

# Uma introdução aos Modelos de Diagnóstico Cognitivo utilizando o software Stan

Eduardo Schneider Bueno de Oliveira  
Marcelo Andrade da Silva  
Jorge Luis Bazán Guzmán

PIPGes - ICMC USP / UFSCar  
*eduardosb@usp.br*



da Silva, M. A., de Oliveira, E. S., Davier, A. A., Bazán, J. L. (2018).  
Estimating the DINA model parameters using the No-U-Turn  
Sampler. *Biometrical Journal*, 60(2), 352-368.

# Sumário

- 1 Introdução
- 2 Principais modelos
  - Modelos DINA e DINO
    - Estimação com enfoque na abordagem Bayesiana
  - Modelo G-DINA
- 3 Stan
- 4 Aplicação
- 5 Discussão e Conclusão
- 6 Referências

# 1. Introdução

# Motivação

- As avaliações são importantes em educação, psicologia e análise biométrica;
- Através de questionários, é possível obter diversas informações, tanto para ranquear quanto para entender forças e fraquezas dos indivíduos;
- Existem diferentes abordagens estatísticas para a análise de resultados de instrumentos de avaliação;
- Dentro desses contextos, os modelos de variável latente tem crescido bastante nas últimas décadas.

# Introdução

- Instrumentos de medidas baseados em modelos da TRI informam as pontuações em um único valor, independente das habilidades avaliadas (ENEM);
- Estas pontuações são úteis para estabelecer a ordem relativa, mas não avaliam os pontos fortes e fracos específicos dos examinados;
- Após o uso da TRI pode ser feito um segundo estágio de análise para determinar critérios de diagnóstico;
- Métodos que fornecem diagnóstico em uma única etapa são de grande valia.

# Introdução

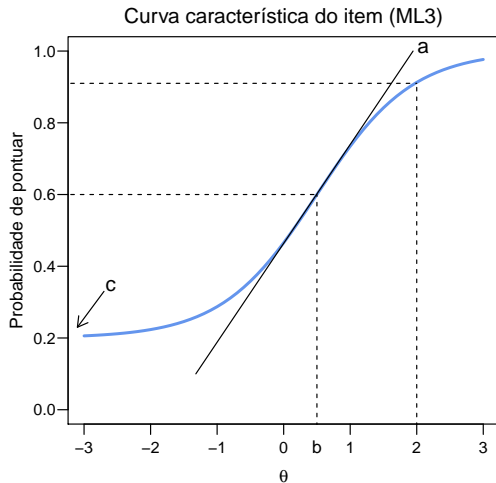
- Uma nova abordagem foi proposta, fornecendo inferências relevantes para estabelecer diagnósticos;
- Estes modelos são conhecidos na literatura por *Cognitive Diagnosis Models* (CDM), Modelos de Diagnóstico Cognitivo, em tradução livre;
- CDM são modelos de variáveis latentes discretas;
- Eles são desenvolvidos especificamente para diagnosticar a presença ou ausência de habilidades envolvidas na avaliação para cada indivíduo;
- Uma boa revisão pode ser vista em George & Robitzsch (2015).

# Introdução - TRI vs CDM

- Diferença fundamental entre TRI e CDMs:
  - TRI: desempenho do indivíduo é baseado em um traço latente contínuo  $\theta$ ;
  - TRI: Indivíduo com traço latente maior tem maior probabilidade de responder corretamente um item;



# Introdução - TRI vs CDM

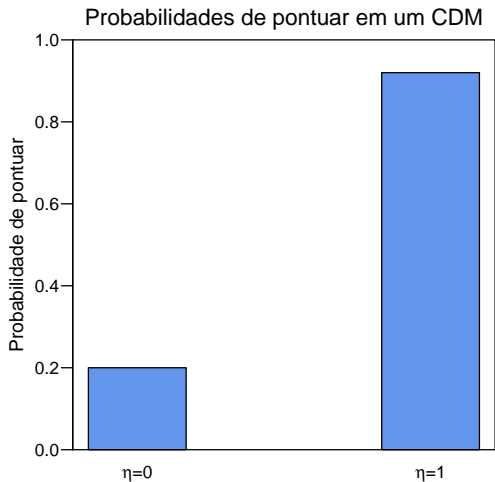


# Introdução - TRI vs CDM

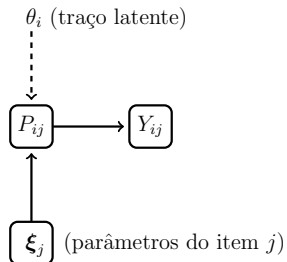
Diferença fundamental entre TRI e CDM:

- TRI: desempenho do indivíduo é baseado em um traço latente contínuo  $\theta$ ;
- Indivíduo com traço latente maior tem maior probabilidade de responder corretamente um item;
- CDM: desempenho do indivíduo baseado em traços latentes discretos;
- Indivíduo é bem sucedido em um item, em sua grande maioria, se ele possuir os traços latentes necessários para responder adequadamente ao que esse item exige.

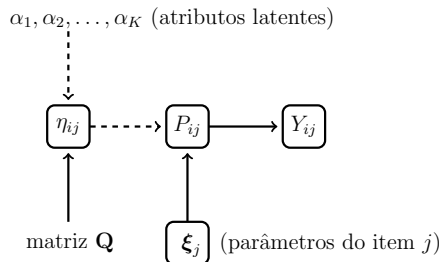
# Introdução - TRI vs CDM



# Construção da TRI e dos CDM



(a) TRI



(b) CDM

# Introdução

A proposta deste minicurso é explorar brevemente três dos modelos pertencentes aos CDM denominados DINA (Deterministic Input, Noisy “And” Gate), DINO (Deterministic Input, Noisy “Or” Gate) e G-Dina (Generalized DINA).

