

UNIVERSIDADE
SÃO FRANCISCO

Teoria de Resposta ao Item: Curva Característica do Item

Dr. Ricardo Primi

Programa de Mestrado e Doutorado em Avaliação Psicológica

Universidade São Francisco

Curva Característica do Item (CCI)

- Idéias importantes
 - Traço latente
 - Métrica (inicialmente arbitrária)
 - Curva descreve a relação entre Probabilidade de acerto e a escala latente
 - Tem duas características principais (localização/dificuldade e inclinação/discriminação) e essas características não se referem à validade do item !! (pg. 8)
 - Noção visual da dificuldade e seu conceito
 - Relação entre discriminação e dificuldade
 - Cada item terá sua curva com suas características particulares
 - A CCI é a "basic building block" da TRI

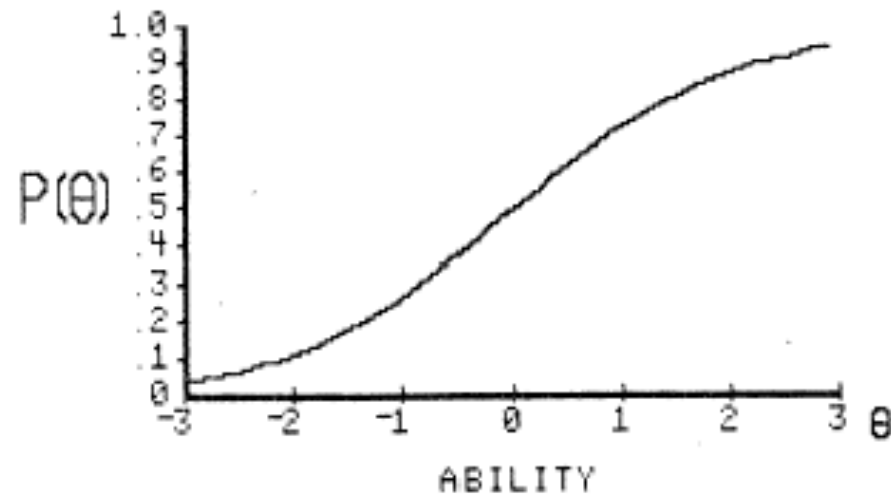
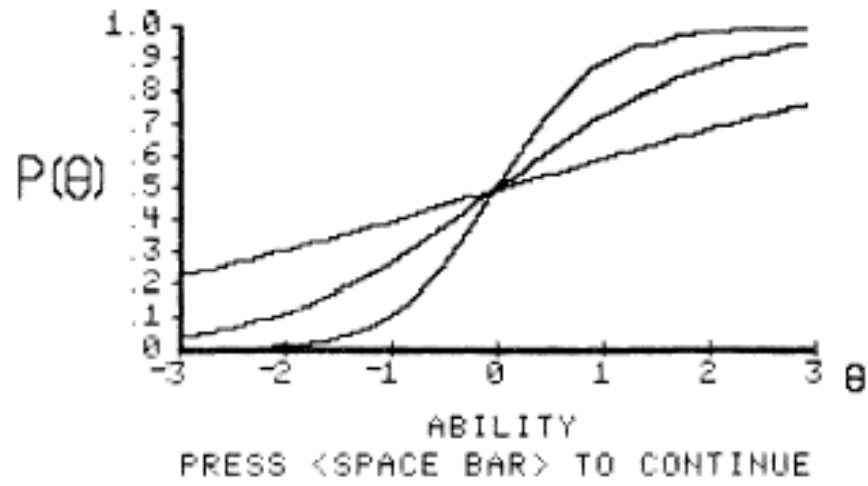
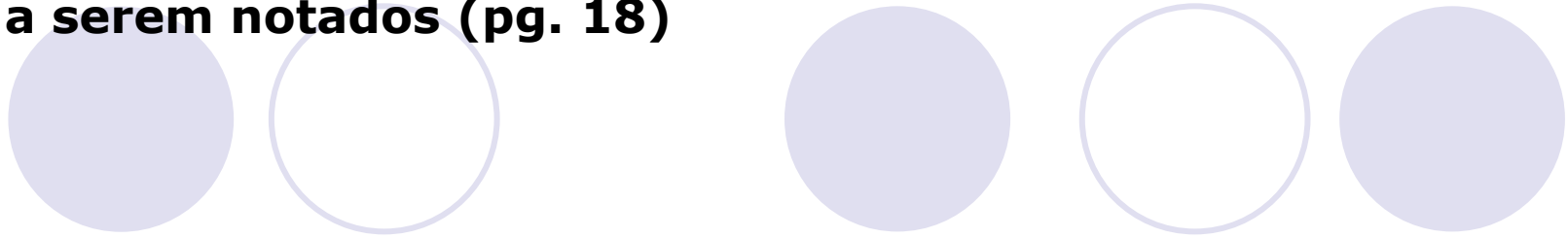


