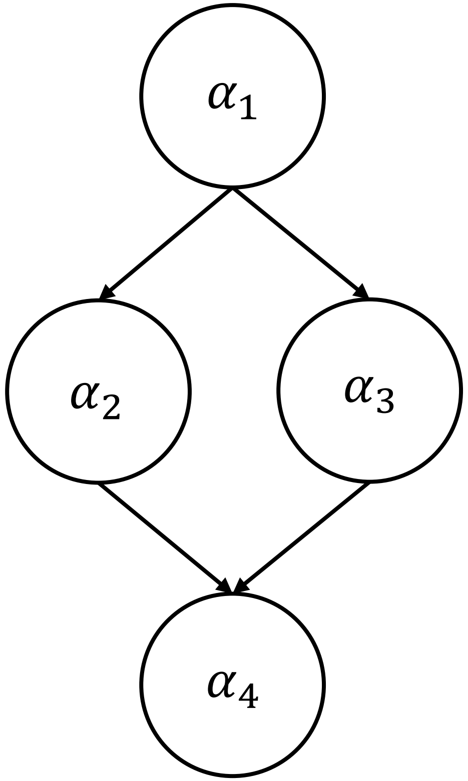
A. Linear



Profiles

	α_1	α_2	α_3	α_4
profile1	1	0	0	0
profile2	1	1	0	0
profile3	1	1	1	0
profile4	1	1	1	1

B. Convergent



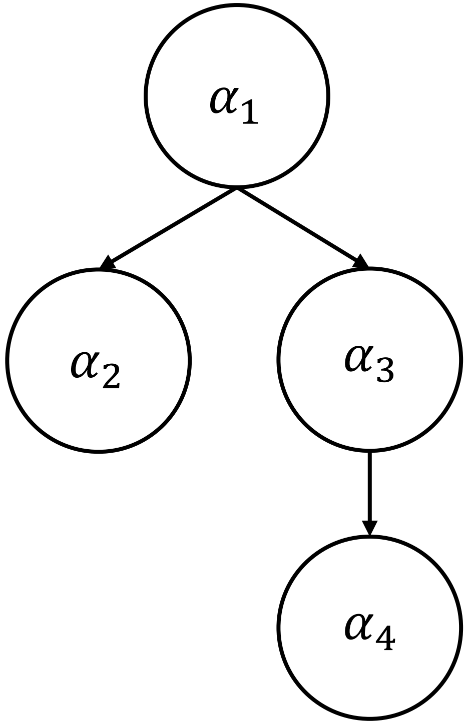
Profiles

	α_1	α_2	α_3	α_4
profile1	1	0	0	0
profile2	1	1	0	0
profile3	1	0	1	0
profile4	1	1	1	0
profile5	1	1	1	1

- Dependent each attributes → Less profiles → Reduce structural parameters to estimate

Prototypical Attribute Hierarchies II

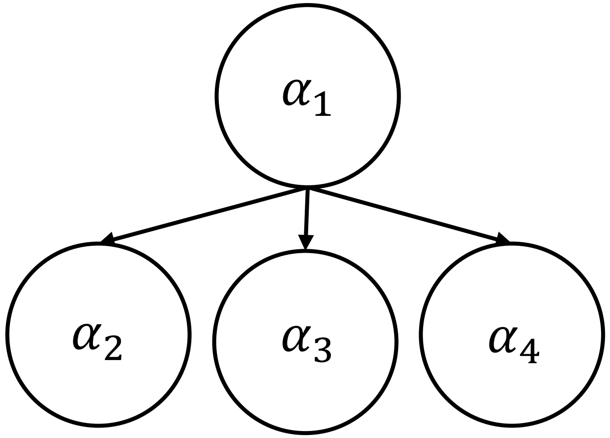
C. Divergent



Profiles

	α_1	α_2	α_3	α_4
profile1	1	0	0	0
profile2	1	1	0	0
profile3	1	0	1	0
profile4	1	1	1	0
profile5	1	0	1	1
profile6	1	1	1	1

D. Unstructured



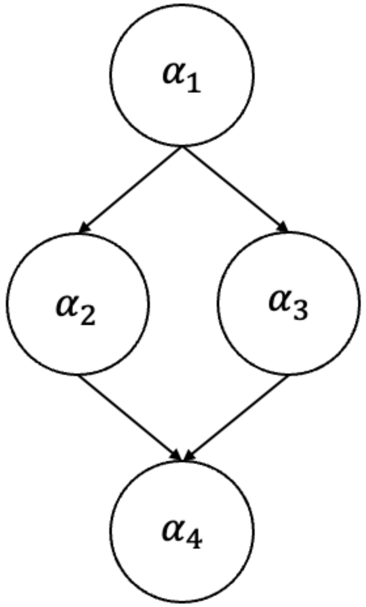
Profiles

	α_1	α_2	α_3	α_4
profile1	1	0	0	0
profile2	1	1	0	0
profile3	1	0	1	0
profile4	1	0	0	1
profile5	1	1	1	0
profile6	1	1	0	1
profile7	1	0	1	1
profile8	1	1	1	1

- Independent each attributes → More profiles

Additional Matrices Implied by Attribute Hierarchies

B. Convergent



How can we get the reduced Q-matrix of the convergent hierarchies?

1. Adjacency Matrix (**A**)

To / From

A	α_1	α_2	α_3	α_4
α_1	0	1	1	0
α_2	0	0	0	1
α_3	0	0	0	1
α_4	0	0	0	0

2. **B** = **A** OR **I** (Boolean Sum)
3. **R** = **BB** : Reachability Matrix

R	α_1	α_2	α_3	α_4
α_1	1	1	1	1
α_2	0	1	0	1
α_3	0	0	1	1
α_4	0	0	0	1

*I : Identity Matrix

$$\begin{bmatrix}1000\\0100\\0010\\0001\end{bmatrix}$$

- **A**: List hierarchically dependent attributes
- **R**: List which attributes can be reached by others

4. Pick one of all possible profiles e.g., (1,0,0,0)
5. If **R** X the profile = the profile,
   → Accept the profile
   If **R** X the profile ≠ the profile, ,
   → Reject the profile

Accept

$$\begin{bmatrix}1111\\0101\\0011\\0001\end{bmatrix}\times\begin{bmatrix}1\\0\\0\\0\end{bmatrix}=\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

6.Reduced Q-matrix

	α_1	α_2	α_3	α_4
profile1	1	0	0	0
profile2	1	1	0	0
profile3	1	0	1	0
profile4	1	1	1	0
profile5	1	1	1	1

Measurement model :
Log-linear Cognitive Diagnosis Model (LCDM)

DCMs are Constrained Latent Class Models

- Choice of structural model not dependent on the measurement component

Observed Data: Probability of observing examinee *r*'s vector of item responses to all items

Measurement Component: Product of Conditional Item Response Probabilities (Item Responses are Independent)

$$P(\mathbf{Y}_r = \mathbf{y}_r) = \sum_{c=1}^{C} \nu_c \prod_{i=1}^{I} \pi_{ic}{}^{y_{ri}} (1 - \pi_{ic})^{1-y_{ri}}$$

Structural component: Proportion of examinees in each class

Log-linear Cognitive Diagnosis Model (LCDM)