- Contribuir para difusão dos CDM;
- Apresentar a estimação bayesiana e divulgar o código do modelo DINA no Stan;
- Aplicar o modelo em dados reais.

## 2. Principais modelos

# Definições

- É importante definir algumas quantidades para os modelos, as quais serão iguais para o DINA e o DINO;
- Esses modelos são aplicados quando queremos avaliar habilidades ou características que um indivíduo possui, com base em respostas à questionários;
- Consideramos  $i$  como o índice para indivíduos ( $i = 1, \dots, N$ ),  $j$  para itens dicotômicos ( $j = 1, \dots, J$ ) e  $k$  para as habilidades ( $k = 1, \dots, K$ ).

- $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{iJ})'$  é o vetor de respostas do indivíduo  $i$ . Esses valores são observados.

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

- $\boldsymbol{\alpha}_i = (\alpha_{i1}, \dots, \alpha_{iK})'$  é o vetor de habilidades do indivíduo  $i$ . Esse vetor é latente.

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

- $\mathbf{q}'_j = (q_{j1}, \dots, q_{jK})$  é o vetor de habilidades avaliadas pelo item  $j$ , o qual forma a Matriz  $\mathbf{Q} = (\mathbf{q}_1, \dots, \mathbf{q}_J)'$ . Esses valores são determinados anteriormente, não sendo latentes nem observados.

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



# Matriz $Q$

- A Matriz  $Q$  é um passo de extrema importância para o correto funcionamento do modelo;
- Para sua construção, é possível utilizar diversos métodos (Henson e Templin, 2009):
  - Realização de *brainstorming* com especialistas da área;
  - Utilização de outras técnicas estatísticas como a TRI;
  - Construção probabilística pelo modelo DINA (de la Torre, 2008)

# Modelo DINA

- Dos vários formatos de teste, o de múltipla escolha (MC) tem sido amplamente utilizado;
- Dentro dos testes de MC há os que envolvem pontuações dicotomizadas (correto/incorreto);
- O modelo DINA para itens dicotômicos é o mais usualmente utilizado, pela sua facilidade de aplicação e interpretação.

# Modelo DINA

$$Y_{ij} | \alpha_i, \Omega_j \sim \text{Bernoulli}(P_{ij}),$$

em que  $\Omega_j = (g_j, s_j)$  é o vetor de parâmetros do item  $j$ ,  $g_j$  é a probabilidade de acertar o item  $j$  adivinhando (*guessing*) e  $s_j$  é a probabilidade de escorregar (*slipping*). Ou seja,

$$g_j = P(Y_{ij} = 1 | \eta_{ij} = 0) \text{ e } s_j = P(Y_{ij} = 0 | \eta_{ij} = 1)$$

Além disso,

$$P_{ij} = P(Y_{ij} = 1 | \alpha_i, \Omega_j) = (1 - s_j)^{\eta_{ij}} g_j^{1 - \eta_{ij}},$$

em que  $\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}} = \mathbb{1}(\alpha_i' \mathbf{q}_j = \mathbf{q}_j' \mathbf{q}_j)$ .

# Modelo DINO

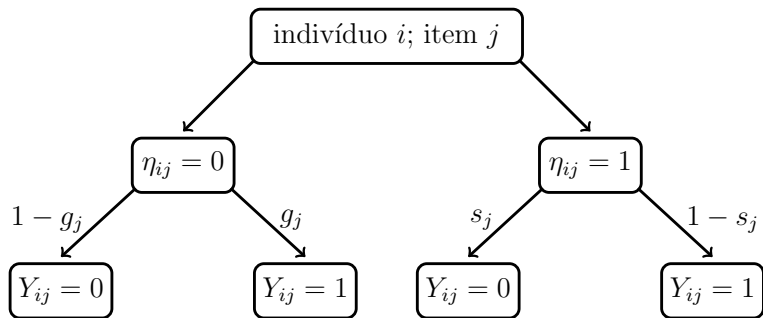
A única diferença para o modelo DINO será para  $\eta$ . Assim teremos:

$$P_{ij} = P(Y_{ij} = 1 | \alpha_i, \Omega_j) = (1 - s_j)^{\eta_{ij}} g_j^{1 - \eta_{ij}},$$

em que

$$\eta_{ij} = \begin{cases} 1, & \text{se o indivíduo } i \text{ possui pelo menos um dos atributos} \\ & \text{exigidos pelo item } j \\ 0, & \text{caso contrário} \end{cases}$$

# Modelos DINA e DINO



# Modelo DINA

Considerando as respostas dos itens independentes e

$\Omega = (\Omega_1, \Omega_2, \dots, \Omega_J)$ , temos

$$P(\mathbf{Y}_i | \alpha_i = \alpha_c, \Omega) = \prod_{j=1}^J P_{cj}^{Y_{ij}} (1 - P_{cj})^{1 - Y_{ij}},$$

em que  $\alpha_i = \alpha_c$  significa que o indivíduo  $i$  pertence à classe de atributos  $c$  e  $P_{cj} = (1 - s_j)^{\eta_{cj}} g_j^{1 - \eta_{cj}}$ .

