

G-DINA: the Generalized DINA model

Renato da Silva Fernandes^{1,2}

¹ Doutorando do Programa Interinstitucional de Pós-Graduação em Estatística (PIPGEs) da UFSCar-USP

² Docente do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo - IFSP

16 de Outubro de 2023

- 1 Introdução
- 2 O modelo G-DINA
- 3 Proposta de restrição dos parâmetros
- 4 A abordagem Bayesiana
- 5 Estudo de simulação
- 6 Aplicação
- 7 Referências

Modelos de Diagnóstico Cognitivo (CDM, em inglês) é uma classe de modelos de variáveis latentes discretas que tem como objetivo determinar quais atributos ou habilidades um indivíduo possui ou não possui com base nas respostas desse indivíduo em um teste (George and Robitzsch, 2015).

Aplicações:

- avaliação educacional (Chen and Chen, 2016; Li, Hunter, and Lei, 2016; Yamaguchi and Okada, 2018).
- avaliação psicométrica (Chen et al., 2019).
- avaliação de desordens psicológicas (Templin and Henson, 2006; de la Torre, Ark, and Rossi, 2018; Silva et al., 2018).

O G-DINA (*Generalized Deterministic inputs, noise “AND” gate*) é uma generalização do modelo DINA proposta por de la Torre (2011). O G-DINA é um modelo geral e diversos CDMs podem ser obtidos como casos especiais desse modelo.

Dentre eles estão:

- DINA (*Deterministic Input Noise “AND” gate*)
- DINO (*Deterministic Input Noise “OR” gate*)
- A-CDM (*Additive CDM*)
- R-RUM (*Reduced Reparameterized Unified Model*)
- LCDM (*Log-linear CDM*)
- LLM (*Linear Logistic Model*)

Considere um questionário com J itens que avaliam K atributos respondidos por N indivíduos. Definimos:

Respostas

Y é uma matriz de dimensão $J \times K$ no qual cada componente Y_{ij} , com $j = 1, \dots, J$, é uma variável binária observável definida como:

$$Y_{ij} = \begin{cases} 1, & \text{se o indivíduo } i \text{ responde ao item } j \text{ corretamente;} \\ 0, & \text{caso contrário} \end{cases}$$

matriz Q

Q é uma matriz binária de dimensão $J \times K$ de valores conhecidos - normalmente definida por especialistas no(s) assunto(s) do questionário - onde o componente q_{jk} , com $j = 1, \dots, J$ e $k \in 1, \dots, K$, definido como:

$$q_{jk} = \begin{cases} 1, & \text{se o atributo } k \text{ é requerido pelo item } j; \\ 0, & \text{caso contrário} \end{cases}$$

Notação:

$\mathbf{q}_j = (q_{j1}, \dots, q_{jK})^T$ é a j -ésima linha da matriz Q .

vetor de atributos

$\alpha_i = (\alpha_{i1}, \dots, \alpha_{iK})$ é o *vetor de atributos* do indivíduo $i = 1, \dots, N$, onde o componente α_{ik} , $k = 1, \dots, K$ é definido como:

$$\alpha_{ik} = \begin{cases} 1, & \text{se o indivíduo } i \text{ possui o atributo } k; \\ 0, & \text{caso contrário.} \end{cases}$$

Função de Resposta ao Item

A *Função de Resposta ao Item* (IRF, em inglês) é a função que relaciona a variável latente que representa a habilidade do indivíduo - no caso, o vetor de atributos - e a probabilidade de sucesso no item em questão.

No caso dos CDMs, a IRF é definida (implicitamente ou explicitamente) em função da quantidade:

$$P_j(\alpha_i) := Pr(Y_{ij} = 1 | \alpha_i)$$

onde $P_j(\alpha_i)$ a probabilidade de um indivíduo com vetor de atributos α_i responder o item j corretamente.

O modelo DINA possui dois parâmetros estruturais para cada item j , denominados *guessing* (g_j) e *slipping* (s_j).

A IRF do modelo DINA é dado pela equação

$$P_j(\alpha_i) = Pr(Y_{ij} = 1 | \alpha_i) = g_j^{1-\eta_{ij}} (1 - s_j)^{\eta_{ij}}$$

Onde $\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}}$, ou alternativamente:

$$\eta_{ij} = \begin{cases} 1, & \text{se } \alpha_{ik} = 1 \text{ para todo } k : q_{jk} = 1 \\ 0, & \text{c.c} \end{cases}$$

Note que no modelo DINA, a IRF pode assumir apenas dois valores distintos, que dependem unicamente do valor de η_{ij}

$$\eta_{ij} = 0 \implies P_j(\alpha_i) = g_j$$

$$\eta_{ij} = 1 \implies P_j(\alpha_i) = (1 - s_j)$$

O modelo G-DINA é uma generalização do modelo DINA, onde a definição de IRF é flexibilizada, podendo assumir uma maior quantidade de valores, de acordo com o vetor reduzido de atributos α_{ij}^* , que será apresentado a seguir.