- The LCDM is a special case of a log-linear model with latent classes (Hagenaars, 1993) and thus is also a special case of the General Diagnostic Model (GDM, von Davier, 2005)
- The LCDM defines the logit of the probability of a correct response as a linear function of the attributes that have been mastered.
  - Parameterized DCMs using a linear model framework
  - Can be compensatory or non-compensatory at the item level
  - Can be estimated using Mplus

- The LCDM doesn’t need for a two-stage approach of lining up and the cutting

Item Response Function in LDCMs

Model	Response	Latent Variable	Equation
LDCM	Discrete	Discrete	$ln\left(\frac{P(y_{ic}=1\|\boldsymbol{\alpha}_c)}{P(y_{ic}=0\|\boldsymbol{\alpha}_c)}\right)=\lambda_{i,0}+\boldsymbol{\lambda}_i h(\boldsymbol{\alpha}_c,\boldsymbol{q}_i)$

The item response is predicted as a function of the set of attributes that is measured by that item.

$$ln\left(\frac{P(\boxed{y_{ic}}=1|\boldsymbol{\alpha}_c)}{P(y_{ic}=0|\boldsymbol{\alpha}_c)}\right)=\lambda_{i,0}+\boldsymbol{\lambda}_i h(\boldsymbol{\alpha}_c,\boldsymbol{q}_i)=\boxed{\lambda_{i,0}}+\sum_{a=1}^{A}\boxed{\lambda_{i,1,(a)}}\alpha_{ca}q_{ia}+\sum_{a=1}\sum_{a'>1}^{A}\boxed{\lambda_{i,2,(a,a')}}\alpha_{ca}\alpha_{ca'}q_{ia}q_{ia'}+\ \ldots$$

The **intercept** represents log-odds of a correct response for an examinee in the reference group who has not mastered any attribute. (=Guessing)

The **main effect**: increase the log-odds of a correct response given mastery of each respective attribute

The **two-way interaction** between two attributes allows the log-odds of a correct response to change given an examinee's mastery of both attributes.

Notation in LDCM

$$\lambda_{i,e,(a)}$$

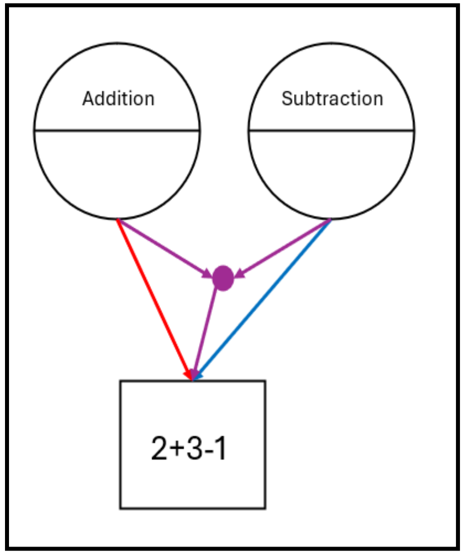
- Subscript #1 - i: the item to which parameters belong
- Subscript #2 - e: the level of the effect
  - 0: intercept
  - 1: main effect
  - 2: two-way interaction
  - 3: three-way interaction
- Subscript #3 - $(a_1, \dots)$: the attributes to which the effect applies
  - Same number of attributes listed as number in Subscript #2

$$ln\left(\frac{P(y_{ic}=1|\boldsymbol{\alpha}_c)}{P(y_{ic}=0|\boldsymbol{\alpha}_c)}\right) = \lambda_{i,0} + \boldsymbol{\lambda}_i h(\boldsymbol{\alpha}_c, \boldsymbol{q}_i) = \boxed{\lambda_{i.0}} + \sum_{a=1}^{A} \boxed{\lambda_{i,1,(a)}} \alpha_{ca} q_{ia} + \sum_{a=1}^{A} \sum_{a'>1} \boxed{\lambda_{i,2,(a,a')}} \alpha_{ca}\alpha_{ca'} q_{ia} q_{ia'} + \dots$$

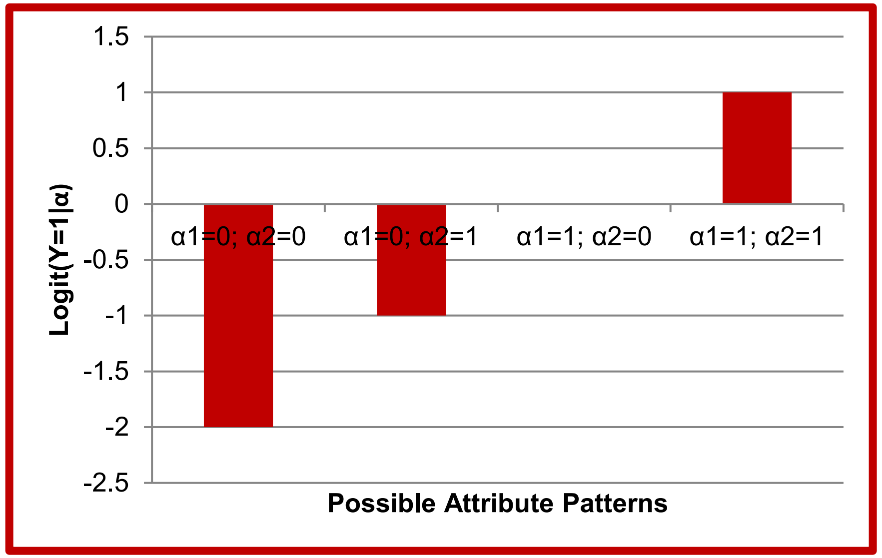
LDCM Predicted Logits and Probabilities

$$ln\left(\frac{P(y_{ic}=1|\boldsymbol{\alpha}_c)}{P(y_{ic}=0|\boldsymbol{\alpha}_c)}\right) = \lambda_{i,0} + \boldsymbol{\lambda}_i h(\boldsymbol{\alpha}_c, \boldsymbol{q}_i) = \lambda_{i,0} + \sum_{a=1}^{A} \lambda_{i,1,(a)} \alpha_{ca} q_{ia} + \sum_{a=1} \sum_{a'>1}^{A} \lambda_{i,2,(a,a')} \alpha_{ca} \alpha_{ca'} q_{ia} q_{ia'} + \ \ldots$$