# Modelo DINA

- Seja  $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots, \pi_C)'$  um vetor de dimensão  $C$ , com  $\pi_c = P(\boldsymbol{\alpha}_i = \boldsymbol{\alpha}_c | \boldsymbol{\pi})$  sendo a probabilidade marginal de um indivíduo pertencer à classe  $c$ ;
- Obviamente,  $0 \leq \pi_c \leq 1$  e  $\sum_{c=1}^C \pi_c = 1$ ;
- A probabilidade de observar  $\mathbf{Y}_i$  condicionado aos parâmetros dos itens  $\boldsymbol{\Omega}$  e ao vetor de probabilidades  $\boldsymbol{\pi}$  é dado pela marginalização dos  $\boldsymbol{\alpha}_i$

$$\begin{aligned} P(\mathbf{Y}_i | \boldsymbol{\Omega}, \boldsymbol{\pi}) &= \sum_{c=1}^C \pi_c P(\mathbf{Y}_i | \boldsymbol{\alpha}_i = \boldsymbol{\alpha}_c, \boldsymbol{\Omega}) \\ &= \sum_{c=1}^C \pi_c \prod_{j=1}^J P(Y_{ij} | \boldsymbol{\alpha}_i = \boldsymbol{\alpha}_c, \boldsymbol{\Omega}). \end{aligned}$$

# Exemplo

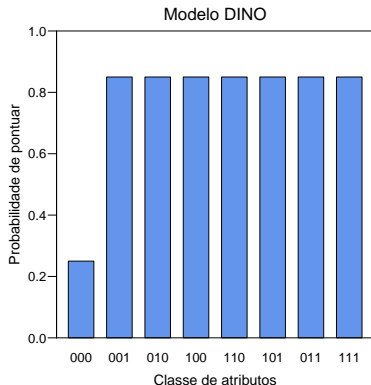
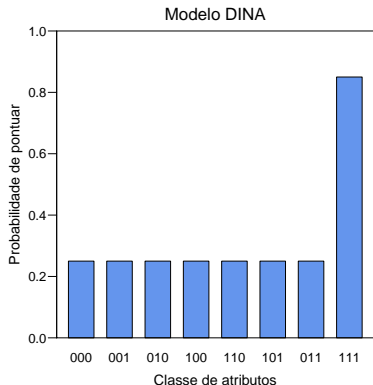
Para melhor entender o funcionamento do modelo imagine a seguinte situação:

- Foi aplicada uma prova de matemática, onde são avaliadas  $K = 3$  habilidades:
  - Habilidade 1: Álgebra;
  - Habilidade 2: Trigonometria;
  - Habilidade 3: Geometria.
- Com isso, haverão  $2^3 = 8$  possíveis perfis de habilidades  $\alpha$ ;
- Cada indivíduo  $i$ , com perfil  $\alpha_i$  terá uma probabilidade diferente de responder corretamente ao item  $j$ , que avalia as habilidades (uma, duas ou as três), definidas em  $\mathbf{q}_j$ .



## DINA vs DINO

- As probabilidades de acerto de um item, em geral, serão bastante maiores para o modelo DINO do que para o DINA.



## 2.1. Estimação dos modelos DINA e DINO

# Estimação

A estimação dos modelos DINA e DINO pode ser feita de diversas maneiras:

- Abordagem Clássica:
  - George & Robitzsch (2015);
  - Uso do algoritmo EM (pacote “CDM” no R).
- Abordagem Bayesiana:
  - Gibbs Sampling (pacote “dina” no R), em Culpepper (2015);
  - Metropolis Hastings (OpenBUGS), código feito para nosso artigo;
  - Hamiltonian Monte Carlo, mais especificamente, No-U-Turn Sampler (Stan), código feito para nosso artigo e apresentado no minicurso.

# Estimação

A função de verossimilhança da observação de uma amostra de  $N$  indivíduos que respondem  $J$  itens é dada por

$$\begin{aligned}\ell(\mathbf{Y}|\Omega, \boldsymbol{\pi}) &= \prod_{i=1}^N \sum_{c=1}^C \pi_c P(\mathbf{Y}_i | \boldsymbol{\alpha}_i = \boldsymbol{\alpha}_c, \Omega) \\ &= \prod_{i=1}^N \sum_{c=1}^C \pi_c \prod_{j=1}^J P(Y_{ij} | \boldsymbol{\alpha}_i = \boldsymbol{\alpha}_c, \Omega).\end{aligned}$$

# Estimação

A estimação Bayesiana dos parâmetros é baseada na distribuição à posteriori

$$p(\Omega, \pi | \mathbf{Y}) \propto \ell(\mathbf{Y} | \Omega, \pi) p(\Omega) p(\pi),$$

em que  $p(\Omega)$  é a priori para os parâmetros dos itens e  $p(\pi)$  é a priori para os perfis de atributos.