FIGURE 1-1. A typical item characteristic curve

Variações de b 's e a 's nas CCI's



Pontos a serem notados (pg. 18)



- Discriminação baixa > curvas praticamente "flat"
- Discriminação alta > curvas em S e com crescimento abrupto na seção mediana
- Dif. Abaixo da mediana > maior parte da CCI $p > 0,50$
- Dif. acima da mediana > maior parte da CCI $p > 0,50$
- A e B são independentes
- Zero discriminação > linha horizontal em $p = 0,50$ porque não se sabe definir o índice de dificuldade
- $P=0,50$ é igual ao índice de dificuldade

Modelos matemáticos

$$P_i(\theta) = \frac{e^{D(\theta - b_i)}}{1 + e^{D(\theta - b_i)}}$$

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

θ = valor da variável latente

b_i = índice de dificuldade

a_i = índice de discriminação

c_i = probabilidade de acerto ao acaso

e = 2,72

D = 1,7

Antes um pouco de matemática....

$$a^1 = a$$

$$a^x \Rightarrow 2^4 = 2.2.2.2 = 16$$

$$a^0 = 1 = a^{n-n} = (a^n)(a^{-n}) = \frac{a^n}{a^n} = 1$$

$$\frac{a^m}{a^n} = a^{m-n} = \frac{1}{a^{n-m}}$$

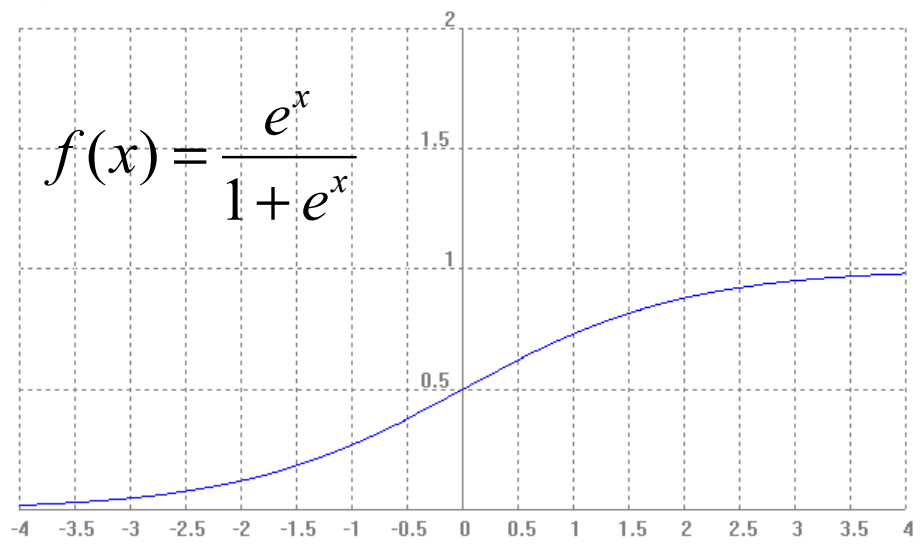
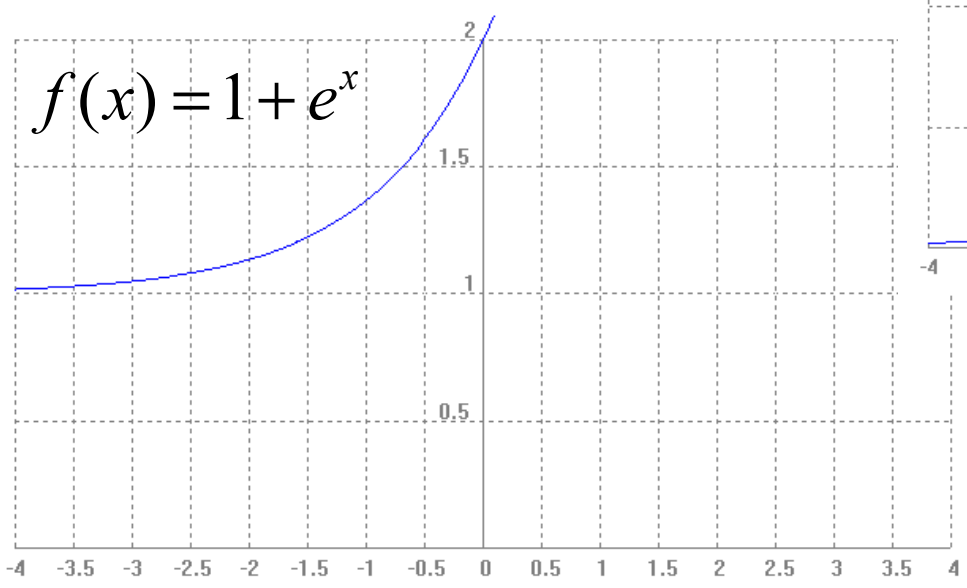
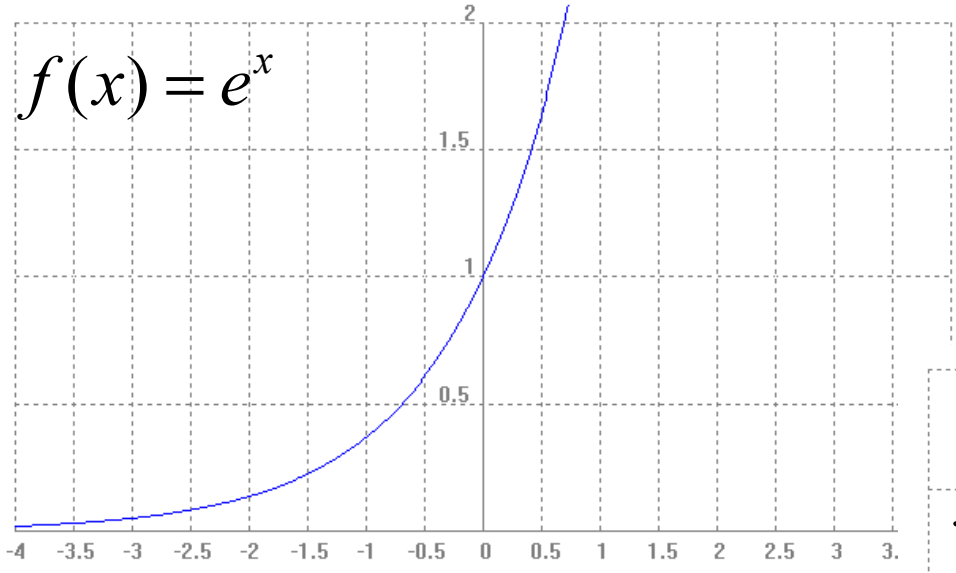
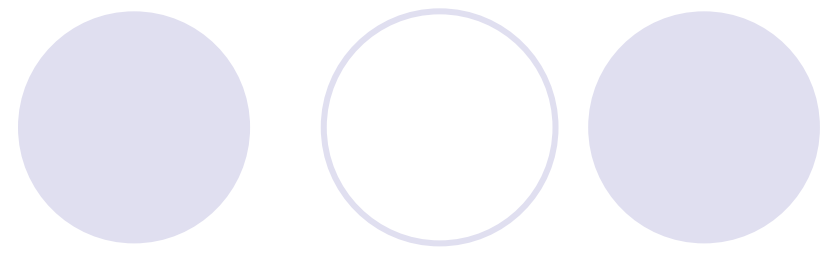
$$a^m \cdot a^n = a^{m+n}$$