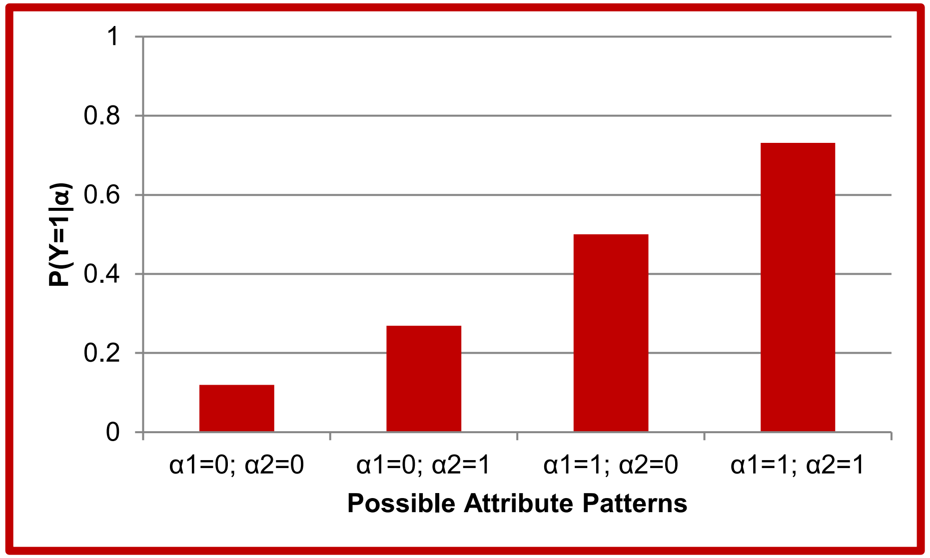
α_{c1}	α_{c2}	LCDM Logit Function	Logit	Probability
0	0	$\boldsymbol{\lambda_{i,0}} + \lambda_{i,1,(1)}(0)(1) + \lambda_{i,1,(2)}(0)(1) + \lambda_{i,2,(1,2)}(0)(0)(1)(1)$	-2	0.12
0	1	$\boldsymbol{\lambda_{i,0}} + \lambda_{i,1,(1)}(0)(1) + \boldsymbol{\lambda_{i,1,(2)}(1)(1)} + \lambda_{i,2,(1,2)}(0)(1)(1)(1)$	-1	0.27
1	0	$\boldsymbol{\lambda_{i,0}} + \boldsymbol{\lambda_{i,1,(1)}(1)(1)} + \lambda_{i,1,(2)}(0)(1) + \lambda_{i,2,(1,2)}(1)(0)(1)(1)$	0	0.5
1	1	$\boldsymbol{\lambda_{i,0}} + \boldsymbol{\lambda_{i,1,(1)}(1)(1)} + \boldsymbol{\lambda_{i,1,(2)}(1)(1)} + \boldsymbol{\lambda_{i,2,(1,2)}(1)(1)(1)(1)}$	1	0.73



Logit Response Function



Probability (Item Characteristic Bar Chart)

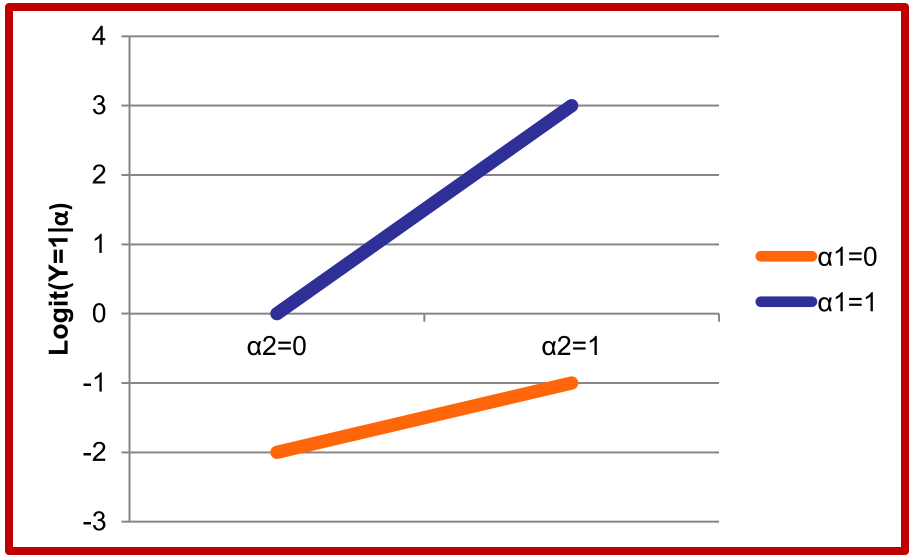


Less Extreme Interactions

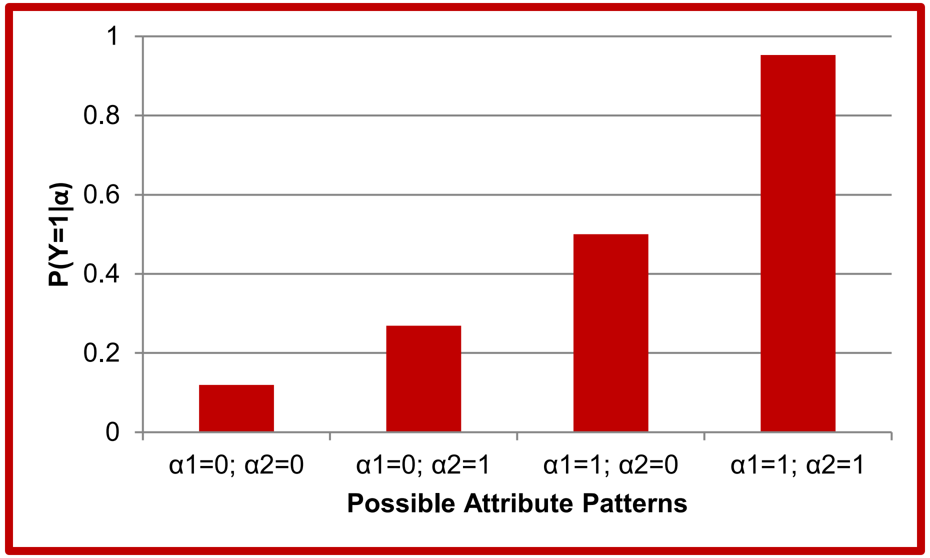
- Extreme interactions are unlikely in practice
- Positive interaction with positive main effects

$$Logit(Y_{ci} = 1|\boldsymbol{\alpha}_c) = \lambda_{i,0} + \lambda_{i,1,(1)}\alpha_{c1} + \lambda_{i,1,(2)}\alpha_{c2} + \lambda_{i,2,(1,2)}\alpha_{c1}\alpha_{c2}$$

Logit Response Function



Probability (Item Characteristic Bar Chart)



Measurement model : Subsumed Models

Previous Popular DCMs

- Because the advent of the GDM and LCDM has been fairly recent, other earlier DCMs are still in use

- Such DCMs are much more restrictive than the LCDM
  - It is anticipated that field will adapt to more general forms

- Each of the subsumed models can be fit using the LCDM
  - Fixing certain model parameters

- Shown for reference purposes
  - See Henson, Templin, & Willse (2009) for more detail

Other DCMs with the LCDM

- The big 6 – DCMs with latent variables:
  - DINA (Deterministic Inputs, Noisy ‘AND’ Gate)
    - Haertel (1989); Junker and Sijtsma (1999)
  - NIDA (Noisy Inputs, Deterministic ‘AND’ Gate)
    - Maris (1995)
  - RUM (Reparameterized Unified Model)
    - Hartz (2002)
  - DINO (Deterministic Inputs, Noisy ‘OR’ Gate)
    - Templin & Henson (2006)
  - NIDO (Noisy Inputs, Deterministic ‘OR’ Gate)
    - Templin (2006)
  - C-RUM (Compensatory Reparametrerized Unified Model)
    - Hartz (2002)

Other DCMs with the LCDM

	A low value on one latent variable **cannot** be compensated for by a high value on another latent variable.			A low value on one latent variable **can** be compensated for by a high value on another latent variable.		
LCDM Parameters	Non-compensatory Models			Compensatory Models		
	DINA	NIDA	NC-RUM	DINO	NIDO	C-RUM
Main Effects	Zero	Positive	Positive	Positive	Positive	Positive
Interactions	Positive	Positive	Positive	Negative	Zero	Zero
Equality Parameter Restrictions	Across Attributes	Across Items	---	Across Attributes	Across Items	---