# Estimação

A formulação Bayesiana do modelo DINA presente em Culpepper (2015) fornece as seguintes priors

$$\alpha_j | \boldsymbol{\pi} \sim \text{Categórica}(\pi_1, \pi_2, \dots, \pi_C), \quad (1)$$

$$\boldsymbol{\pi} \sim \text{Dirichlet}(\boldsymbol{\delta}_0), \boldsymbol{\delta}_0 = (\delta_{01}, \delta_{02}, \dots, \delta_{0C}), \quad (2)$$

$$s_j \sim \text{Beta}(a_s, b_s), \quad (3)$$

$$g_j \sim \text{Beta}(a_g, b_g). \quad (4)$$

- A escolha do vetor de parâmetros  $\boldsymbol{\delta}_0$  em (2) depende do conhecimento prévio da prevalência das classes de atributos;
- Se não houver essa informação prévia sobre  $\boldsymbol{\pi}$ , uma priori não informativa pode ser obtida tomando  $\delta_{0c} = 1$  para todo  $c$ .

# Estimação

- Note que as prioris (3) e (4) tornam-se não informativas tomando  $a_s = b_s = a_g = b_g = 1$ ;
- Culpepper (2015) explicitamente insere a restrição de monotonicidade nos parâmetros dos itens da seguinte forma

$$p(\Omega_j) = p(s_j, g_j) \propto s_j^{a_s-1} (1-s_j)^{b_s-1} g_j^{a_g-1} (1-g_j)^{b_g-1} \mathbb{1}((s_j, g_j) \in \mathcal{P}), \quad (5)$$

em que  $\mathcal{P} = \{(s, g) : 0 \leq s + g < 1, 0 \leq s < 1, 0 \leq g < 1\}$

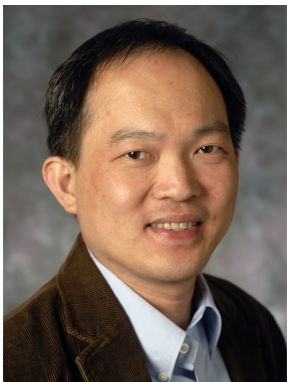
# Estimação

- Para o modelo DINO as mesmas prioris podem ser utilizadas;
- A forma da verossimilhança também será escrita da mesma maneira;
- É importante lembrar que a definição dos  $\eta_{ij}$  é diferente para o DINO.



# Modelo G-DINA

de la Torre, J. (2011) The generalized DINA model framework.  
Psychometrika 76 (2) 179-199.



Jimmy de la Torre, atualmente Professor na Universidade de Hong Kong.

## Modelo G-DINA: motivação

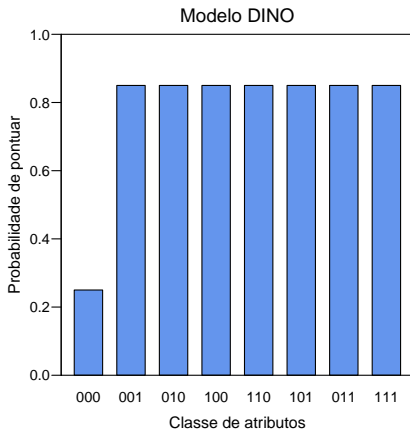
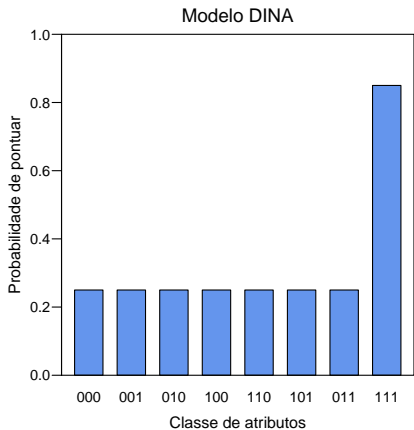
- Nos modelos DINA e DINO, fixado um item  $j$ , as classe de atributos  $\alpha_c = (\alpha_{c1}, \alpha_{c2}, \dots, \alpha_{cK})$  são divididas em dois grupos: o grupo em que  $\eta_{cj} = 0$  e o grupo em que  $\eta_{cj} = 1$ ;
- Indivíduos com classes de atributos em um mesmo grupo tem a mesma probabilidade de responder corretamente o item  $j$ ;

## Modelo G-DINA: motivação

Exemplo: Supondo um item  $j$  com  $\mathbf{q}'_j = (1, 1, 1)$ ,

- a partir do modelo DINA, um indivíduo  $i$  com classe de atributo  $\alpha_i \neq (1, 1, 1)'$ , terá mesma probabilidade de responder corretamente o item.
- a partir do modelo DINO, um indivíduo  $i$  com classe de atributo  $\alpha_i \neq (0, 0, 0)'$ , terá mesma probabilidade de responder corretamente o item.

# Modelo G-DINA: motivação



# Modelo G-DINA

- O modelo G-DINA foi proposto para lidar com esta limitação;
- A matriz  $\mathbf{Q}$  continua sendo necessária;
- O número de atributos requeridos pelo item  $j$  é  $K_j^* = \sum_{k=1}^K q_{jk}$
- Por simplicidade, definimos  $\alpha_{ij}^* = (\alpha_{ij1}, \alpha_{ij2}, \dots, \alpha_{ijK_j^*})'$  contendo apenas os atributos requeridos para o item  $j$ ;
- Exemplo:

$$\mathbf{q}'_j = (1, 0, 1)$$

$$\begin{array}{ccc} & \downarrow & \downarrow \\ \alpha_{ij}^* & = & (\alpha_{ij1}, \alpha_{ij2})' \end{array}$$