$$(a^m)^n = a^{m \cdot n}$$

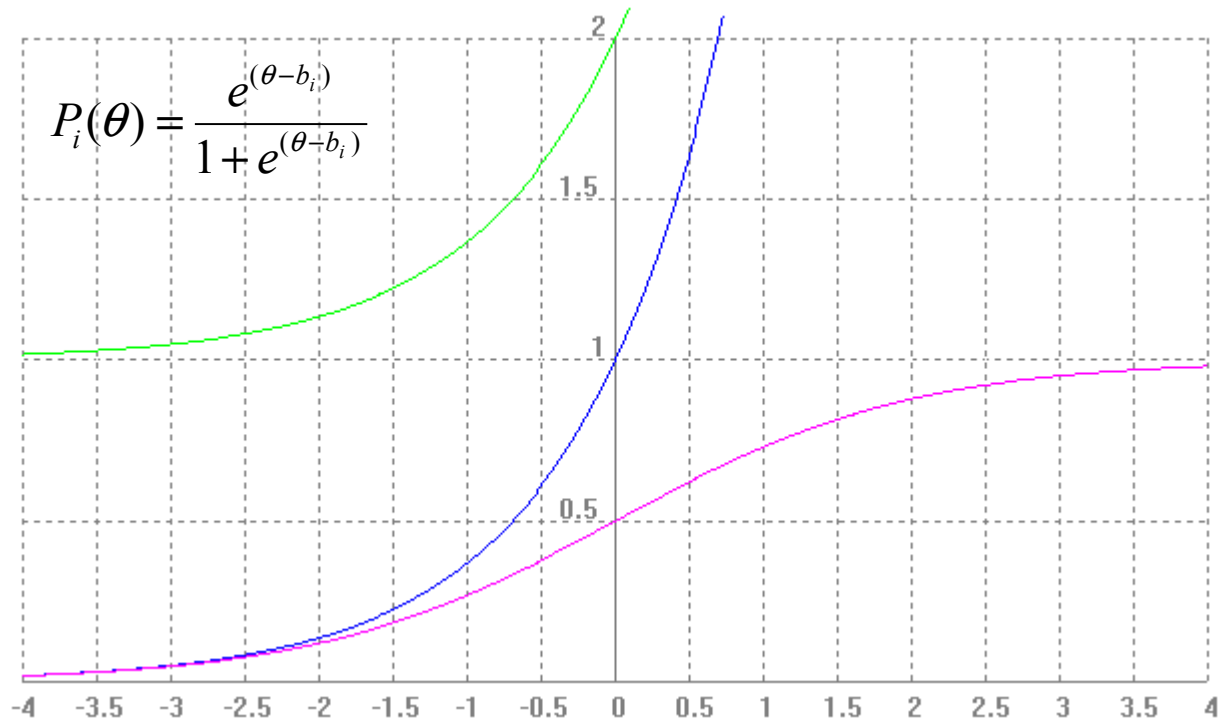
$$a^{-m} = \frac{1}{a^m}$$

$$\begin{aligned} P_i(\theta) &= \frac{1}{1+e^{-L}} = \frac{1}{1+e^{-a(\theta-b_i)}} = \frac{1}{1+\frac{1}{e^{a(\theta-b_i)}}} \\ &= \frac{1}{1+\frac{1}{e^{a(\theta-b_i)}}} \cdot \frac{e^{a(\theta-b_i)}}{e^{a(\theta-b_i)}} = \frac{e^{a(\theta-b_i)}}{e^{a(\theta-b_i)} + \frac{e^{a(\theta-b_i)}}{e^{a(\theta-b_i)}} \cdot \frac{e^{a(\theta-b_i)}}{e^{a(\theta-b_i)} + 1} \\ &= \frac{e^{a(\theta-b_i)}}{1+e^{a(\theta-b_i)}} \end{aligned}$$