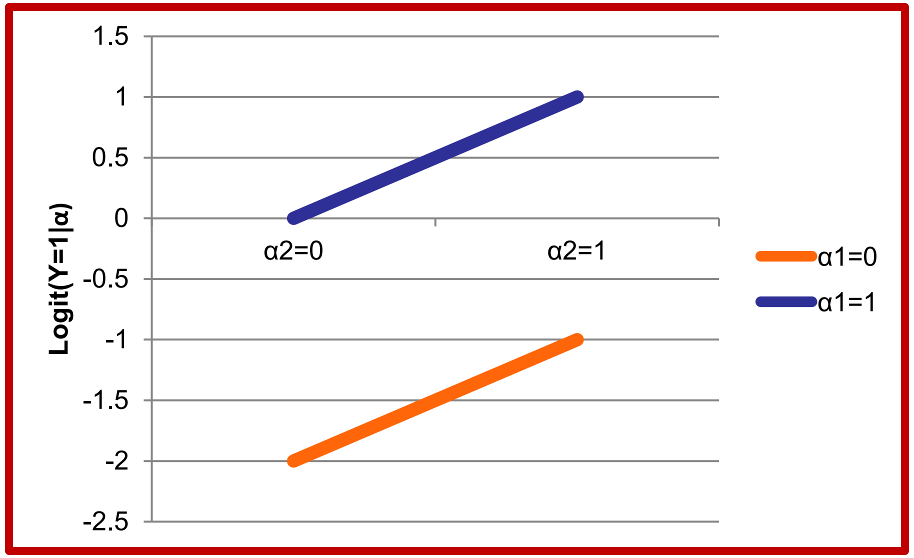
Adapted from: Rupp, Templin, Henson (2010)

C-RUM

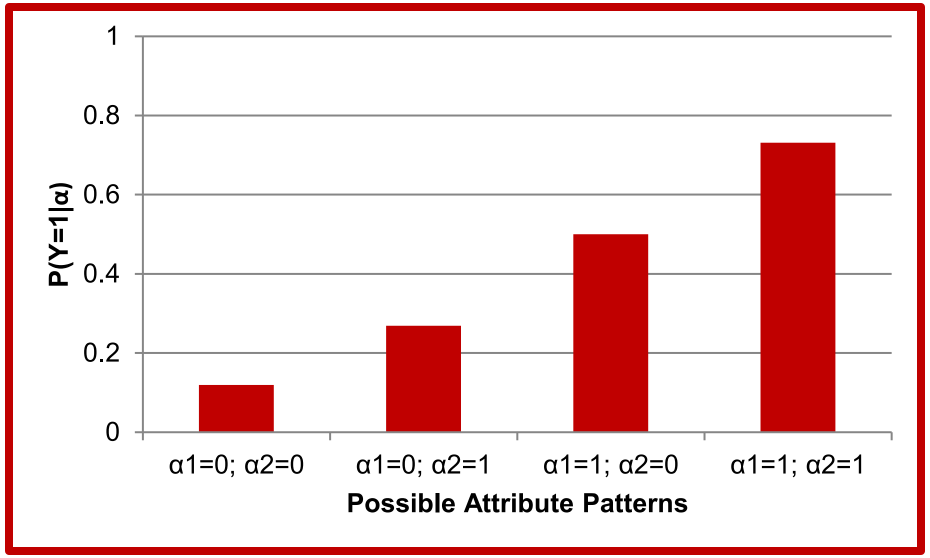
- Compensatory model
- No interactions in model: Parallel lines for the logit

$$Logit(Y_{ci} = 1|\boldsymbol{\alpha}_c) = \lambda_{i,0} + \lambda_{i,1,(1)}\alpha_{c1} + \lambda_{i,1,(2)}\alpha_{c2}$$

Logit Response Function



Probability (Item Characteristic Bar Chart)

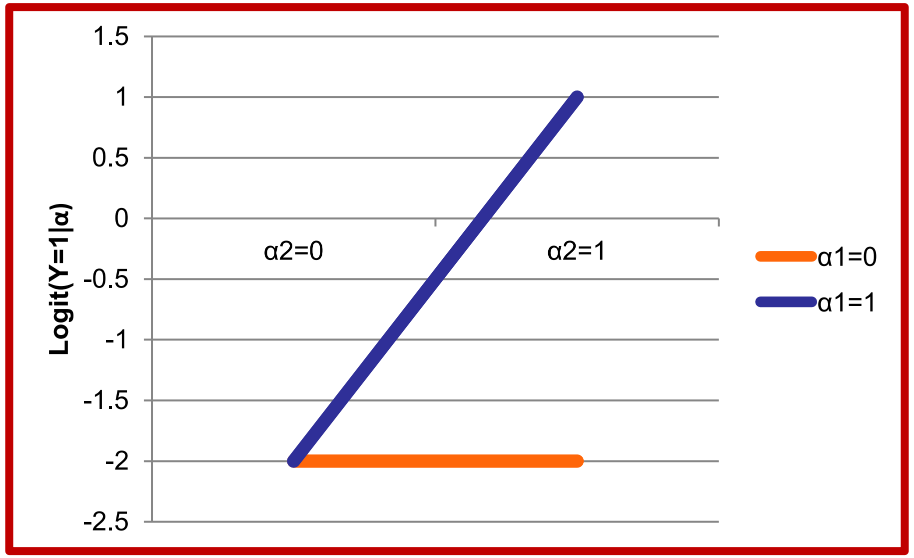


DINA Model

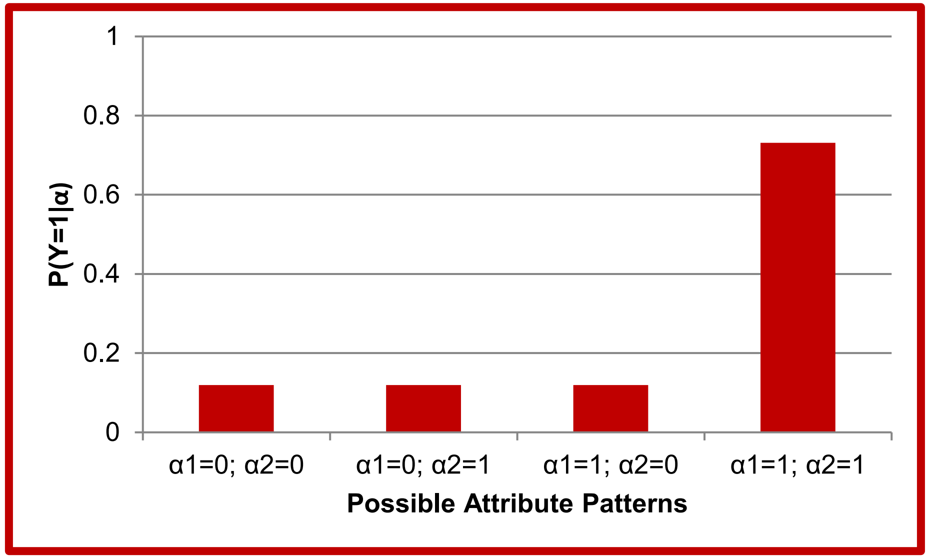
- Conjunctive model (i.e., all-or-none)
- **Positive** interaction: over-additive logit model
- Highest interaction parameter is non-zero
- All main effects (and lower interactions) zero

$$Logit(Y_{ci} = 1|\boldsymbol{\alpha}_c) = \lambda_{i,0} + \lambda_{i,2,(1,2)}\alpha_{c1}\alpha_{c2}$$