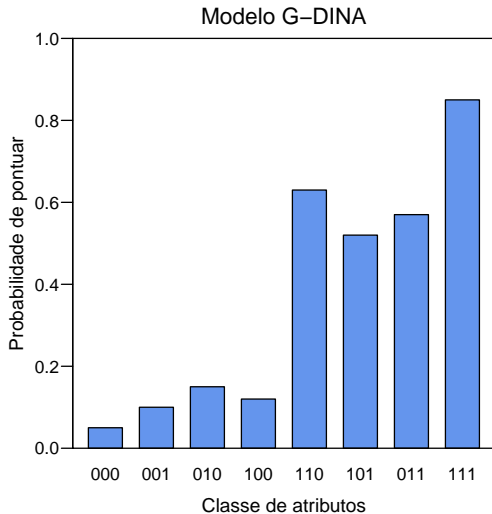
# Modelo G-DINA

- A probabilidade do indivíduo  $i$  responder corretamente o item  $j$  de acordo com o modelo G-DINA é

$$P_{ij} = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{ijk} + \sum_{k' > k}^{K_j^*} \sum_{k=1}^{K_j^* - 1} \delta_{jkk'} \alpha_{ijk} \alpha_{ijk'} + \dots + \delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{ijk}$$

- $\delta_{j0}$  é o intercepto para o item  $j$ ;
- $\delta_{jk}$  é o efeito principal devido ao  $\alpha_{ijk}$ ;
- $\delta_{jkk'}$  é o efeito da interação entre  $\alpha_{ijk}$  e  $\alpha_{ijk'}$ ;
- $\delta_{j12\dots K_j^*}$  é o efeito da interação entre  $\alpha_{ij1}, \alpha_{ij2}, \dots, \alpha_{ijK_j^*}$ .

# Modelo G-DINA





# Modelo G-DINA

- O modelo DINA pode ser obtido a partir do modelo G-DINA definindo todos os parâmetros, exceto  $\delta_{j0}$  e  $\delta_{j12\dots K_j^*}$ , iguais a zero. Assim, temos

$$P_{ij} = \delta_{j0} + \delta_{j12\dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{ijk}, \quad (6)$$

em que  $g_j = \delta_{j0}$  e  $1 - s_j = \delta_{j0} + \delta_{j12\dots K_j^*}$ .

# Modelo G-DINA

- O modelo DINO também pode ser obtido a partir do modelo G-DINA definindo a seguinte restrição nos parâmetros

$$\delta_{jk} = -\delta_{jk'k''} = \dots = (-1)^{K_j^*+1} \delta_{j12\dots K_j^*},$$

- Assim, os valores, em módulo, dos efeitos são idênticos, porém os sinais são alternados de acordo com a ordem de interação.
- Então, temos

$$P_{ij} = \begin{cases} \delta_{j0}, & \text{se } \alpha_{ij}^* = \mathbf{0}_{K_j^*} \\ \delta_{j0} + \delta_{jk}, & \text{caso contrário} \end{cases}$$

- Em termos do G-DINA,  $g_j = \delta_{0j}$  e  $1 - s_j = \delta_{j0} + \delta_{jk}$ .

## 3. Stan

# Stan

- Stan é uma plataforma para modelagem estatística lançada em 2012;
- É uma alternativa frente a outros softwares conhecidos: BUGS e JAGS;
- Alguns estudos comparativos mostram que o Stan é uma opção eficiente para a estimação dos parâmetros de modelos latentes.

# Stan

- Os usuários do Stan podem escolher entre dois eficientes métodos MCMC: HMC e NUTS;
- O Stan funciona a partir da interação com alguns softwares de análise de dados: Python, R, Stata, Julia, entre outros;
- Especificamente para o software R, a interação com o Stan é feita através do pacote “RStan”.

# Stan

- Carpenter, B.; Gelman, A.; Hoffman, M.; Lee, D.; Goodrich, B.; Betancourt, M.; Brubaker, M.; Guo, J.; Li, P. and Riddell, A. (2017) Stan: A probabilistic programming language. *Journal of Statistical Software* 76 (1), 1–32.
- Stan Development Team. (2017). Stan Modeling Language Users Guide and Reference Manual, Version 2.17.0.
- <http://mc-stan.org>

# Utilizando o Stan no software R

Orientações para a instalação do pacote “RStan”:

- <https://github.com/stan-dev/rstan/wiki/RStan-Getting-Started>

## Utilizando o Stan no software R

A função “stan” no R:

```
1 library(rstan)
2 stanfit <- stan(file = <arquivo com código Stan>,
3               pars = <vetor de parâmetros de
4                   retorno>,
5               algorithm = <``NUTS`` ou ``HMC``>,
6               chains = <número de cadeias>,
7               thin = <intervalo entre observações>,
8               iter = <número total de iterações>,
9               warmup = <número de iterações
10                      iniciais descartadas>)
```



# A estrutura da linguagem Stan

A construção de um modelo no Stan é feita em blocos.

- functions;
- data;
- transformed data;
- parameters;
- transformed parameters;
- model;
- generated quantities.

