

Teorias de Avaliação - CE095

Adilson dos Anjos¹

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

Curitiba, PR
14 de junho de 2016

Teoria da Resposta ao Item –Teste Adaptativo Computadorizado–

Introdução

- CAT = **C**omputerized **A**daptive **T**esting
- TAI = **T**este **A**daptativo **I**nformatizado

Introdução

O que é?

- Um procedimento onde os itens são apresentados de acordo com a habilidade do respondente;
- Utiliza modelos da TRI: dicotômicos ou politômicos
- Permite estimar um traço latente com poucos itens (em relação a um teste convencional);
- Necessidade de utilização de recursos computacionais;

Introdução

Questões:

- Qual item apresentar primeiro?
- Como estimar o θ ?
- Quais itens utilizar?
- Quando parar?

Introdução

Aplicações:

- Plataforma Concerto;
- catR no R (simulação);
- catIrt no R (simulação);
- Plataformas/Sistemas próprios: livres e pagos.

Introdução

Aplicações

- TOEFL: exame de inglês
- NAEP: avaliação do progresso de estudantes do ensino fundamental e médio dos EUA;

Introdução

- Um conjunto de itens para cada indivíduo;
- Necessidade de um banco de itens;
- Em geral, pelo menos 3 vezes o tamanho do instrumento de medida (definido por simulação)

Introdução

- Banco de itens calibrados pela TRI;
- Tamanho do banco depende do tipo de aplicação:
- Testes de seleção (educacional): banco maior (segurança)
- Testes de avaliação: banco menor

Característica de um CAT:

Introdução

- Em geral são aplicados menos itens do que um teste convencional sem perda de precisão;
- Economiza itens e causa menos *cansaço* nos respondentes;
- Em geral são mais seguros: itens são apresentados parcialmente;

Característica de um CAT:

Introdução

- Pode-se utilizar o CAT para pré-testar itens novos;
- Permite o uso de tecnologias diferentes: por exemplo, um vídeo, áudio, animação;
- Pode ser utilizado no EAD.

Introdução

Outras características:

- Critério de exposição dos itens (não repetir itens);
- Seleção de itens por conteúdo: balanceamento e melhoria da estimação da proficiência;
- Avaliação de DIF;
- Uso de vários modelos da TRI

Introdução

Algoritmo geral para aplicação do teste:

- 1 Seleção do primeiro Item;
- 2 Estimação provisória da proficiência;
- 3 Regra de parada ou seleção do próximo item e atualização da proficiência;
- 4 Finalização: estimativa final da proficiência (e interpretação da escala);

Introdução

Seleção inicial do item:

- Iniciar com um item médio ($\theta = 0$);
- Iniciar com um item ao redor de $\theta = 0$: sorteio;
- Iniciar com vários itens (n=3 a 6).
- Itens menos discriminativos no início e mais discriminativos no final.

Introdução

Seleção dos itens seguintes:

- Itens mais informativos próximos de $\hat{\theta}$
- Sorteio de itens próximos de $\hat{\theta}$;
- entre outros métodos;

Introdução

Regra de parada:

- EP da proficiência menor do que um valor pré-definido;
- Diminuição sucessiva do EP: tende a se aproximar do θ ;
- Estimativa de θ alcançada;
- Número máximo de itens: evitar fadiga;
- Término dos itens

Introdução

Outras características:

- Não é necessário que todos realizem o teste simultaneamente;
- Tem alto custo: banco de itens + recursos tecnológicos
- Permite a inclusão de novos itens ao longo do tempo

Introdução

Sugestões de Leitura/Referências

- Dissertação USP
- <http://www.psychometrics.cam.ac.uk/newconcerto>Concerto
- <https://repositorio.ufsc.br/bitstream/handle/123456789/95506/299657.pdf?sequence=1>Tese UFSC