Vetor reduzido de atributos

$\alpha_{ij}^* = (\alpha_{i1}^*, \dots, \alpha_{iK_j^*}^*)$ é o *vetor reduzido de atributos* do indivíduo i com respeito ao item j , onde $K_j^* = \sum_{k=1}^K q_{jk}$ é o número de atributos requeridos pelo item j .

α_{ij}^* é obtido tomando o vetor (completo) de atributos α_i e mantendo apenas os componentes requeridos para responder o item j (isto é, mantendo os componentes α_{ik} tais que $q_{jk} = 1$).

Exemplo ilustrativo ($K=4$)

q_j	vetor de atributos (α_i)	vetor reduzido de atributos (α_{ij}^*)
(1,0,0,1)	(1,1,1,0)	(1,0)
(1,0,0,1)	(1,0,1,0)	(1,0)
(1,0,0,1)	(0,1,1,0)	(0,0)

A seguir, definir o conceito de classes e grupos latentes que serão amplamente utilizados desse ponto em diante.

Note que existe uma quantidade finita de valores que o vetor de atributo, α_i , e o vetor reduzido de atributos, bas , podem assumir.

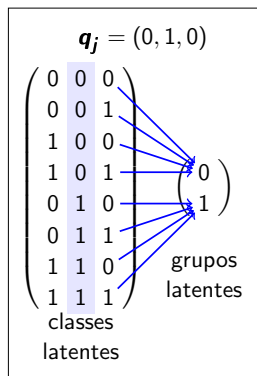
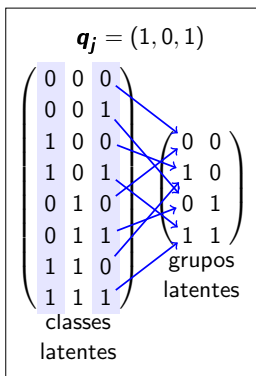
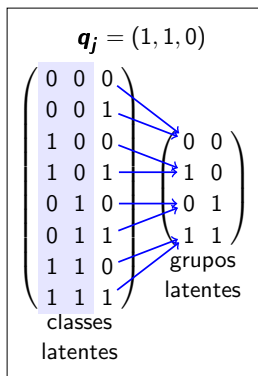
Existem $C = 2^K$ possíveis valores para o vetor α_i . Sendo assim, podemos tratar o vetor de atributos como uma variável categórica. Desse modo, utilizaremos $\alpha_{(c)}$ para denotar a classe latente $l = 1, \dots, C$.

Analogamente, existem $C_j = 2^{K_j^*}$ possíveis valores para o vetor α_{ij}^* . e utilizaremos $\alpha_{(l)j}^*$ para denotar o grupo latente $l = 1, \dots, C_j$ em relação ao item j .

Vetor reduzido de atributos

Para cada item j , o conjunto dos perfis latentes $\{\alpha_{(1)}, \dots, \alpha_{(C)}\}$ é mapeado de forma unívoca com o conjunto de grupos latentes $\{\alpha_{(1)j}^*, \dots, \alpha_{(C)j}^*\}$. Esse mapeamento depende apenas de \mathbf{q}_j .

Abaixo temos a ilustração de alguns mapeamentos com $K = 3$.



No G-DINA, a probabilidade $P_j(\alpha_i) = Pr(Y_{ij} = 1|\alpha_i)$ depende apenas do vetor reduzido de atributos α_{ij}^* . Isto é:

$$P_j(\alpha_i) = Pr(Y_{ij} = 1|\alpha_i) = Pr(Y_{ij} = 1|\alpha_{ij}^*) = P(\alpha_{ij}^*)$$

onde $P(\alpha_{ij}^*)$ representa a probabilidade do indivíduo que possui vetor de atributos reduzidos α_{ij}^* responder ao item j corretamente.

Definição do Modelo G-DINA

A IRF do modelo G-DINA é dada pelo *Modelo Linear Generalizado* abaixo:

$$g(P(\alpha_{ij}^*)) := \underbrace{\delta_{j0}}_{\text{intercepto}} + \underbrace{\sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{ik}^*}_{\text{efeitos principais}} + \underbrace{\sum_{k=1}^{K_j^*-1} \sum_{k'=k+1}^{K_j^*} \delta_{jkk'} \alpha_{ik}^* \alpha_{ik'}^*}_{\text{Efeitos de interação de 1ª ordem}} + \dots + \underbrace{\delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{ik}^*}_{\text{Efeitos de interação de ordem } K_j^* - 1}$$

onde:

- $\delta_{j0}, \delta_{j1}, \dots, \delta_{j12\dots K_j^*}$ são os parâmetros estruturais do item j .
- os componentes do vetor reduzido de atributos α_{ij}^* fazem papel de variáveis preditoras.
- g é uma função de ligação, que pode ser a função *identidade*, *log* ou *logit* (de la Torre, 2011)

Interpretação dos parâmetros

Considerando a função de ligação identidade, os parâmetros δ_j podem ser interpretados da seguinte maneira:

- δ_{j0} é o intercepto do item j e representa a probabilidade de *linha de base*, que, nesse contexto, corresponde à probabilidade de um indivíduo responder corretamente o item j , quando esse indivíduo não possui nenhum dos atributos requeridos por esse item.
- δ_{jk} é o efeito principal do componente α_{ik} e representa a diferença entre probabilidade de sucesso no item j de um indivíduo, que, dentre os atributos requeridos por esse item, possui apenas o k -ésimo atributo, e a probabilidade de linha de base, δ_{j0} .
- $\delta_{jkk'}$ é o efeito de interação (de primeira ordem) entre os componentes α_{ik} e $\alpha_{ik'}$ e representa a diferença entre probabilidade de sucesso no item j de um indivíduo, que, dentre os atributos requeridos por esse item, possui apenas k -ésimo e k' -ésimo atributos, e a soma da probabilidade de linha de base δ_{j0} e dos efeitos principais, δ_{jk} e $\delta_{jk'}$.
- E assim por diante...

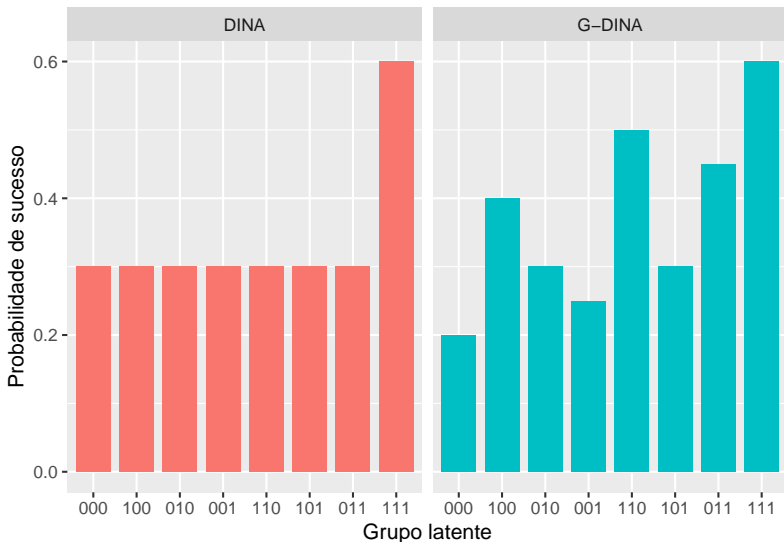
Para ilustrar a expressão da IRF do modelo G-DINA, considere um item que requer 3 atributos (ie, $K_j^* = 3$). Considere também que a função de ligação g utilizada é a função identidade.

- $\alpha_{ij}^* = (0, 0, 0) \implies P(\alpha_{ij}^*) = \delta_{j0}$
- $\alpha_{ij}^* = (0, 1, 0) \implies P(\alpha_{ij}^*) = \delta_{j0} + \delta_{j2}$
- $\alpha_{ij}^* = (1, 1, 0) \implies P(\alpha_{ij}^*) = \delta_{j0} + \delta_{j1} + \delta_{j2} + \delta_{j12}$
- $\alpha_{ij}^* = (1, 1, 1) \implies P(\alpha_{ij}^*) = \delta_{j0} + \delta_{j1} + \delta_{j2} + \delta_{j3} + \delta_{j12} + \delta_{j13} + \delta_{j23} + \delta_{j123}$

De forma geral, os parâmetros δ_j podem assumir quaisquer valores, desde que todas as possíveis combinações que surgem da IRF resultem em valores dentro do intervalo $(0, 1)$.

Comparação DINA e G-DINA

Exemplo ilustrativo da probabilidade de sucesso das classes latentes dos modelos DINA e G-DINA para um item com $K_j^* = 3$.



Como dito anteriormente, o modelo DINA pode ser obtido como um caso especial do modelo G-DINA.

Para isso, basta tomar a função de ligação identidade e fixar todos os parâmetros dos items igual a 0, exceto δ_{j0} e $\delta_{j12\dots K_j^*}$.

Desse modo, a IRF do modelo DINA pode ser escrita como:

$$P(\alpha_{ij}^*) = \delta_{j0} + \delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{ik}^*$$

Forma matricial e parâmetros estruturais

Sem nenhuma restrição, cada item j possui C_j parâmetros a serem estimados.

Para cada item j , temos a seguinte equação matricial:

$$g(\mathbf{P}_j) = \mathbf{M}_j \delta_j$$

Onde:

- **Vetor de probabilidades:** $\mathbf{P}_j = (p_{j1}, \dots, p_{jC_j})^T$ é um vetor de dimensão C_j sendo $p_{jl} = P(\alpha_{(l)j}^*)$, $l = 1, \dots, C_j$ a probabilidade de sucesso no item j , dado que o indivíduo possui vetor reduzido de atributos igual à $\alpha_{(l)j}^*$ em relação à esse item.
- **Vetor de parâmetros:** $\delta_j = (\delta_{j0}, \delta_{j1}, \dots, \delta_{j12\dots K_j^*})^T$ é um vetor de dimensão C_j
- **Matriz de design:** \mathbf{M}_j é uma matriz binária de dimensão $C_j \times C_j$ que relaciona os grupos latentes $\alpha_{(l)j}^*$ com o vetor de parâmetros δ_j .

