

Teorias de Avaliação - CE095

Adilson dos Anjos¹

¹Departamento de Estatística
Universidade Federal do Paraná
aanjos@ufpr.br

Curitiba, PR
28 de agosto de 2014

Teoria da Resposta ao Item

Introdução

–Motivação–

Teoria da Resposta ao Item

	item 1	item 2	item 3	item 4	item 5	item 6	item 7	item 8	item 9	item 10
	■	■	■	■	■	■	■	■	■	■
A	×	●	×	●	×	●	×	●	×	●
B	×	●	×	●	×	●	×	●	×	●

Teoria Clássica dos Testes

 $A \Rightarrow 5 \text{ acertos}$ $B \Rightarrow 5 \text{ acertos}$

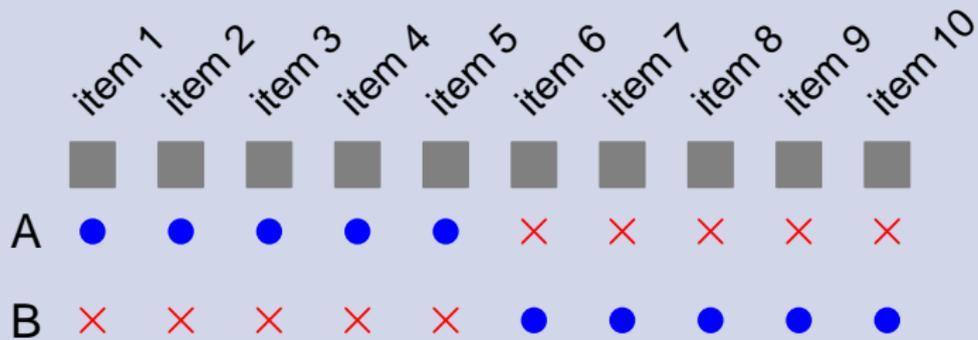
Resultado

score A = score B

Conclusão

 $A = B$

Teoria da Resposta ao Item



Teoria Clássica dos Testes

 $A \Rightarrow 5 \text{ acertos}$ $B \Rightarrow 5 \text{ acertos}$

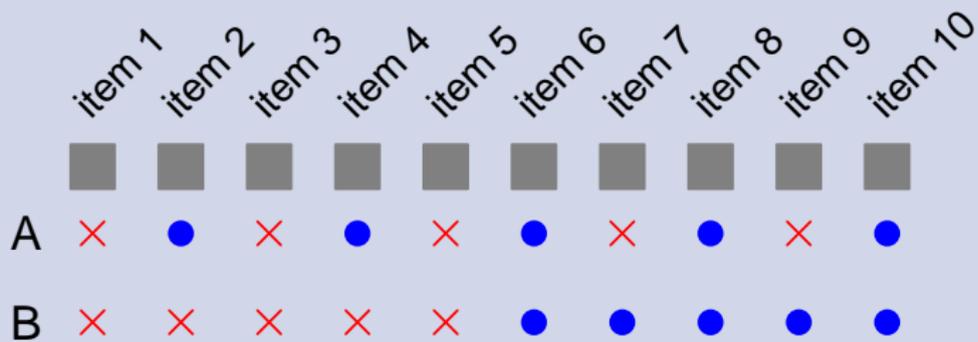
Resultado

score A = score B

Conclusão

 $A = B$

Teoria da Resposta ao Item



Teoria Clássica dos Testes

 $A \Rightarrow 5 \text{ acertos}$ $B \Rightarrow 5 \text{ acertos}$

Resultado

score A = score B

Conclusão

 $A = B$

Teoria da Resposta ao Item

Fácil

Difícil

item 1

item 2

item 3

item 4

item 5

item 6

item 7

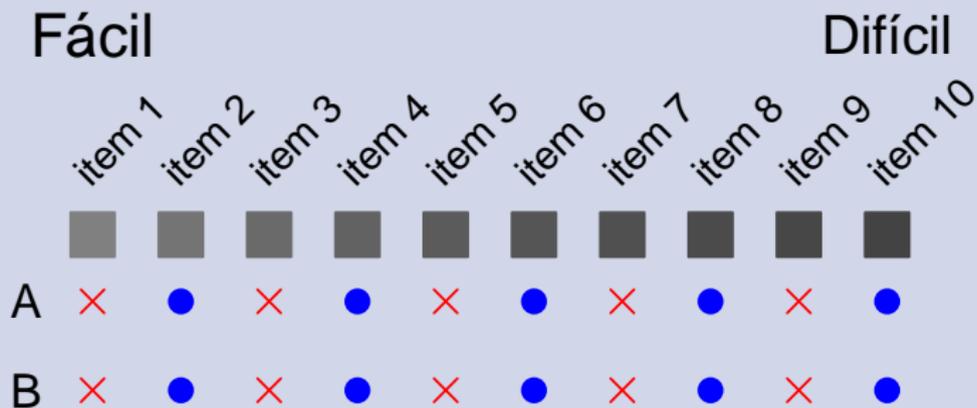
item 8

item 9

item 10



Teoria da Resposta ao Item



Teoria Clássica dos Testes

 $A \Rightarrow 5 \text{ acertos}$ $B \Rightarrow 5 \text{ acertos}$

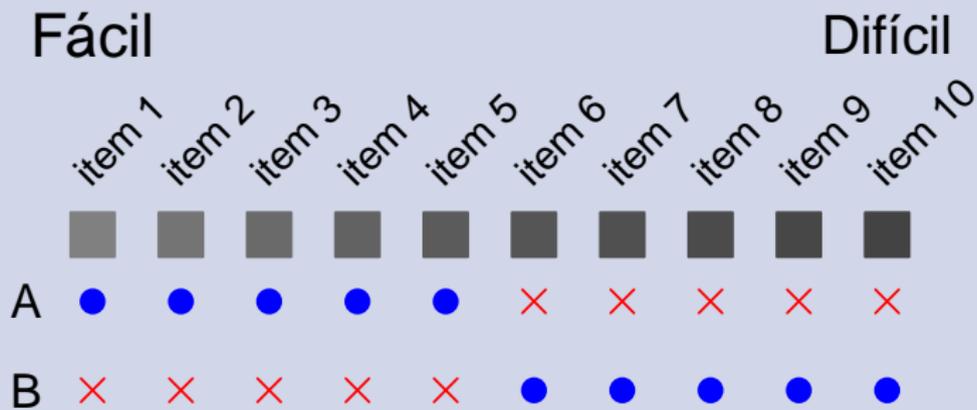
Resultado

score A = score B

Conclusão

 $A = B$