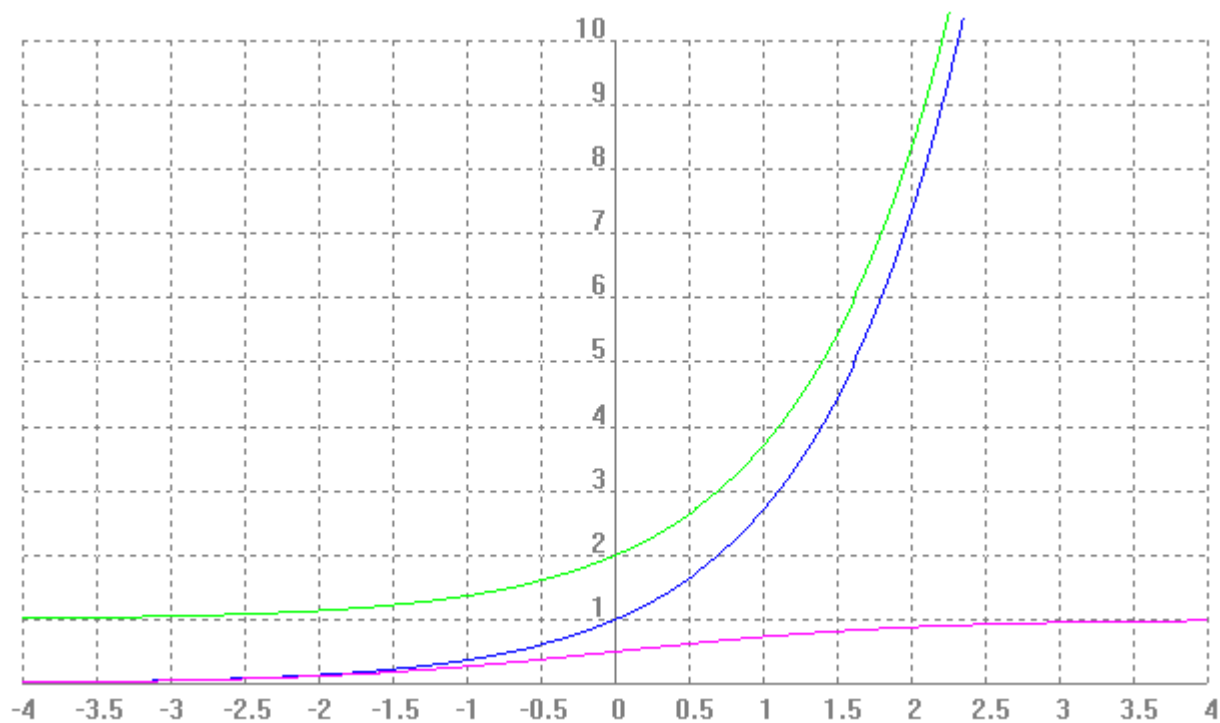
Função exponencial...