Exemplo ilustrativo: $K_j^* = 3$ e função de ligação identidade

$$\mathbf{P}_j = \mathbf{M}_j \cdot \boldsymbol{\delta}_j$$
$$\begin{pmatrix} p_{j1} \\ p_{j2} \\ p_{j3} \\ p_{j4} \\ p_{j5} \\ p_{j6} \\ p_{j7} \\ p_{j8} \end{pmatrix} = \begin{pmatrix} P((0, 0, 0)) \\ P((1, 0, 0)) \\ P((0, 1, 0)) \\ P((0, 0, 1)) \\ P((1, 1, 0)) \\ P((1, 0, 1)) \\ P((0, 1, 1)) \\ P((1, 1, 1)) \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \delta_{j0} \\ \delta_{j1} \\ \delta_{j2} \\ \delta_{j3} \\ \delta_{j12} \\ \delta_{j13} \\ \delta_{j23} \\ \delta_{j123} \end{pmatrix}$$

Ponto chave

\mathbf{M}_j é invertível e cada componente de \mathbf{P}_j pode variar livremente entre 0 e 1, independentemente dos demais componentes. Por essas razões, na versão saturada do G-DINA, podemos utilizar os componentes de \mathbf{P}_j como parâmetros estruturais do item j .

Estimação dos parâmetros na abordagem clássica

Na abordagem clássica, os parâmetros dos itens do modelo G-DINA são obtidos à partir da maximização da função de log-verossimilhança marginalizada.

Verossimilhança completa:

$$L_Y(\mathbf{A}; \mathbf{P}) := P(\mathbf{Y}|\mathbf{A}; \mathbf{P}) = \prod_{i=1}^N \prod_{j=1}^J P(\alpha_{ij}^*)^{Y_{ij}} (1 - P(\alpha_{ij}^*))^{1-Y_{ij}}$$

log-verossimilhança marginalizada:

$$l_Y(\pi; \mathbf{P}) = \log \prod_{i=1}^N \sum_{c=1}^C p(\alpha_{(c)}) \prod_{j=1}^J P_j(\alpha_{(c)})^{Y_{ij}} (1 - P_j(\alpha_{(c)}))^{1-Y_{ij}}$$

onde $p(\alpha_{(c)})$ representa a probabilidade a priori do indivíduo i pertencer à classe c e $P_j(\alpha_{(c)}) = Pr(Y_{ij} = 1 | \alpha_{(c)})$.

A função $l_Y(\pi; \mathbf{P})$ é maximizada utilizando um algoritmo EM (Expectation - Maximization).

O modelo G-DINA é um modelo bastante flexível e, na forma apresentada, não impõe qualquer restrição sobre as probabilidades p_{jl} , além de terem que pertencer ao intervalo $(0, 1)$.

Esta falta de restrições permite o surgimento de resultados contra intuitivos.

Exemplo: Indivíduos que possuem todos os atributos exigidos para um item tenham menor probabilidade de sucesso neste item do que indivíduos que não possuem nenhum desses atributos.

Esta situação é incomum em aplicações reais e pode ser indesejável em muitos casos. Para evitar alguns resultados contra-intuitivos, como a situação anteriormente mencionada, podemos incluir restrições sobre os parâmetros estruturais.

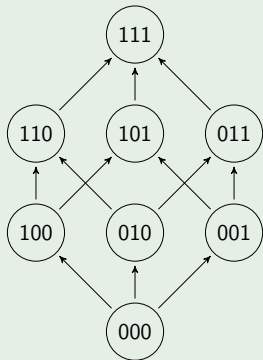
Proposta existente na literatura:

- **Restrição monotônica (MC):** (de la Torre, 2011) O domínio de um atributo adicional implica o aumento (ou inalteração) da probabilidade de sucesso nesse item.
- **Restrição da menor probabilidade (LPC):** (Fernandes, Bazán, and Cúri, 2022) Indivíduos que não possuem nenhum dos atributos requeridos pelo item j tem a menor probabilidade de sucesso nesse item.
- **Restrição da maior probabilidade (HPC):** (Fernandes, Bazán, and Cúri, 2022) Indivíduos que possuem todos os atributos requeridos pelo item j tem a maior probabilidade de sucesso nesse item.

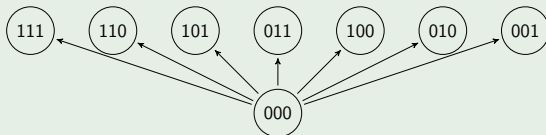
Ilustração das propostas de restrição

Ilustração das restrições impostas para um item com $K_j^* = 3$

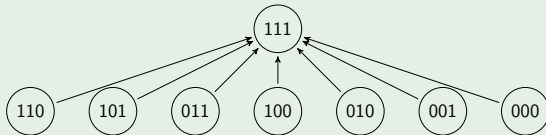
Restrição monotônica



Restrição da menor probabilidade



Restrição da maior probabilidade



Para a abordagem Bayesiana, dividimos os parâmetros em dois níveis: os parâmetros relacionados aos indivíduos e os relacionados aos itens.