Teoria da Resposta ao Item



Teoria Clássica dos Testes

 $A \Rightarrow 5 \text{ acertos}$ $B \Rightarrow 5 \text{ acertos}$

Resultado

score A = score B ??

Conclusão

 $A = B ??$

Existe metodologia 'mais adequada'?

TEORIA DA RESPOSTA AO ITEM TRI

Teoria da Resposta ao Item –Introdução–

Introdução

- 1 Primeiro modelo, Lord, 1952: unidimensional de 2 parâmetros;
- 2 Considera a **Probabilidade** de resposta (correta) a um item;
- 3 'Escore TRI': chamado de *proficiência* ou *habilidade* (θ);
- 4 Foco no Item e não na Prova;
- 5 Item: possui parâmetros;
- 6 No Brasil: utilizada em 1995 no SAEB.

Modelos

- Dependem do tipo de item:
 - dicotômico (dicotomizado);
 - politômico (ordinal).
- Dependem do traço latente:
 - cumulativo;
 - não cumulativo (desdobramento).

Modelo cumulativo

Quanto maior a proficiência do indivíduo maior é a probabilidade de responder corretamente o item.

Modelos

A probabilidade de resposta (correta) a um item é modelada em função da proficiência do indivíduo e dos parâmetros dos itens.

Introdução

- População e grupos
 - grupo: uma amostra de indivíduos de uma população;
 - população depende do interesse do estudo:

Uma população poderia ser composta por 'alunos do primeiro ano de cursos superiores diurnos e noturnos em ciências exatas da UFPR'.

Uma amostra de alunos desses cursos poderia ser considerada um *grupo*.