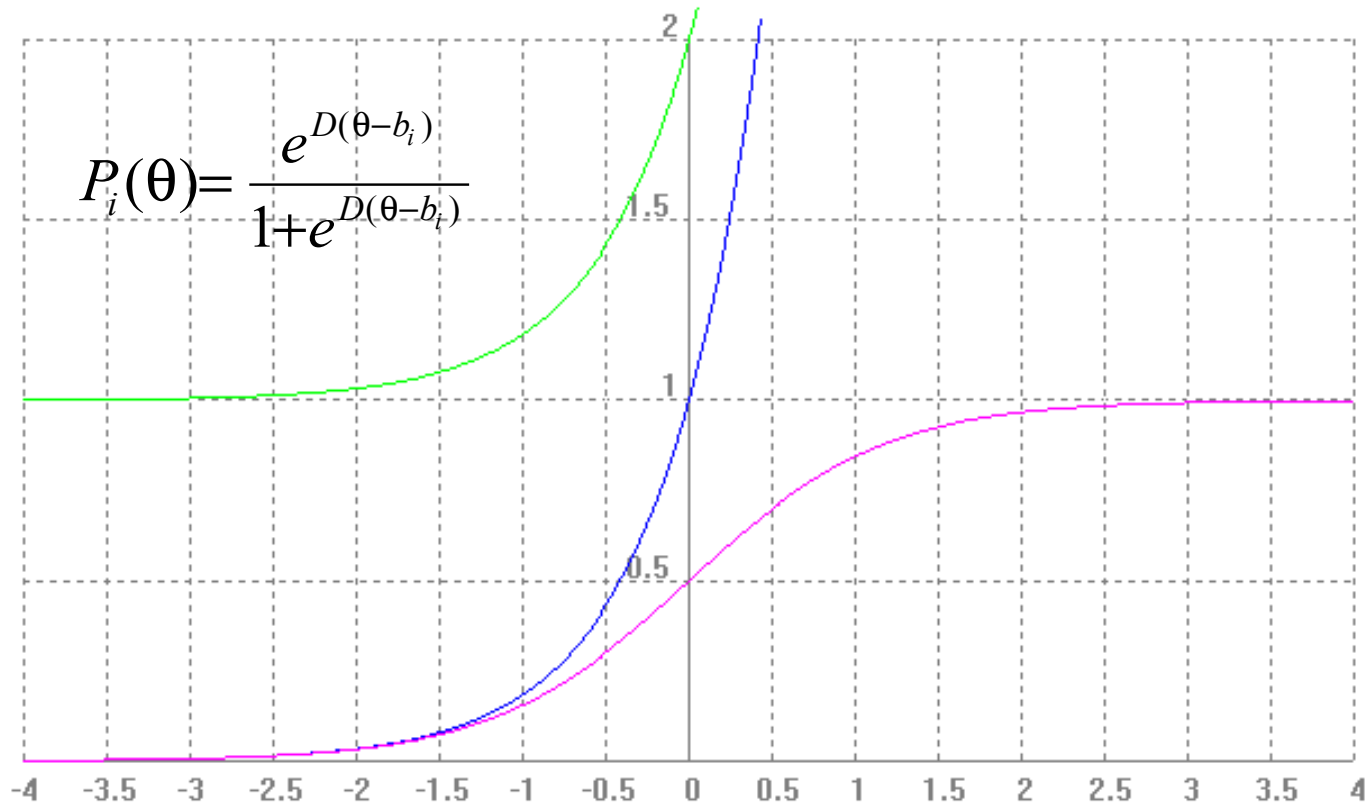
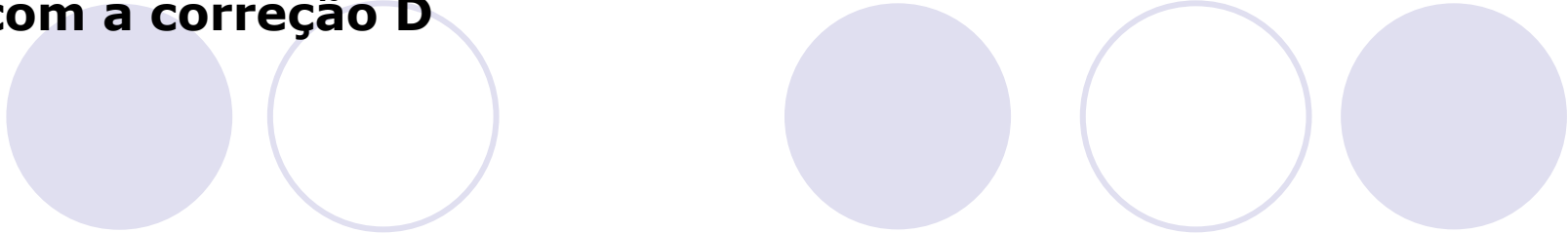
Resultando