Parâmetros dos indivíduos

$$\alpha_i | \pi \sim \text{Categórica}(\pi)$$

$$\text{para } i = 1, \dots, N$$

$$\pi \sim \text{Dirichlet}(\lambda)$$

onde $\lambda = (\lambda_1, \dots, \lambda_C)$, com $\lambda_c > 0$, para $c = 1, \dots, C$.

- $\pi = (\pi_1, \dots, \pi_C)$ representa o vetor de proporções das classes latentes.
- λ é o parâmetro de concentração de π .
- Para obter uma priori pouco informativa, tomamos $\lambda = (1, \dots, 1)$.

Parâmetros dos itens

$$Y_{ij} | \alpha_{ij}^* = \alpha_{(l)j}^*, p_{jl} \sim \text{Bernoulli}(p_{jl})$$

$i = 1, \dots, N; j = 1, \dots, J$, onde $l = l(i, j)$ é tal que $\alpha_{ij}^* \equiv \alpha_{(l)j}^*$

$$p_{jl} \sim \text{Beta}(a_{jl}, b_{jl})$$

$j = 1, \dots, J$, e $l = 1, \dots, C_j$, para cada j .

- Para obter prioris pouco informativa, tomamos $a_{jl} = b_{jl} = 1$ para todo j e todo l .
- Para garantir a identificabilidade do modelo, impomos a *restrição da menor probabilidade*.
- Para obter uma amostra aproximada da distribuição a posteriori de todos os parâmetros, foi utilizado um algoritmo MCMC *Gibbs Sampling* implementado com o software **JAGS**.

O segundo estudo de simulação foi realizado para avaliar a performance do método de estimação Bayesiano na recuperação dos parâmetros e compará-la à performance da implementação frequentista.

Scenario	J	K	Satisfies the Higher Probability Constraints?	Satisfies the Lower Probability Constraints?	Satisfies the Monotonic Constraints?
I	10	3	yes	yes	yes
II	30	5	yes	yes	yes
III	20	4	yes	yes	no
IV	20	4	yes	yes	no

Tamanho da amostra (N): 50, 200, 1000

Número de replicações: 100

Obs.: Para os cenários III e IV foi utilizado a mesma matriz Q , mas com diferentes valores para os parâmetros dos itens.

Medidas de acurácia:

- 1 Parâmetros estruturais
 - Média da raiz do erro quadrático médio (RMSE) sobre todos os parâmetros dos itens.
- 2 Vetor de atributos
 - Média do AACCR (*Average Attribute Correct Classification Ratio*).
 - Média do PCCR (*Pattern Correct Classification Ratio*).

Table: Resultados do Estudo de simulação - RSME médio

Scenario	N	MMLE		MMLE - MC		Bayesian - HPC		Bayesian - LPC	
I	50	0.199	(0.041)	0.152	(0.021)	0.126	(0.019)	0.148	(0.022)
	200	0.097	(0.018)	0.087	(0.014)	0.080	(0.012)	0.079	(0.013)
	1000	0.040	(0.007)	0.038	(0.006)	0.038	(0.007)	0.038	(0.007)
II	50	0.195	(0.018)	0.144	(0.013)	0.136	(0.010)	0.134	(0.010)
	200	0.095	(0.008)	0.076	(0.006)	0.080	(0.006)	0.078	(0.005)
	1000	0.037	(0.002)	0.032	(0.002)	0.036	(0.002)	0.037	(0.002)
III	50	0.209	(0.022)	0.202	(0.021)	0.131	(0.015)	0.142	(0.013)
	200	0.103	(0.012)	0.132	(0.006)	0.086	(0.009)	0.090	(0.031)
	1000	0.042	(0.004)	0.110	(0.001)	0.041	(0.004)	0.039	(0.004)
IV	50	0.303	(0.027)	0.256	(0.024)	0.156	(0.010)	0.151	(0.010)
	200	0.212	(0.029)	0.197	(0.026)	0.121	(0.018)	0.119	(0.014)
	1000	0.074	(0.009)	0.096	(0.007)	0.063	(0.006)	0.064	(0.006)

* MMLE: Marginalized Maximum Likelihood Estimation (without constraints on the parameters); MMLE - MC: MMLE with the monotonic constraints; Bayesian - HPC: Bayesian estimation method with the higher probability constraints; Bayesian - LPC: Bayesian estimation method with the lower probability constraints; the value inside parenthesis is the standard deviation.