Vantagens

- Permite comparação entre populações (com itens em comum);
- Permite comparação entre indivíduos da mesma população (provas diferentes);
- Evolução do desempenho de populações entre anos (estudos longitudinais);

Vantagens

Os parâmetros dos itens e as proficiências dos indivíduos são **INVARIANTES** (exceto pela escala escolhida (métrica)).

Desvantagens?

- Estimaco de parmetros:
 - Vrios mtodos de estimaco;
 - Vrios softwares;
- Tamanho de amostra: padres de resposta (importante para calibrao);
- Necessrio respondentes posicionados ao longo de toda a escala de habilidade.
- **Apenas uma amostra grande no garante uma boa estimaco dos parmetros dos itens;**

Estimação de parâmetros

A estimação de parâmetros do itens é chamada de **Calibração**.

Estimação de parâmetros

Vários métodos:

- máxima verossimilhança conjunta - muitos parâmetros;
- máxima verossimilhança marginal - menor custo computacional;
- métodos bayesianos;

Modelos para itens dicotômicos

- Modelos baseados na função logística;
- Rasch, 1, 2 e 3 parâmetros;
- 4 parâmetros (d): indivíduos com proficiência elevada não tem uma probabilidade alta de responder afirmativamente um item: por exemplo, em um instrumento para medir depressão, respondentes com alta probabilidade à depressão podem não responder 'sim' a chance de suicídio.
- ou ainda, d=parâmetro de 'descuido';

Escala

- O 'escore TRI' possui uma interpretação (pedagógica);
- Construída juntamente com especialistas;
- Escala: conjunto dos números reais - uma 'régua' (ver site do INEP/ENEM);

O que é Medida de Traço latente?

- É o processo por meio do qual os números estão atribuídos aos objetos (itens e respondentes), tais que a relação entre objetos é representada pela relação entre números.
- Itens têm diferentes níveis do traço latente.

Medindo altura com questionário

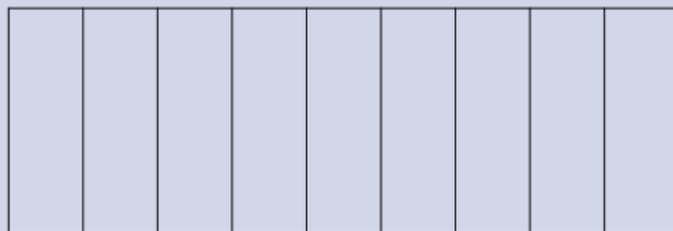
Tabela 1 : Questionário com itens para estimar a altura de pessoas.

Item	Descrição (pergunta): Assinale 1 para 'sim' e 0 para 'não'.
1	Na cama, eu frequentemente sinto frio nos pés.
2	Eu frequentemente desço as escadas de dois em dois degraus.
3	Eu acho que me daria bem em um time de basquete.
4	Como policial eu impressionaria muito.
5	Na maioria dos carros eu me sinto desconfortável.
6	Eu literalmente olho para meus colegas de cima para baixo
7	Você é capaz de pegar um objeto no alto de um armário sem usar escada?
8	Você abaixa quando vai passar por uma porta?
9	Você consegue guardar a bagagem no porta-malas do avião?
10	Você regula o banco do carro para trás?
11	Normalmente, quando você está andando de carona, lhe oferecem o banco da frente?
12	Quando você e várias outras pessoas vão tirar fotos, formando-se três fileiras, onde ninguém ficará agachado, você costuma ficar atrás?
13	Você tem dificuldade para se acomodar no ônibus?
14	Em uma fila, por ordem de tamanho, você é sempre colocado atrás?

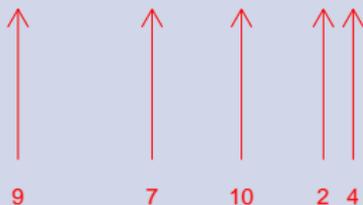
Posicionamento dos indivíduos e dos itens na mesma escala