Logit Response Function



Probability (Item Characteristic Bar Chart)

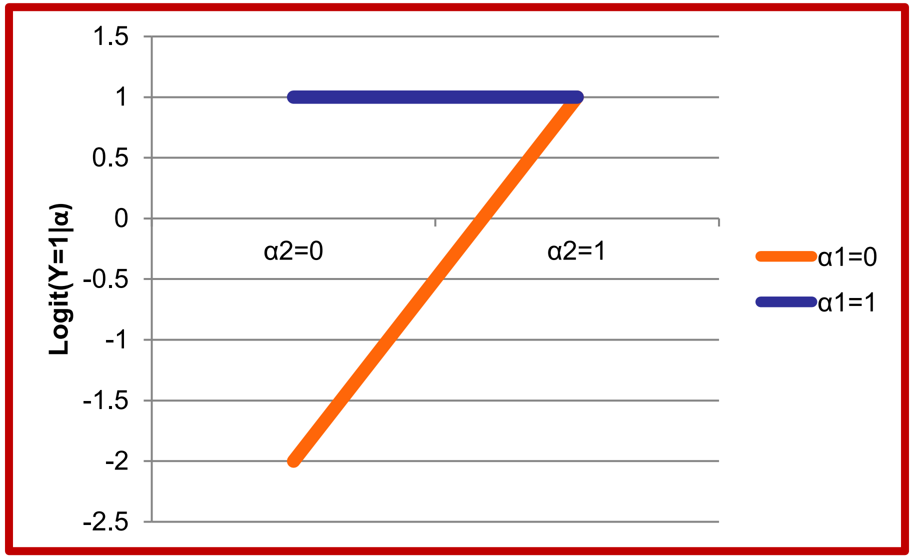


DINO Model

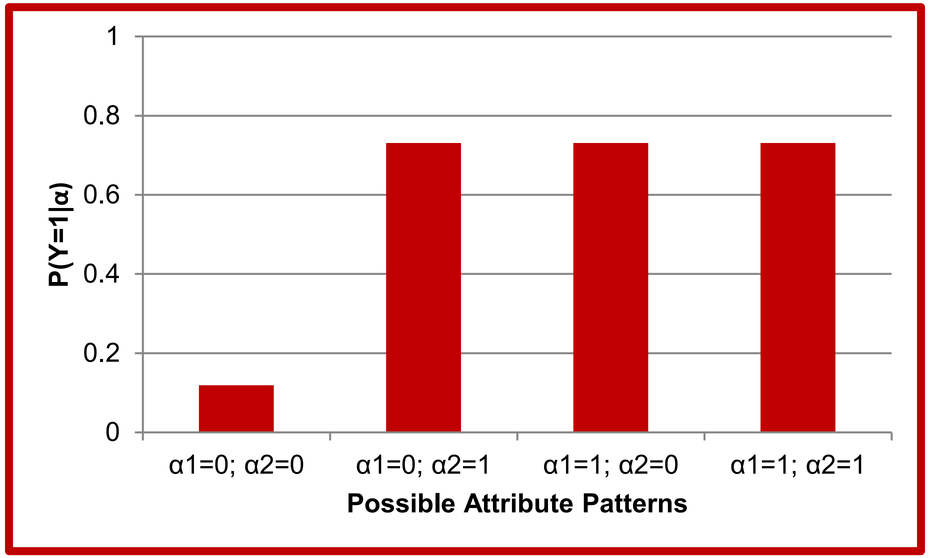
- Disjunctive model (i.e., one-or-more)
- **Negative** interaction: under-additive logit model
- All main effects are equal
- No attributes mastered: $\lambda_{i,0}$
- One or more are mastered: $\lambda_{i,0} + \lambda_{i,1}$

$$Logit(Y_{ci} = 1|\boldsymbol{\alpha}_c) = \lambda_{i,0} + \lambda_{i,1}\,\alpha_{c1} + \lambda_{i\,,1}\alpha_{c2} - \lambda_{i\,,1}\alpha_{c1}\alpha_{c2}$$

Logit Response Function



Probability (Item Characteristic Bar Chart)



Structural Models:
Modeling the Attribute Space in DCMs

Structural Models in DCMs

- For a test measuring A attributes, 2^A profiles are possible
  - The structural model provides the probability for each profile

- The **base rates** represent the probability that any respondent has a given attribute profile
  - Part of respondent diagnosis: the attribute “base rates”
  - Distribution of attribute profiles in a sample
    - Proportion of masters for any given attribute
    - Correlation of attributes

Information from Structural Model

- The parameter for the structural model is υ_c
- The base-rate is the probability of attribute profile c: $\upsilon_c = P(\boldsymbol{\alpha}_c)$
- Each attribute profile $\boldsymbol{\alpha}_c$ has one
- Such information is helpful in determining nature of attributes in a population of interest

Unstructured Structural Model

c	υ_c	α_1	α_2	α_3	α_4
1	.212	0	0	0	0
2	.070	0	0	0	1
3	.056	0	0	1	0
4	.084	0	0	1	1
5	.038	0	1	0	0
6	.026	0	1	0	1
7	.050	0	1	1	0
8	.153	0	1	1	1
9	.002	1	0	0	0
10	.003	1	0	0	1
11	.002	1	0	1	0
12	.017	1	0	1	1
13	.003	1	1	0	0
14	.011	1	1	0	1
15	.017	1	1	1	0
16	.255	1	1	1	1
	1.00	.31	.55	.63	.62

1. Proportion of Masters/ Marginal Base Rates ($P(\alpha_c)$)

2. Pairwise Estimates

	$\alpha_2 = 0$	$\alpha_2 = 1$
$\alpha_1 = 0$	0.422	0.267
$\alpha_1 = 1$	0.024	0.286

3. Tetrachoric Correlation (in R)

```
> library(psych)
> print(cont)
      [,1]  [,2]
[1,] 0.422 0.267
[2,] 0.024 0.286
> tetrachoric(cont)
Call: tetrachoric(x = cont)
tetrachoric correlation
[1] 0.78

 with tau of
[1]  0.49 -0.13
```

Differing Structural Models

- Because there are numerous υ_c parameters, interpretation is difficult
  - Useful for detecting attribute hierarchies

- The structural model of a DCM has the potential to have an overwhelming number of parameters
  - For A attributes total estimated: $2^A - 1$
  - Saturated model

- Multiple structural model exist
  - All reduce the number of parameters
  - All use categorical data analysis techniques to model υ_c

- Analogous to latent variable covariance structure in structural equation modeling
  - Distribution of attributes is categorical, not continuous
  - Can help to determine nature of attribute relationships

Types of Structural Models

- **Log-linear model**
  - Henson and Templin (2005)
  - Models the natural logarithm of υ_c by the attributes in each profile
  - Allows for varying levels of complexity
    - Most:– Saturated Model – full set of parameters
    - Least: Independent Attributes Model no parameters
  - Implemented in Mplus

- **Tetrachoric correlation model**
  - Provides an item factor model for latent attributes
  - Uses only bivariate information for pairs of attributes
  - Allows for covariance structures to be estimated
  - Not available in any software packages (see Templin & Henson, 2006)

- **Hierarchical factors model**
  - Special case of tetrachoric correlation model (see de la Torre & Douglas, 2004)

- **Mixture models**
  - Henson and Templin (2005): used to evaluate types of pathological gamblers
  - Also given by von Davier (2008)