Table: Resultados do estudo de simulação -ACCR e PCCR médios

Scenario	N	MMLE		MMLE - MC		Bayesian - HPC		Bayesian - LPC	
		AACCR	PCCR	AACCR	PCCR	AACCR	PCCR	AACCR	PCCR
I	50	86.1 %	65.1 %	86.6 %	66.7 %	86.1 %	64.9 %	85.2 %	62.1 %
	200	88.5 %	71.0 %	88.6 %	71.2 %	88.9 %	71.7 %	89.0 %	72.0 %
	1000	90.1 %	74.9 %	90.1 %	75.0 %	90.1 %	74.9 %	90.1 %	74.8 %
II	50	91.6 %	66.9 %	91.9 %	67.7 %	91.9 %	68.2 %	91.7 %	67.1 %
	200	93.3 %	73.2 %	93.8 %	74.5 %	93.8 %	75.2 %	93.9 %	75.1 %
	1000	94.6 %	77.8 %	94.6 %	78.0 %	94.6 %	77.9 %	94.6 %	77.8 %
III	50	90.3 %	71.5 %	89.1 %	65.1 %	91.2 %	74.0 %	90.8 %	72.8 %
	200	93.0 %	79.2 %	92.4 %	75.5 %	93.4 %	80.3 %	92.8 %	78.5 %
	1000	94.4 %	83.3 %	93.6 %	79.5 %	94.4 %	83.4 %	94.4 %	83.4 %
IV	50	71.7 %	30.2 %	71.2 %	27.3 %	72.6 %	35.9 %	72.7 %	34.4 %
	200	73.2 %	37.3 %	73.0 %	31.5 %	76.3 %	45.3 %	76.1 %	43.7 %
	1000	78.3 %	50.5 %	77.9 %	45.8 %	78.6 %	51.3 %	78.5 %	51.1 %

Parâmetros estruturais:

O método de estimação bayesiano (com ambas as restrições) possui precisão semelhante ou melhor comparado ao MMLE sem restrições, sendo a diferença em favor do bayesiano é mais perceptível com o pequeno tamanho da amostra. O método bayesiano apresenta precisão semelhante ou melhor em comparação ao MMLE com restrições monotônicas; neste caso, o método bayesiano é superior (apenas) nos cenários III e IV.

Atributos latentes

Todos os métodos recuperam os atributos dos indivíduos (ACCR) com precisão semelhante e não há diferença relevante entre o método de estimação Bayesiano (com ambas restrições) e métodos MMLE neste critério.

O método de estimação bayesiano (com ambas as restrições) apresenta PCCR semelhante aos métodos MMLE (com e sem o MC) nos cenários I e II.

Por outro lado, o método bayesiano apresenta PCCR um pouco melhor que o MMLE sem restrição para o tamanho amostral pequeno no cenário III e para o tamanho amostral pequeno e médio no cenário IV.

Além disso, o método bayesiano apresenta PCCR consideravelmente melhor para os cenários III e IV, quando comparado ao MMLE com a restrição monotônica.

Outro destaque é em relação ao efeito da inclusão das restrições monotônicas no método de estimação frequentista. Podemos notar que o impacto positivo (em termos de melhoria da precisão) da inclusão adequada das restrições monotônicas é muito menor do que o impacto negativo da inclusão inadequada dessas restrições. Assim, é aconselhável empregar as restrições monotônicas com precaução.

O método de estimação Bayesiano demonstra ser confiável, uma vez que a precisão deste método é comparável ao método de estimação frequentista. Além disso, espera-se que o método de estimativa proposto tenha um desempenho melhor que o método MMLE em amostras pequenas.

Resumo da proposta de aplicação:

- Como exemplo de aplicação, é apresentado um novo método de classificação da depressão para respostas do teste *Beck Depression Inventory* (BDI) derivado do método apresentado por Silva et al. (2018).
- Nesse novo método, o modelo G-DINA é utilizado no ajuste dos dados, no lugar do modelo DINA, utilizado originalmente por Silva et al. (2018).

O teste BDI

- O BDI é um teste psicométrico auto-reportado de múltipla escolha com 21 itens.
- Cada item do teste avalia um sintoma diferente da depressão.
- Cada item tem quatro possíveis respostas que são pontuadas de 0 a 3.
- O método tradicional de diagnóstico é baseado na soma dos pontos de todos os itens.

- Seguindo Fragoso and Cúri (2013) e Silva et al. (2018), foi considerado duas dimensões para a depressão: *cognitiva* e *somático-afetiva*.
- O vetor de atributos é um vetor bidimensional (α_1, α_2) onde α_1 é a indicadora de sintomas na dimensão cognitiva; e α_2 é a indicadora de sintomas na dimensão somático-afetiva.
- Existem quatro possíveis diagnósticos:
 - $\alpha_{(1)} = (0, 0)$: não depressivo;
 - $\alpha_{(2)} = (1, 0)$: sintomático na dimensão cognitiva;
 - $\alpha_{(3)} = (0, 1)$: sintomático na dimensão somático-afetiva;
 - $\alpha_{(4)} = (1, 1)$: sintomático nas duas dimensões.