# A estrutura da linguagem Stan

O bloco “data” para o modelo DINA:

```
1 data {  
2   int<lower=1> N;  
3   int<lower=1> J;  
4   int<lower=1> K;  
5   int<lower=1> C;  
6   int<lower=0, upper=1> Y[N, J];  
7   int<lower=0, upper=1> Q[J, K];  
8   int<lower=0, upper=1> As[C, K];  
9 }
```

# A estrutura da linguagem Stan

O bloco “transformed data” para o modelo DINA:

```
1 transformed data{  
2   vector[C] delta;  
3   delta = rep_vector(1, C);  
4 }
```

# A estrutura da linguagem Stan

O bloco “parameters” para o modelo DINA:

```
1 parameters {  
2   simplex[C] pi;  
3   simplex[3] Omega[J];  
4 }
```

```
1 transformed parameters {
2   vector<lower=0, upper=1>[J] prob[C];
3   vector<lower=0, upper=1>[J] eta[C];
4   vector<lower=0, upper=1>[J] g;
5   vector<lower=0, upper=1>[J] s;
6   vector[C] lp[N];
7   g = to_vector(Omega[, 1]);
8   s = to_vector(Omega[, 2]);
9   for (c in 1:C) {
10    for (j in 1:J) {
11      eta[c, j] = 1;
12      for (k in 1:K)
13        eta[c, j] = eta[c, j] * As[c, k] ^ Q[j, k];
14      prob[c, j] = (1-s[j])^eta[c, j]*(g[j])^(1-eta[c,
15        j]);
16    }
17  }
```

## A estrutura da linguagem Stan

O bloco “transformed parameters” para o modelo DINA:

```
17 //continuação
18 for (i in 1:N) {
19     for (c in 1:C) {
20         lp[i, c] = log(pi[c]);
21         for (j in 1:J)
22             lp[i, c] += bernoulli_lpmf(Y_sim[i, j] | prob
23                                     [c, j]);
24     }
25 }
```

## A estrutura da linguagem Stan

$$P(\mathbf{Y}_i | \Omega, \pi) = \sum_{c=1}^C \pi_c \prod_{j=1}^J P(Y_{ij} | \alpha_i = \alpha_c, \Omega)$$

$$\begin{aligned} \log P(\mathbf{Y}_i | \Omega, \pi) &= \log \sum_{c=1}^C \pi_c \prod_{j=1}^J P(Y_{ij} | \alpha_i = \alpha_c, \Omega) \\ &= \log \sum_{c=1}^C \exp \left\{ \log \left( \pi_c \prod_{j=1}^J P(Y_{ij} | \alpha_i = \alpha_c, \Omega) \right) \right\} \\ &= \log \sum_{c=1}^C \exp \left\{ \log(\pi_c) + \sum_{j=1}^J \log \left( P(Y_{ij} | \alpha_i = \alpha_c, \Omega) \right) \right\} \end{aligned}$$

A função `log_sum_exp()` fornece estabilidade aritmética e eficiência para os cálculos, segundo o Manual do Stan.

# A estrutura da linguagem Stan

O bloco “model” para o modelo DINA:

```
1 model {
2   for (j in 1:J) {
3     Omega[j, 1] ~ beta(1, 1);
4     Omega[j, 2] ~ beta(1, 1);
5   }
6   pi ~ dirichlet(delta);
7   for (i in 1:N)
8     target += log_sum_exp(lp[i, 1:C]);
9 }
```



## A estrutura da linguagem Stan

O bloco “generated quantities” para o modelo DINA:

```

1 generated quantities {
2   int<lower=1, upper=C> alpha[N];
3   vector<lower=0, upper=1>[J] u;
4   for (i in 1:N)
5     alpha[i] = categorical_rng(softmax(lp[i, 1:C]));
6   u = 1-s;
7 }
```

em que  $\text{softmax}(lp_i) = \left( \frac{\exp(lp_{i,1})}{\sum_{c=1}^C \exp(lp_{i,c})}, \dots, \frac{\exp(lp_{i,C})}{\sum_{c=1}^C \exp(lp_{i,c})} \right)$ .

# Stan

## A saída da função “stan” no R:

```

1 > stanfit
2 Inference for Stan model: CodeDINamodel.
3 1 chains, each with iter=4000; warmup=2000; thin=1;
4 post-warmup draws per chain=2000, total post-warmup draws=2000.
5
6           mean    se_mean    sd   2.5%   25%   50%   75%   97.5%   n_eff   Rhat
7 g[1]      0.47      0.00   0.02   0.43   0.46   0.47   0.48    0.51   2000    1
8 g[2]      0.19      0.00   0.02   0.16   0.18   0.19   0.20    0.22   2000    1
9 g[3]      0.05      0.00   0.01   0.02   0.04   0.04   0.05    0.07   2000    1
10 g[4]      0.31      0.00   0.02   0.27   0.29   0.31   0.32    0.35   2000    1
11 g[5]      0.04      0.00   0.01   0.02   0.03   0.04   0.05    0.06   2000    1
12 g[6]      0.11      0.00   0.02   0.08   0.10   0.11   0.13    0.15   2000    1
13 g[7]      0.24      0.00   0.02   0.20   0.23   0.24   0.26    0.29   2000    1
14 g[8]      0.42      0.00   0.02   0.38   0.41   0.42   0.44    0.47   2000    1
15 g[9]      0.03      0.00   0.01   0.02   0.03   0.03   0.04    0.05   2000    1
16 g[10]     0.14      0.00   0.01   0.12   0.13   0.14   0.15    0.17   2000    1
17 [ reached getOption("max.print") — omitted 1169 rows ]
18
19 Samples were drawn using NUTS(diag_e) at Wed May 09 16:41:48 2018.
20 For each parameter, n_eff is a crude measure of effective sample size,
21 and Rhat is the potential scale reduction factor on split chains (at
22 convergence, Rhat=1).
```

# Stan

- Na prática, o usuário escreve o modelo Bayesiano em linguagem Stan, que “traduz” para C++, executa e retorna uma amostra MCMC.