Structural Models – Log-Linear Models

The Logic Behind Log-Linear Models

- Log-linear models take the set of probabilities from the structural model and re-express them on the log scale

$$\mu_c = \log \upsilon_c$$

- Re-expression on the log scale is convenient as these terms can now be modeled (predicted) by other features in the model
  - The attributes themselves
  - Covariates (if any)

- Because of the re-expression, redundant terms can be removed from the model
  - Simplifying estimation, improving parsimony

- In a structural model, there are 2^A probabilities…but they all add up to 1.0
  - Therefore, there can only be at most $2^A - 1$ log-linear model parameters

Log-Linear Structural Models

- The log-linear structural model is the easiest to implement with Mplus

Natural Logarithm	Back to Probabilities
$\mu_c = \log \upsilon_c$	$\upsilon_c = \dfrac{\exp(\mu_c)}{\sum_{i=1}^{2^A} \exp(\mu_i)}$
μ_c is relative	υ_c depends on the other terms

- Mplus fix the value of the last class “mean” to zero

Notation

$$\mu_c = \sum_{a=1}^{A} \gamma_{1,(a)} \alpha_{ca} + \sum_{a=1}^{A-1} \sum_{a'=1}^{A} \gamma_{2,(a,a')} \alpha_{ca} \alpha'_{ca} + \cdots + \gamma_{A,(a,a',\ldots)} \prod_{a=1}^{A} \alpha_{ca}$$

$$\gamma_{e,(a_1,\ldots)}$$

- Subscript #1 – e: the level of the effect
  - 0:would be the intercept – but we won't have one
  - 1 is the main effect
  - 2 is the two-way interaction
  - 3 is the three-way interaction
- Subscript #2 – $(a_1, \ldots)$: the attributes the effect applies to
  - Same number of attributes listed as number in Subscript #2

Log-Linear Model Explained

- For profile 1: $\boldsymbol{\alpha}_1 = [\alpha_{11} = 0; \alpha_{12} = 0; \alpha_{13} = 0; \alpha_{14} = 0]$:

$$\mu_1 = \gamma_{1,(1)}(0) + \gamma_{1,(2)}(0) + \gamma_{1,(3)}(0) + \gamma_{1,(4)}(0) + \gamma_{2,(1,2)}(0)(0) + \gamma_{2,(1,3)}(0)(0) + \gamma_{2,(1,4)}(0)(0) + \gamma_{2,(2,3)}(0)(0) + \gamma_{2,(2,4)}(0)(0) + \gamma_{2,(3,4)}(0)(0) + \gamma_{3,(1,2,3)}(0)(0)(0) + \gamma_{3,(1,2,4)}(0)(0)(0) + \gamma_{3,(2,3,4)}(0)(0)(0) + \gamma_{4,(1,2,3,4)}(0)(0)(0)(0)$$

As all attributes are zero, the predicted value of $\mu_1 = 0$
15 parameters are estimated fixing μ_1 to zero

- For profile 6: $\boldsymbol{\alpha}_6 = [\alpha_{61} = 0; \alpha_{62} = 1; \alpha_{63} = 0; \ \alpha_{64} = 1]$:

$$\mu_6 = \gamma_{1,(1)}(0) + \boxed{\gamma_{1,(2)}(1)} + \gamma_{1,(3)}(0) + \boxed{\gamma_{1,(4)}(1)} + \gamma_{2,(1,2)}(0)(1) + \gamma_{2,(1,3)}(0)(0) + \gamma_{2,(1,4)}(0)(1) + \gamma_{2,(2,3)}(1)(0) + \boxed{\gamma_{2,(2,4)}(1)(1)} + \gamma_{2,(3,4)}(0)(1) + \gamma_{3,(1,2,3)}(0)(1)(0) + \gamma_{3,(1,2,4)}(0)(1)(1) + \gamma_{3,(2,3,4)}(1)(0)(1) + \gamma_{4,(1,2,3,4)}(0)(1)(0)(1)$$

The main effects of attribute 2 and attribute 4, and interaction between attributes 2 and 4 apply

$$\mu_6 = \gamma_{1,(2)} + \gamma_{1,(4)} + \gamma_{2,(2,4)}$$

Interpretations of Model Parameters

Main Effect	Interaction	Characteristics
No	No	All attribute profiles are equally likely
Yes	No	Uncorrelated attributes
		Main effects are essentially equal to attribute base rate
Yes	Yes	Statistically equivalent to the saturated structural model

- Two-way interactions: equal to bivariate correlations in categorical models
- Higher-level interactions represent higher level of characteristics of attribute distribution (i.e. skewness, kurtosis, etc…)
- Higher-level interactions can be removed if not significantly different from zero

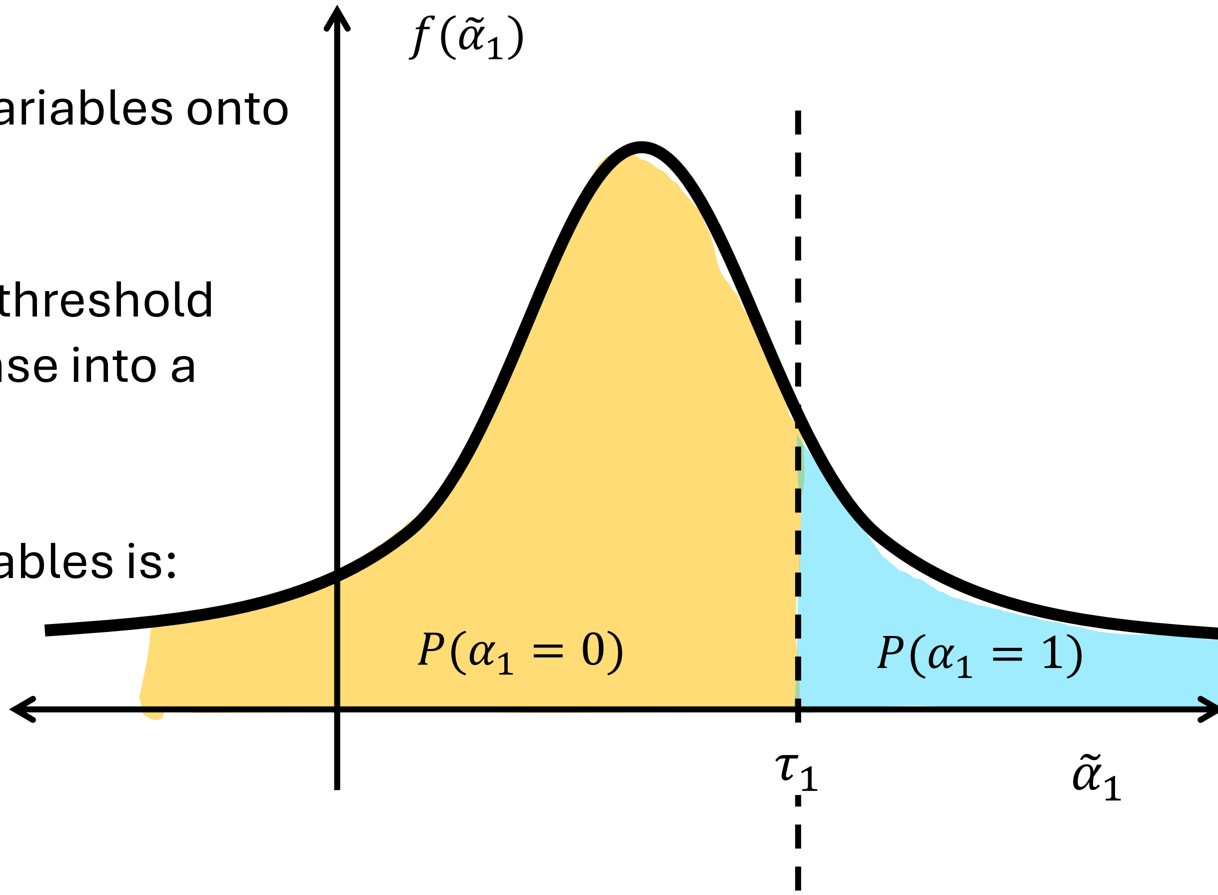
Structural Models – Tetrachoric Correlation Models