Figure: Matriz Q da aplicação

Item	Q (dimensões)	
	α_1	α_2
1. Tristeza	1	1
2. Pessimismo	1	1
3. Senso de fracasso	1	0
4. Falta de prazer	1	1
5. Sentimentos de culpa	1	0
6. Senso de punição	1	0
7. Auto desgosto	1	0
8. Auto acusação	1	0
9. Desejos suicidas	1	1
10. Crises de choro	1	1
11. Irritabilidade	0	1
12. Falta de sociabilidade	1	1
13. Indecisão	1	1
14. Distorção da imagem própria	1	1
15. Inibição para o trabalho	1	1
16. Distúrbio do sono	0	1
17. Fatigabilidade	0	1
18. Perda de apetite	1	1
19. Perda de peso	0	1
20. Preocupação com a saúde	1	1
21. Perda de libido	0	1

- As respostas foram dicotomizadas de modo que $Y = 0$ representa as respostas do teste com pontuação 0 e $Y = 1$, as respostas com pontuação 1, 2 ou 3.
- Em nosso estudo, foi utilizado um conjunto de dados referente às respostas do teste BDI de 1111 estudantes universitários.
- Foi realizado uma análise comparativa entre a metodologia que utiliza o G-DINA e a metodologia que utiliza o DINA.
- Ambos modelos foram ajustados utilizando algoritmos MCMC Gibbs Sampling implementados com o software JAGS.
- Foram geradas 2 cadeias de Markov com 10000 iterações cada, sendo as primeiras 5000 descartadas.

Seleção de Modelo:

Table: Comparação de ajuste dos modelos DINA e G-DINA (LPC) para os dados do BDI.

model	DIC	WAIC	RMSEA médio
DINA	25538.6	34116.6	0.0365
G-DINA	24684.4	33280.3	0.0257

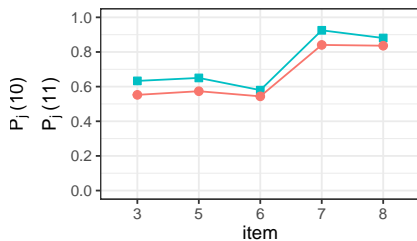
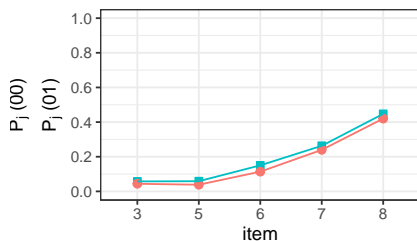
DIC: Deviance Information Criterion (Spiegelhalter et al., 2002)

WAIC: Widely Applicable Information Criterion (**vehtari2017practical**)

RMSEA: Root Mean Square Error of Approximation (Lei and Li, 2016)

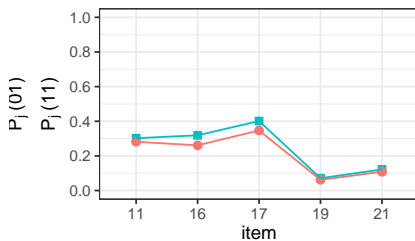
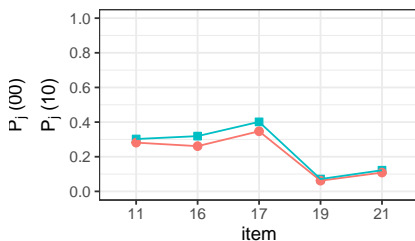
$$RMSEA_j = \sqrt{2 \sum_{c=1}^C \pi_c \left(\hat{P}(\alpha) - R_{jc}/I_c \right)^2}$$

Figure: Probabilidade de sucesso estimada dos itens com $\mathbf{q}_j = (1, 0)$



Model —●— DINA —■— G-DINA

Figure: Probabilidade de sucesso estimada dos itens com $\mathbf{q}_j = (0, 1)$



Model —●— DINA —■— G-DINA

Figure: Probabilidade de sucesso estimada dos itens com $\mathbf{q}_j = (1, 1)$

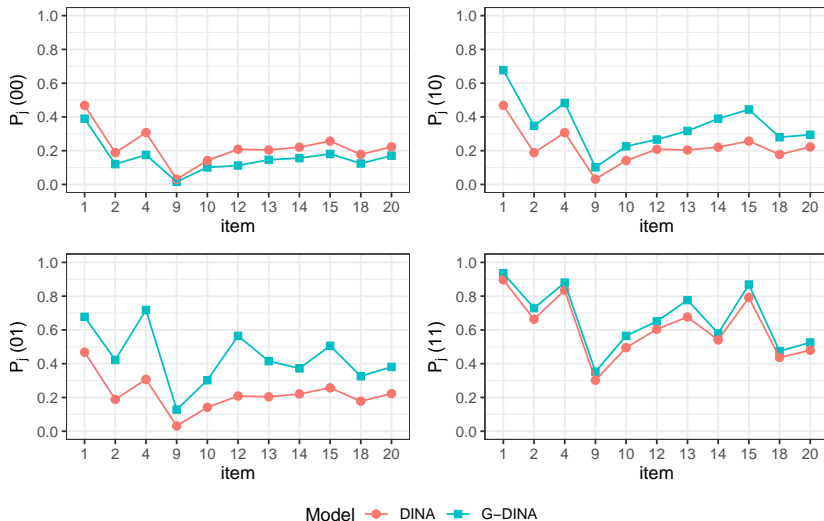


Figure: Proporção dos perfis latentes estimada com o modelo DINA e G-DINA

classe	Descrição	$\hat{\pi}_c$ (DINA)		$\hat{\pi}_c$ (G-DINA)	
		Média	DP	Média	DP
1	Não depressivo	0.360	0.023	0.377	0.021
2	Sintomático na dimensão cognitiva	0.124	0.016	0.173	0.019
3	Sintomático na dimensão somático-afetiva	0.124	0.021	0.190	0.021
4	Sintomático nas duas dimensões	0.392	0.019	0.260	0.018