Medindo traço latente: Escala

Posicionamento dos indivíduos e dos itens na mesma escala

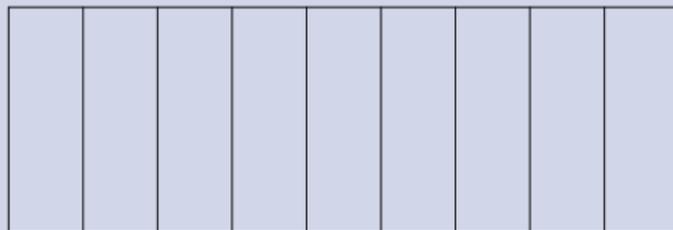


1,50m 1,55m 1,60m 1,65m 1,70m 1,75m 1,80m 1,85m

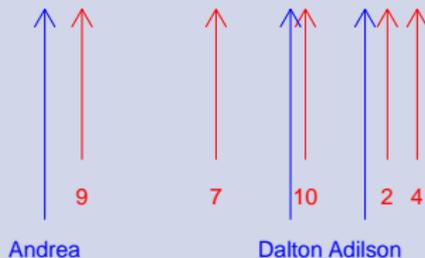


Medindo traço latente: Escala

Posicionamento dos indivíduos e dos itens na mesma escala

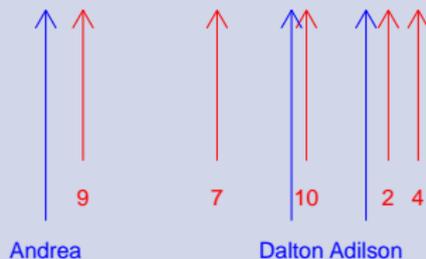
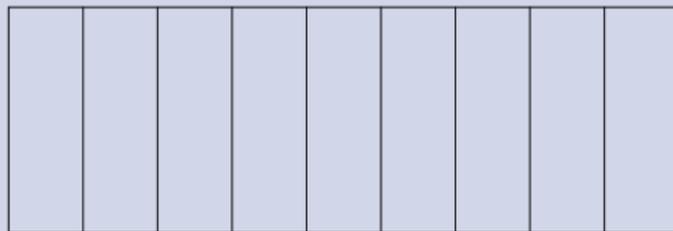