Defining Tetrachoric Correlations

- The tetrachoric correlation is a measure of the association between two **binary** variables (α_1, α_2)
- The correlation comes from mapping the binary variables onto two **underlying continuous** variables ($\tilde{\alpha}_1, \tilde{\alpha}_2$)
- Each of the continuous variables is bisected by a threshold ($\tau_{\alpha_1}, \tau_{\alpha_2}$) which transforms the continuous response into a categorical outcome(α_1, α_2)
- The distribution of the underlying continuous variables is:

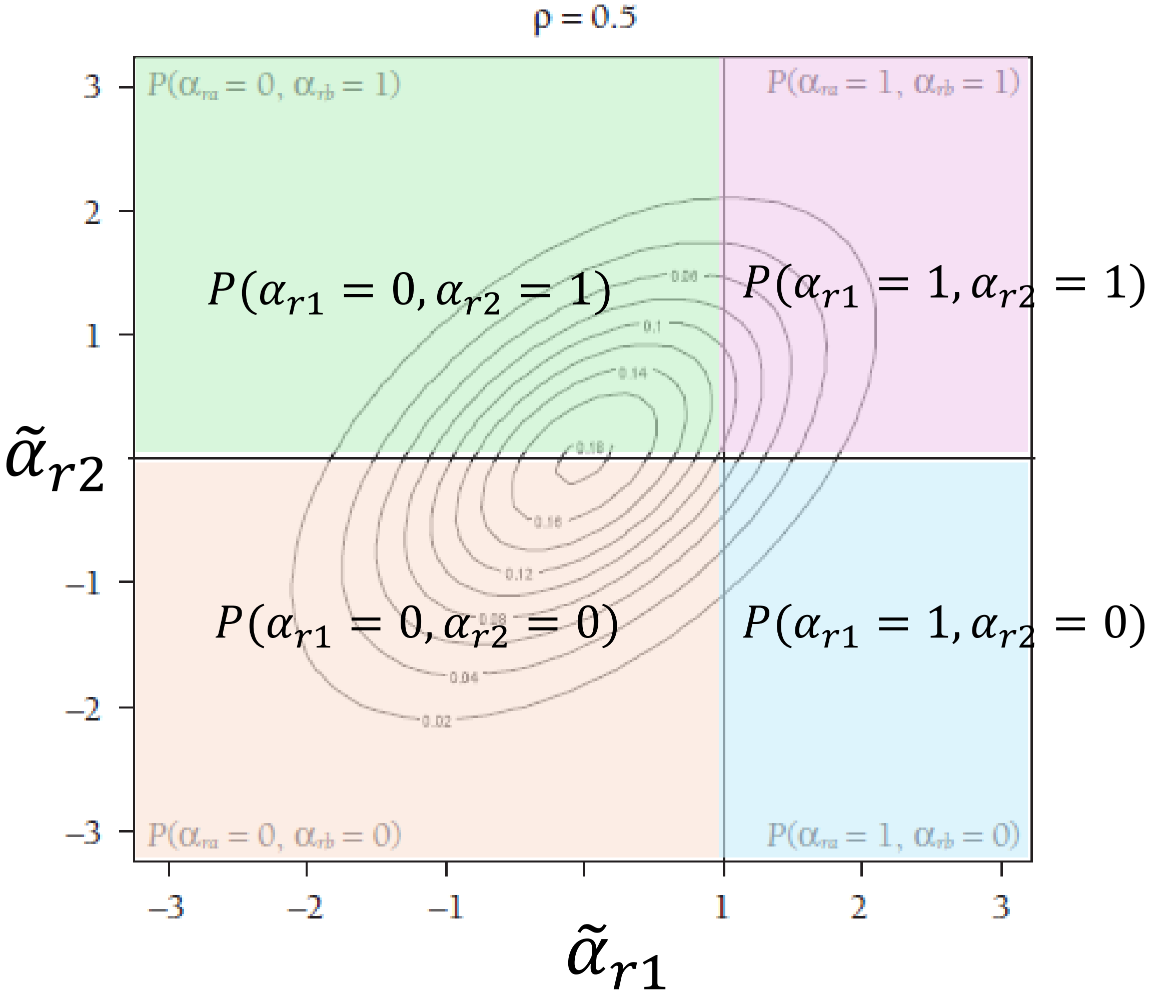
$$\begin{bmatrix}\tilde{\alpha}_1\\ \tilde{\alpha}_2\end{bmatrix} \sim N\left(\boldsymbol{\mu} = \begin{bmatrix}0\\0\end{bmatrix}, \boldsymbol{\Xi} = \begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right)$$

- ρ is the tetrachoric correlation coefficient



Tetrachoric Correlation Explained

Contingency Table		
	$\alpha_2 = 0$	$\alpha_2 = 1$
$\alpha_1 = 0$	0.422	0.267
$\alpha_1 = 1$	0.024	0.286



Note. Revised from Rupp, Templin, Henson (2010, p.180). The graph is not aligned with the contingency table on the left.

Technical Specifics: Multivariate Attributes

- The tetrachoric models assume use the following function to model the probability of an attribute profile:

$$\upsilon_c = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \dots \int_{a_A}^{b_A} \boxed{\frac{1}{(2\pi)^{A/2} \boxed{|\boldsymbol{\Xi}|}^{1/2}} \exp\left[-\frac{1}{2} \widetilde{\boldsymbol{\alpha}}' \boldsymbol{\Xi}^{-1} \widetilde{\boldsymbol{\alpha}}\right]} \partial\widetilde{\alpha}_A \dots \partial\widetilde{\alpha}_2 \partial\widetilde{\alpha}_1$$

Tetrachoric Correlation Matrix

Multivariate Normal Density

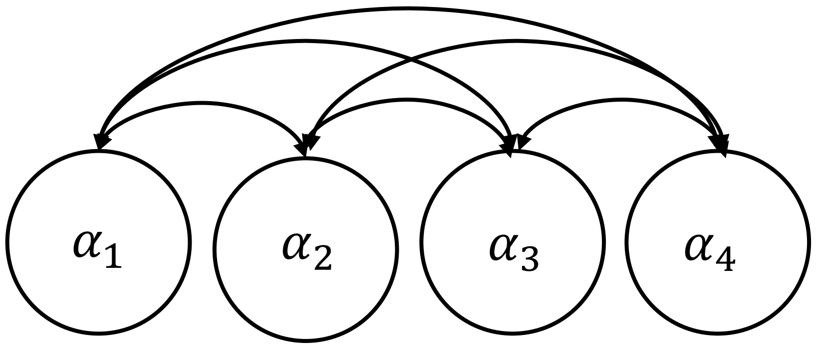
$$b_a = \begin{cases} \infty & if \quad \alpha_{cb} = 1 \\ \tau_a & if \quad \alpha_{cb} = 0 \end{cases}$$

$$a_a = \begin{cases} \tau_a & if \quad \alpha_{ca} = 1 \\ -\infty & if \quad \alpha_{ca} = 0 \end{cases}$$