# Stan

Alguns benefícios de utilizar o Stan:

- fácil implementação de modelos bayesianos;
- interação com softwares populares de análise de dados;
- uso de métodos MCMC eficientes;
- fácil obtenção de critérios de comparação de modelos.

## 4. Aplicação - Estudo sobre depressão usando o Modelo DINA

# Dados

- Utilizamos dados de respostas ao *Beck Depression Inventory* (BDI; Beck et al. 1961);
- Ele é um questionário com 21 perguntas de múltipla escolha, com respostas de 0 a 3, sobre questões relacionadas à depressão;
- Foi respondido por 1111 estudantes, aplicado pelo Dr. Teng Chei-Tung, do Hospital das Clínicas, Universidade de São Paulo;
- A abordagem tradicional (Kendall et. al, 1987) apenas soma os valores do questionário e classifica os indivíduos.

# Dados

Na abordagem tradicional temos:

- Não depressivo (Score total do BDI entre 0 e 15);
- Disfórico (Score total do BDI entre 16 e 20);
- Depressivo (Score total do BDI entre 21 e 63).

# Dados

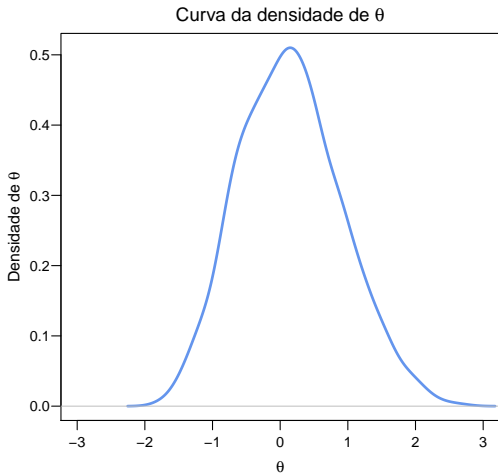
- Para nosso estudo os dados foram dicotomizados, como em Fragoso e Cúri, 2013, anteriormente;
- Os indivíduos que respondem 0 continuam com 0 (não sintomáticos) e os que respondem 1, 2 ou 3 recebem 1 (sintomáticos);
- Para a aplicação, usamos o modelo DINA, mostrando a possibilidade de obtenção de diferentes informações em relação à TRI.



# Análise com a TRI

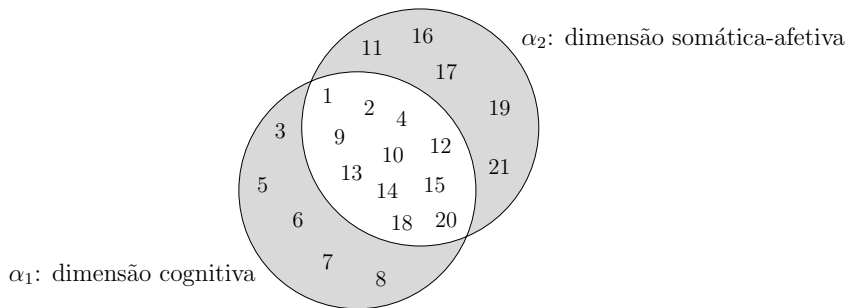
Item	a	b	c
BDI1	1.636	-0.244	0.156
BDI2	1.628	0.559	0.038
BDI3	2.136	0.650	0.000
BDI4	1.652	-0.044	0.000
BDI5	1.814	0.664	0.000
BDI6	1.185	0.734	0.000
BDI7	1.853	-0.159	0.000
BDI8	1.211	-0.584	0.000
BDI9	1.629	1.589	0.000
BDI10	1.158	1.024	0.000
BDI11	1.032	-0.032	0.000
BDI12	1.189	0.603	0.000
BDI13	1.442	0.434	0.000
BDI14	0.893	0.830	0.000
BDI15	1.621	0.126	0.000
BDI16	0.849	0.039	0.000
BDI17	1.086	-0.459	0.000
BDI18	0.806	1.336	0.000
BDI19	1.729	2.877	0.083
BDI20	0.742	1.115	0.000
BDI21	0.709	1.844	0.000

# Análise com a TRI



## Matriz $Q$

- Um passo importante para usar o modelo DINA é especificar a Matriz  $Q$ ;
- Para isso, usamos a divisão da depressão em duas dimensões, proposta por Fragoso e Cúri, 2013, que utilizaram TRI para tal:



## Estimativa dos parâmetros

- Uma vez especificada a matriz  $Q$ , ajustamos o modelo DINA, utilizando o NUTS para obter estimativas para os parâmetros;
- Foram avaliados os parâmetros de itens e também o dos indivíduos;
- Além disso, podemos fazer a comparação dos resultados obtidos com a abordagem tradicional.