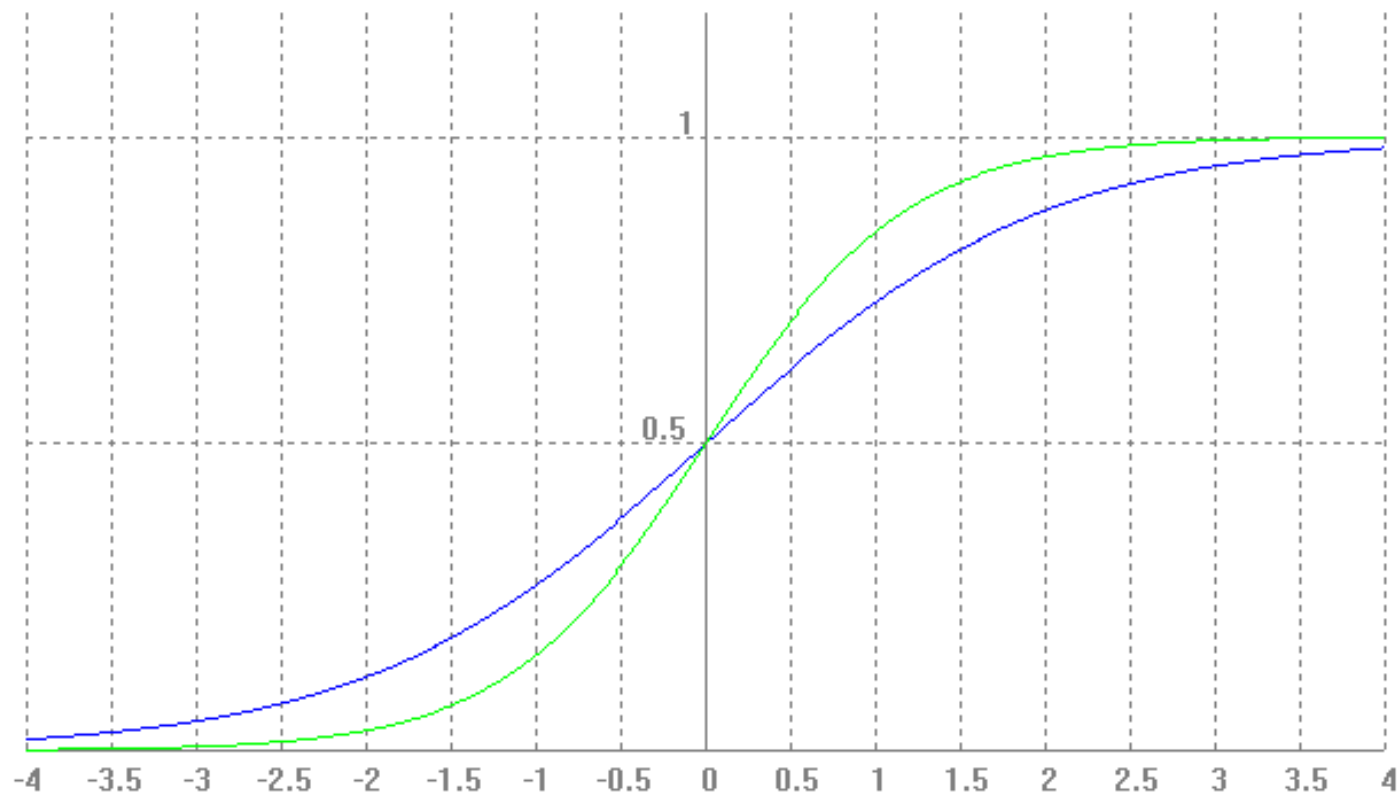
Resultando 2



Agora com a correção D



Curvas logísticas original e ajustada...



Modelo de 3 parâmetros .. Efeito colateral ao parâmetro “b”

A side effect of using the guessing parameter c is that the definition of the difficulty parameter is changed. Under the previous two models, b was the point on the ability scale at which the probability of correct response was .5. But now, the lower limit of the item characteristic curve is the value of c rather than zero. The result is that the item difficulty parameter is the point on the ability scale where:

$$\begin{aligned}P(\theta) &= c + (1 - c) (.5) \\ &= (1 + c)/2\end{aligned}$$

This probability is halfway between the value of c and 1.0. What has happened here is that the parameter c has defined a floor to the lowest value of the probability of correct response. Thus, the difficulty parameter defines the point on the ability scale where the probability of correct response is halfway between this floor and 1.0.

The discrimination parameter a can still be interpreted as being proportional to the slope of the item characteristic curve at the point $\theta = b$. However, under the three-parameter model, the slope of the item characteristic curve at $\theta = b$ is actually $a(1 - c)/4$.

Interpretação do parâmetro de discriminação e dificuldade (Baker, 2001 pg. 35)

- Discriminação

Verbal label	Range of values
none	0
very low	.01 - .34
Low	.35 - .64
moderate	.65 - 1.34
High	1.35 - 1.69
Very high	> 1.70
Perfect	+ infinity

Table 2-4. Labels for item discrimination parameter values

- Dificuldade

- Significado relativo ao ponto mediano (iremos discutir isso quando falarmos sobre equalização e métrica da escala de theta)
- Maneira mais apropriada é dizer sobre "onde", na escala, o item funciona melhor em termos de discriminação.
- A inclinação (*slope*) da CCI é mais acentuada na região da escala ao redor da dificuldade do item (o máximo é exatamente no ponto da escala correspondente à dificuldade do item)
- O item diferencia melhor pessoas com habilidades próximas à sua dificuldade
- Portanto a dificuldade nos diz que região da escala o item funciona melhor. Um significado associado à sua localização.
- Não se diz se o item é fácil ou difícil mas sim em qual lugar ele tem dificuldade mediana pois nesses locais ele é mais discriminativo.