Figure: Tabela de contingência dos vetores de atributos estimados





		G-DINA				Total
		$\hat{\alpha}_{(1)}$	$\hat{\alpha}_{(2)}$	$\hat{\alpha}_{(3)}$	$\hat{\alpha}_{(4)}$	
DINA	$\hat{\alpha}_{(1)}$	378	11	46	0	435
	$\hat{\alpha}_{(2)}$	16	109	1	0	126
	$\hat{\alpha}_{(3)}$	37	4	65	0	106
	$\hat{\alpha}_{(4)}$	0	63	92	289	444
Total		431	187	204	289	1111





- $\hat{\alpha}_{(1)}$: não depressivo;
- $\hat{\alpha}_{(2)}$: sintomático na dimensão cognitiva;
- $\hat{\alpha}_{(3)}$: sintomático na dimensão somático-afetiva;
- $\hat{\alpha}_{(4)}$: sintomático nas duas dimensões.

- O modelo DINA tende a superestimar a proporção de sintomáticos nas duas dimensões.
- Essa tendência já havia sido notada por Silva et al. (2018), que atribuíram tal comportamento à dicotomização das respostas.
- Nossos resultados mostram que essa tendência de superestimação também está relacionada à escolha do modelo.

- Os estudos de simulação mostram que o método de estimação Bayesiano proposto recupera todos os parâmetros com boa acurácia.
- O custo computacional da estimação Bayesiana é consideravelmente maior do que o do método de estimação frequentista. Porém, o método Bayesiano fornece mais informações sobre os parâmetros por meio da distribuição a posteriori.
- Os resultados do estudo de simulação indica que é mais seguro impor restrições menos restritivas, como o LPC e o HPC, do que o MC, devido ao impacto negativo na precisão da estimativa de assumir erroneamente condições mais rigorosas.
- Com relação à metodologia de classificação da depressão apresentada, é importante destacar que nosso foco foi na análise estatística. Para a real aplicação dessa metodologia, investigações mais aprofundadas, envolvendo especialistas em psiquiatria, se faz necessário.

-  Chen, Huilin and Jinsong Chen (2016). “Exploring reading comprehension skill relationships through the G-DINA model”. In: *Educational Psychology* 36.6, pp. 1049–1064.
-  Chen, Yi-Hsin et al. (2019). “Examining psychometric properties and level classification of the van Hiele Geometry Test using CTT and CDM frameworks”. In: *Journal of Educational Measurement* 56.4, pp. 733–756.
-  de la Torre, Jimmy (2011). “The generalized DINA model framework”. In: *Psychometrika* 76.2, pp. 179–199.
-  de la Torre, Jimmy, L Andries van der Ark, and Gina Rossi (2018). “Analysis of clinical data from a cognitive diagnosis modeling framework”. In: *Measurement and Evaluation in Counseling and Development* 51.4, pp. 281–296.
-  Fernandes, Renato S, Jorge L Bazán, and Mariana Cúri (2022). “A Bayesian approach for the G-DINA model”. In.

-  Fragoso, Tiago M and Mariana Cúri (2013). “Improving psychometric assessment of the beck depression inventory using multidimensional item response theory”. In: *Biometrical Journal* 55.4, pp. 527–540.
-  George, Ann Cathrice and Alexander Robitzsch (2015). “Cognitive diagnosis models in R: A didactic”. In: *The Quantitative Methods for Psychology* 11.3, pp. 189–205.
-  Lei, Pui-Wa and Hongli Li (2016). “Performance of fit indices in choosing correct cognitive diagnostic models and Q-matrices”. In: *Applied Psychological Measurement* 40.6, pp. 405–417.
-  Li, Hongli, C Vincent Hunter, and Pui-Wa Lei (2016). “The selection of cognitive diagnostic models for a reading comprehension test”. In: *Language Testing* 33.3, pp. 391–409.

-  Silva, Marcelo A. da et al. (2018). “Estimating the DINA model parameters using the No-U-Turn Sampler”. In: *Biometrical Journal*. DOI: [10.1002/bimj.201600225](https://doi.org/10.1002/bimj.201600225).
-  Spiegelhalter, David J et al. (2002). “Bayesian measures of model complexity and fit”. In: *Journal of the royal statistical society: Series b (statistical methodology)* 64.4, pp. 583–639.
-  Templin, Jonathan L and Robert A Henson (2006). “Measurement of psychological disorders using cognitive diagnosis models.”. In: *Psychological methods* 11.3, p. 287.
-  Yamaguchi, Kazuhiro and Kensuke Okada (2018). “Comparison among cognitive diagnostic models for the TIMSS 2007 fourth grade mathematics assessment”. In: *PloS one* 13.2, e0188691.