# Estimativa dos parâmetros de itens

Item	Q (dimensões)		$\hat{g}$		$\hat{s}$	
	$\alpha_1$	$\alpha_2$	Média	dp	Média	dp
1. Tristeza	1	1	0.468	0.020	0.101	0.017
2. Pessimismo	1	1	0.189	0.017	0.334	0.026
3. Senso de fracasso	1	0	0.045	0.012	0.444	0.024
4. Falta de prazer	1	1	0.309	0.020	0.163	0.020
5. Sentimentos de culpa	1	0	0.039	0.011	0.423	0.024
6. Senso de punição	1	0	0.115	0.017	0.453	0.023
7. Auto desgosto	1	0	0.242	0.022	0.158	0.019
8. Auto acusação	1	0	0.422	0.023	0.163	0.017
9. Desejos suicidas	1	1	0.032	0.008	0.694	0.023
10. Crises de choro	1	1	0.142	0.014	0.502	0.026
11. Irritabilidade	0	1	0.283	0.024	0.279	0.023
12. Falta de sociabilidade	1	1	0.210	0.017	0.394	0.025
13. Indecisão	1	1	0.205	0.016	0.320	0.025
14. Distorção da imagem própria	1	1	0.222	0.017	0.458	0.025
15. Inibição para o trabalho	1	1	0.259	0.018	0.206	0.023
16. Distúrbio do sono	0	1	0.262	0.028	0.288	0.022
17. Fatigabilidade	0	1	0.348	0.030	0.162	0.019
18. Perda de apetite	1	1	0.178	0.016	0.560	0.026
19. Perda de peso	0	1	0.062	0.012	0.851	0.016
20. Preocupação com a saúde	1	1	0.223	0.016	0.518	0.026
21. Perda de libido	0	1	0.109	0.017	0.645	0.022

# Estimativa dos parâmetros de indivíduos

Diagnóstico de depressão usando as probabilidades a posteriori dos respondentes do BDI ( $n = 1111$ )

c		Dimensões		$\hat{\pi}$	
		$\alpha_1$	$\alpha_2$	Média	dp
1	(não depressivo)	0	0	0.363	0.024
2	(sintomático na dimensão cognitiva)	1	0	0.124	0.016
3	(sintomático na dimensão somático-afetiva)	0	1	0.124	0.021
4	(ambos os sintomas)	1	1	0.389	0.019

# Comparação das abordagens

Distribuição de  $\alpha$  por grupos do BDI.

Diagnóstico proposto pelo modelo DINA	Grupos de acordo com o proposto por Kendall		
	Depressivo	Disfórico	Não depressivo
Não depressivo	0(0%)	0(0%)	442(51.64%)
Sintomático na dimensão cognitiva	0(0%)	5(4.39%)	116(13.55%)
Sintomático na dimensão somático-afetiva	0(0%)	0(0%)	106(12.38%)
Ambos os sintomas	141(100%)	109(95.61%)	192(22.43%)

## 5. Discussão e Conclusão



## Discussão e Conclusão

- Os modelos CDM tem grande valia em contextos educacionais e também psicométricos;
- O avanço computacional permite o amplo uso dessa e de outras técnicas, com diferentes pacotes;
- Vantagens de utilizar o Stan: simples para programar, método eficiente, vários meios de suporte (fórum, manual, etc);
- Ainda há bastante espaço para o desenvolvimento de técnicas nos CDM e também para a aplicação a diferentes áreas do saber;
- Para trabalhos práticos, é importante a presença de equipes interdisciplinares.

## Importância dos CDM para a Biometria

- Metodologias para classificar usando traços latentes, contínuos ou discretos, podem ser utilizadas na agricultura, na saúde e muitas áreas de interesse na Biometria;
- A Matriz  $Q$  tem grande importância. Conhecimento prévio por especialistas faz-se de suma importância;
- Com este minicurso desejamos atrair a mais pesquisadores e usuários para desenvolver mais trabalhos na área.

## Grupo de modelos de variáveis latentes

Grupo de modelos de variáveis latentes com encontros regulares na USP, São Carlos-SP. Alguns temas já abordados:

- Modelos de Diagnóstico Cognitivo;
- Modelos de TRI para dados politômicos;
- Modelos de TRI com testlet;
- Fraude em testes.

## 6. Referências



BECK, A. T.; WARD, C. H.; MENDELSON, M.; MOCK, J.; ERBAUGH, J. (1961)

An inventory for measuring.  
*Archives of general psychiatry* 4, 561 – 571.



CULPEPPER, S. A. (2015)

Bayesian estimation of the DINA model with gibbs sampling.  
*Journal of Educational and Behavioral Statistics* 40 (5), 454 – 476.



FRAGOSO, T. M.; CÚRI, M. (2013)

Improving psychometric assessment of the beck depression inventory using multidimensional item response theory.  
*Biometrical Journal* 55, 527 – 540.



GEORGE, A. C.; ROBITZSCH, A. (2015)

Cognitive diagnosis models in R: a didactic.  
*The quantitative methods for psychology* 11(3), 189 – 205.



HENSON, R. A.; TEMPLIN, J. L. (2009)

Q-matrix construction.  
Paper presented at 2009 NCME training session. San Francisco, CA.



KENDALL, P. C.; HOLLON, S. D.; BECK, A. T.; HAMMEN, C. L.; INGRAM, R. E. (1987)

Issues and recommendations regarding use of the Beck Depression Inventory.  
*Cognitive therapy and research* 11 (3), 289 – 299.



de la TORRE, J.; DOUGLAS, J. A. (2004)

Higher-order latent trait models for cognitive diagnosis.  
*Journal of Educational and Behavioral Statistics* 69, 333 – 353.



de la TORRE, J.; DOUGLAS, J. A. (2008)

An Empirically Based Method of Q-Matrix Validation for the DINA Model: Development and Applications.  
*Journal of educational measurement* 45 (4), 342 – 362.

Link deste material:

<http://conteudo.icmc.usp.br/pessoas/jlbazan/presentations.html>

Obrigado!

Perguntas?