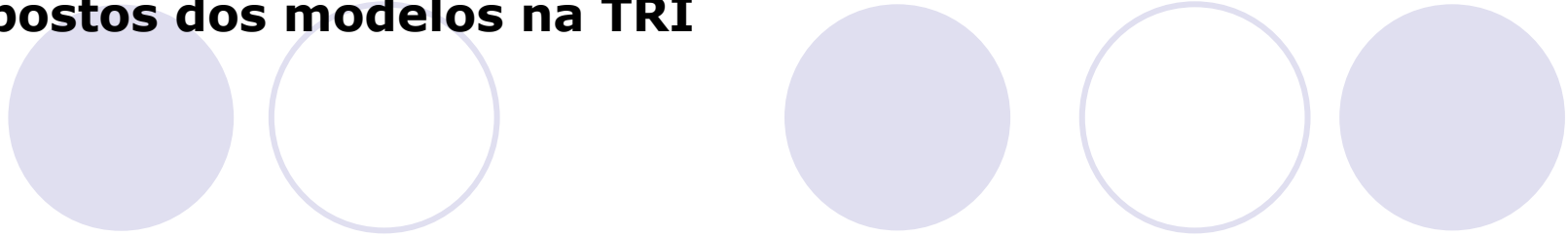
Pressupostos dos modelos na TRI



- **Unidimensionalidade**

- As teorias do traço latente (que vêm dos anos 30) afirmam que existe um conjunto de traços latentes que estão por detrás de um desempenho comportamental qualquer.
- Entretanto, em sua maioria, os modelos da TRI postulam que há apenas uma capacidade (θ) responsável pela realização de um conjunto de tarefas (itens de um teste).
- Parece pacífico que qualquer desempenho humano é sempre multideterminado ou multimotivado, dado que mais de um traço latente entra na execução de qualquer tarefa, como o modelo geral do traço latente afirma.
- Contudo, para satisfazer o postulado da unidimensionalidade, é suficiente admitir que haja uma aptidão *dominante* (um fator ou traço dominante) responsável pelo desempenho num conjunto de itens de um teste. Esse fator é o que se supõe estar sendo medido pelo teste.


Pressupostos dos modelos na TRI



- **Independência local**

- Este postulado afirma que, tendo controlando o fator dominante, as respostas dos sujeitos a quaisquer dois itens são estatisticamente independentes. Isso implica que o desempenho do sujeito num item não interfere no seu desempenho em outro item, já que o θ desse sujeito não se altera ao longo do teste.
- Dizendo de outro modo, a probabilidade do sujeito acertar um item i é função da diferença entre sua capacidade e dificuldade do item. No próximo item do teste, $i+1$, a probabilidade de acerto continua sendo governada por essa diferença.
- O fato do sujeito ter acertado ou não o item anterior não interfere no acerto no próximo item. Assim a resposta a cada item está exclusivamente em função do tamanho do seu θ dominante.

Independência local

- 
- Embora pareça improvável que os comportamentos (respostas) de um mesmo sujeito não estejam correlacionados, a independência local afirma que, se houver correlação, esta se deve à influência de fatores outros que não o fator dominante.
 - Se estes outros fatores forem controlados (mantidos constantes), o fator dominante será a única fonte de variação e as respostas se tornam independentes, porque o sujeito responde exclusivamente em função do tamanho do seu θ . Assim, a independência local implica também a unidimensionalidade, já que a única causa da resposta do sujeito é este suposto θ dominante (Lord, 1980; Lord & Novick, 1968)
 - Se esse pressuposto é atendido, a probabilidade de um vetor de resposta é igual a multiplicação das probabilidades de cada resposta individual.