1,50m 1,55m 1,60m 1,65m 1,70m 1,75m 1,80m 1,85m



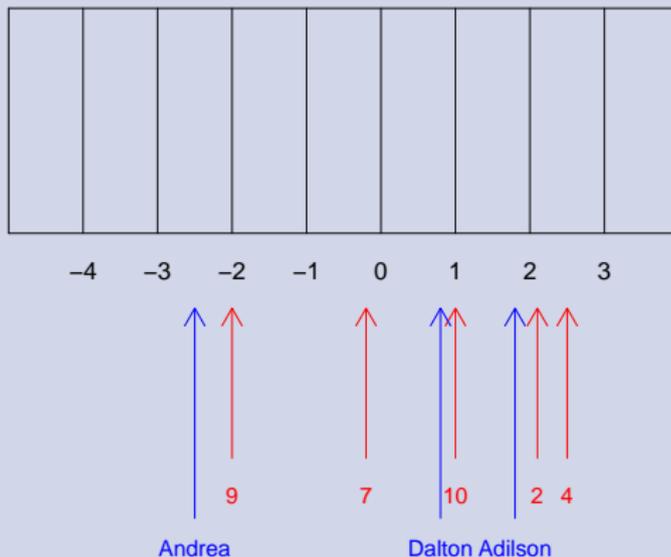
Medindo traço latente: Escala

Posicionamento dos indivíduos e dos itens na mesma escala



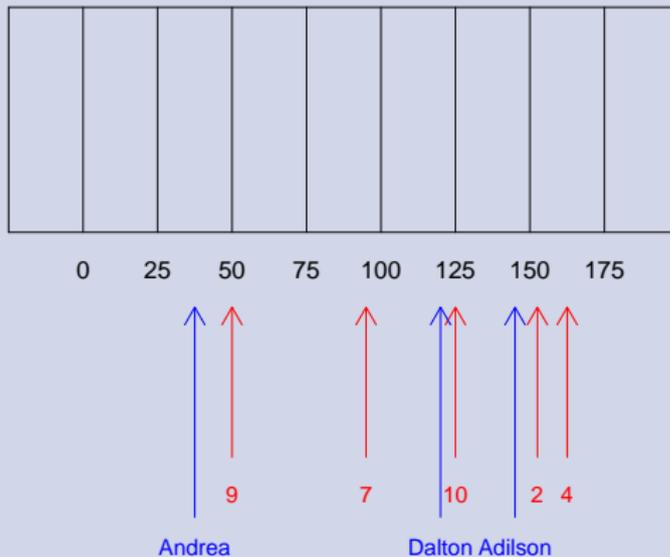
Medindo traço latente: Escala

Posicionamento dos indivíduos e dos itens na mesma escala



Medindo traço latente: Escala

Posicionamento dos indivíduos e dos itens na mesma escala



Comparação TCT vs TRI

Característica	TCT	TRI
Medida	número de acertos	proficiência θ
Escala	número entre 0 e 1	números reais
Escala	depende do teste	independente do teste
Modelo	$X_i = V_i + Erro_i$	probabilístico
Parâmetros dos itens		

Comparação TCT vs TRI

Parâmetros dos itens	TCT	TRI
Dependentes dos respondentes	Sim	Não
Dificuldade vs Escore	escala diferente	mesma escala

Modelos Logísticos para itens dicotômicos ou “dicotomizados”

- 3LP: a , b e c - caso do ENEM (Decreto do MEC)
- 2LP: a e b ($c = 0$) - caso da Altura
- 1LP: b ($c = 0$ e a , o mesmo para todos os itens) - Rasch

Modelo de 3 parâmetros

$$P(U_{ij} = 1 | \theta_j, a_i, b_i, c_i) = c_i + (1 - c_i) \frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}}$$

em que,

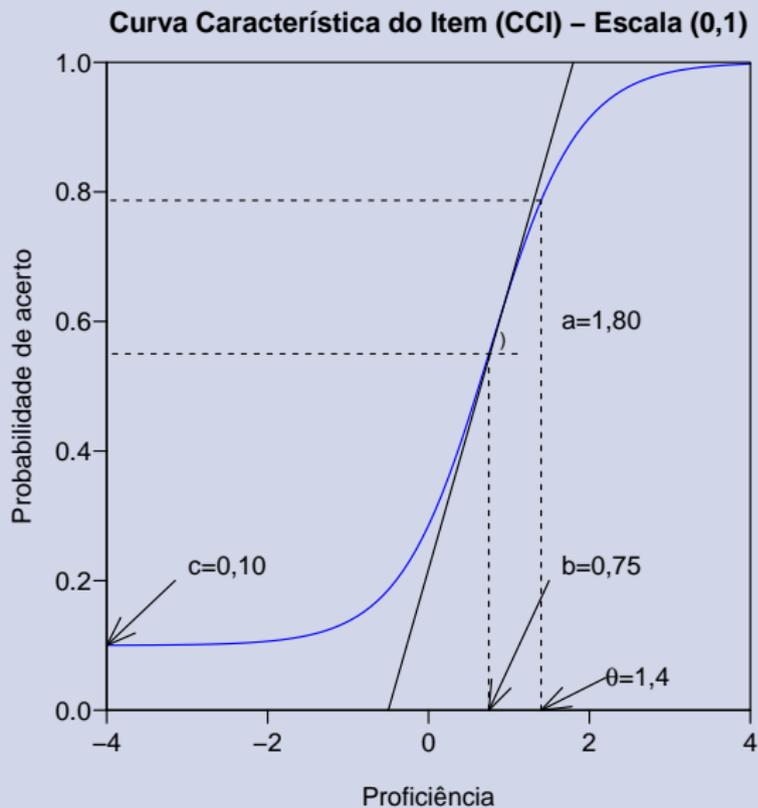
$P(U_{ij} = 1 | \theta_j, a_i, b_i, c_i)$ é a probabilidade do indivíduo j com habilidade θ_j acertar o item i ;

b_i é o parâmetro de dificuldade (ou de posição) do item i , medido na mesma escala de habilidade;

a_i é o parâmetro de discriminação (ou inclinação) do item i , com valor proporcional à inclinação da Curva Característica do Item no ponto b_i ;

c_i é o parâmetro do item que representa a probabilidade de indivíduos com baixa habilidade responderem corretamente o item i (também chamado de probabilidade de acerto casual).

Teoria da Resposta ao Item



Uma função interativa no R:

```
> require(shiny)
> runUrl('http://www.ufpr.br/~aanjos/CE095/01_logis.zip')
```