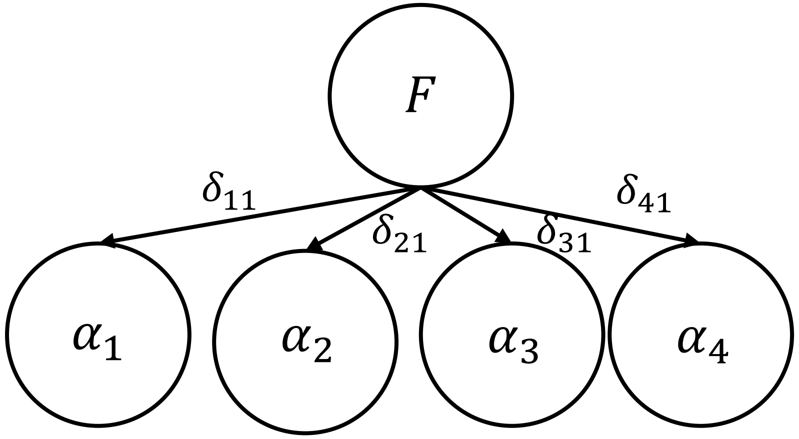
Structured Tetrachoric Models

- Placing a structure on the Ξ tetrachoric correlation matrix expands the model to mimic SEM (Templin & Henson, 2006)

Unstructured tetrachoric model	Structured tetrachoric model
$\boxed{A} + \boxed{\frac{A(A-1)}{2}}$	$\boxed{A} + \boxed{}$
# of Thresholds # of tetrachoric correlations (Ξ)	# of Thresholds Reduce # of tetrachoric correlations (Ξ)



$$\Xi = \Delta\Phi\Delta^T + \Psi$$

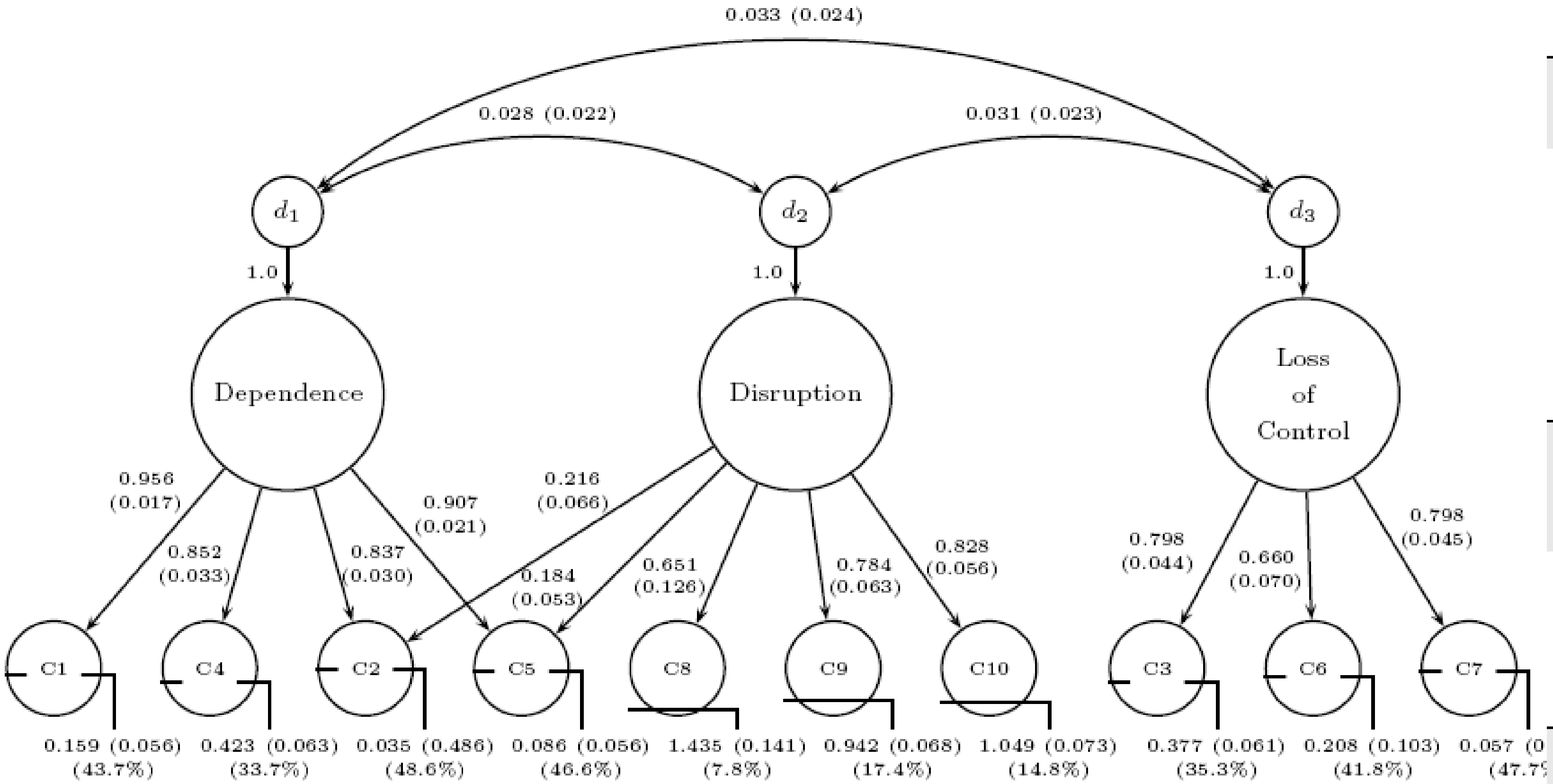


$$\tilde{\alpha}_{ra} = \delta_{a1} f_{r1} + \cdots + \delta_{a1} f_{r1} + \epsilon_{ra}$$

e.g.) When A=4, # of parameters: $4 + \frac{4(4-1)}{2} = 10$

e.g.) When A=4, # of parameters: $4 + 4 = 8$

Structured Tetrachoric Structural Model



Note. From Rupp, Templin, Henson (2010), p.180

Model	**# of parameters**
Structured Tetrachoric Structural Model	
10 Latent thresholds 12 Factor loadings 3 Higher-order interfactor correlations	25
Unstructured Tetrachoric Structural Model	
$A + \frac{A(A-1)}{2} = 10 + \frac{10 * 9}{2}$	55
Unstructured Structural Model	
$2^A - 1 = 2^{10} - 1$	1023

Take Away

1. DCMs are a family of statistical models designed to analyze discrete item responses and discrete latent variables.
2. Reliability of estimates from DCMs are high with multiple latent variables, sacrificing continuous measurement.
3. LCDM is a general measurement model of DCMs subsuming other models.
4. The subsumed models can be modeled by constraining the parameters in LCDM.
5. The Structural model is the other essential component modeling the attributes spaces.
6. Through the structural model, the total number of parameters can be reduced.