

Teoria da Medição de Variáveis Latentes: história e utilização atual

Carlos A.Q. Coimbra
MAST / MCT

Variáveis Latentes

- * Definição mais comum:
é a variável que não pode ser medida diretamente
- * Definição de Bollen, 2001
é a variável que não consta da base
- * Esta definição inclui
os construtos teóricos mais estudados:
inteligência, nível socio-econômico, qualidade de vida...
e também outras características do modelo:
efeitos aleatórios, dados faltantes...

A Medição de Qualidades

- * Está nas origens do método científico no século 17
- * Projeto *Characteristica* de Leibnitz, 1672:
um cálculo de qualidades
- * Cálculo infinitesimal:
um cálculo de quantidades

Aplicação dos Modelos de Variáveis Latentes

- * *Evaluation, Assessment, Appraisal*
- * Avaliação estatística ou medição estatística
avaliação de desempenho, avaliação de capacidade,
avaliação de qualidade
- * Avaliação em ciências humanas e sociais
 - técnicas de descrição de fenômenos humanos e sociais
 - discriminação de pessoas ou de instituições
- * Um exemplo: Educação:
"os alunos aprenderam o que V ensinou?"

Os 3 Aspectos da Avaliação

- * Golstein & Lewis, 1996
- * O objetivo ou função da avaliação para que será usada?
- * O desenho ou método de avaliação longitudinal, seccional...
- * A unidade a ser avaliada o indivíduo, o coletivo, as unidades sociais...

○ Objetivo da Avaliação

- * I Certificação ou qualificação de indivíduos
são os mais antigos até recentemente os mais importantes
- * II Acompanhamento e direcionamento
de atividades, da aprendizagem, de projetos
- * III Comparação ou qualificação de instituições
grande importância política atualmente

Objetivo I: qualificação de indivíduos

- * Juízes, 12:5-6: um teste com um único item
- * China 2200AC: funcionários públicos testados a cada três anos
China 1115AC: testes para ingresso ao serviço público
- * Uma amostra relativamente pequena do desempenho de indivíduos podia produzir uma previsão bastante precisa sobre seu desempenho sob as mais diversas condições

- * O programa chinês influenciou os filósofos do iluminismo
Voltaire e Quesnay: procedimento adotado em 1791
depois abandonado por Napoleão
- * O programa chinês foi citado na reforma inglesa de 1833
foi adotado na escolha de funcionários do serviço civil na Índia
- * O relatório Jenckes do congresso norte-americano, 1868,
continha descrição detalhada dos sistemas chinês e britânico
serviu de base para o Civil Service Act de 1883.

- * Testes de certificação ou seleção tiveram grande desenvolvimento a partir do final do século 19 e começo do 20: necessidade de criar novas formas de seleção em massa
- * Testes de certificação ou seleção sempre foram importantes e sempre provocaram controvérsia: ninguém discute a necessidade dos testes
- * O que se discute é a qualidade da medição: a comparabilidade dos resultados e as propriedades da escala
- * "A escolha dos procedimentos de avaliação reflete as preferências políticas e culturais daqueles que controlam os sistemas de avaliação"

Objetivo II: acompanhamento da aprendizagem

- * É o tipo de avaliação menos controverso: a aprendizagem deve estar interessada em seu progresso
- * Mesmo assim sua prática é recente
Universidade de Bolonha, 1219: prova oral
Robert de Sorbon, 1257: grupo que deu origem à Sorbonne.
- * Testes escritos nas universidades: Jesuítas, século 16
Documento de 1599: 11 regras sobre a condução de provas praticamente idênticas às regras usadas hoje
- * Testes escritos nas universidades: século 19

- * Século 20: começam as pesquisas sobre as propriedades dos testes
- * Outras disciplinas se interessam
 - Psicologia: teste de Binet, 1905 (inteligência)
 - Estatística: Spearman, 1904 (análise fatorial)
 - Educação: testes de múltipla escolha: Thonrdike e Thurstone
 - Psicologia: Terman e Otis 1916: teste Stanford-Binet
- * Primeira Guerra: Army Alpha, primeiro teste em massa
2 milhões de recrutas examinados

Objetivo III: comparação entre organizações

- * É o tipo mais recente
- * Grande interesse para cientistas sociais
administradores de políticas públicas
profissionais e usuários envolvidos
- * Sociologia: Coleman, 1966: eficácia escolar
Raudenbush e Willms, 1995: eficácia e equidade
- * Avaliação educacional nos EUA: modelos de valor agregado
avaliando escolas e professores

Problemas e críticas

- * Desvios: uma avaliação criada com um objetivo não deve ser usada para outro, desvaloriza os dois. Testes de ensino não devem ser usados para seleção (Broadfoot, 1996)
- * Causalidade: a atribuição de causalidade aos efeitos do modelo nem sempre é justificada: os modelos não medem o que propõem a medir (Rubin, 2004)
- * Ajuste: como é difícil conseguir bom ajuste de modelo há uma tendência à tolerância excessiva: as hipóteses do modelo não são verificadas (Reckase, 2004)

Variáveis Latentes: Psicologia e Educação

- * Relações entre indicadores e variáveis latentes

William James, 1909, *The Meaning of Truth*

- * "O que se quer dizer quando se diz que uma pessoa é prudente? de modo geral significa que a pessoa faz seguro contra riscos, que não joga na roleta que olha antes de pular, etc..."
- * "Se estes são hábitos desta pessoa podemos, por conveniência e brevidade, chamá-la de prudente, adotando esta abstração para seus atos"

variáveis latentes ou construtos teóricos

- * Imagem (latente, abstrata, teórica) →
Indicadores ditados pela experiência cotidiana
- * " Na verdade não esperamos que uma pessoa prudente sempre faça seguro contra riscos ou que nunca jogue na roleta,
- * mas sim, que a probabilidade dela executar estes atos seja maior do que ocorreria com uma pessoa imprudente"

a medição estatística

- * Francis Galton, 1884: as dimensões da mente
analogia com grandezas físicas
- * Psicometria: encontrar maneira de medir tais dimensões
Pearson, físico, aluno de Galton: componentes principais
Spearman, psicólogo, aluno de Galton: análise de fatores
Thurstone, engenheiro: múltiplos fatores
- * 1935: Psychometric Society e revista Psychometrika
1947: Educational Testing Service

teoria clássica dos testes

- * A unidade de análise é o teste
procedimentos altamente formalizados
- * Escore total = Traço latente + Erro

$$T_i = \theta_i + e_i$$

- * confiabilidade do teste, validade do teste
os resultados dependem da população
- * Teoria da Generalizabilidade, Teoria das Facetas

teoria da resposta ao item

- * Loevinger, 1947: unidade de análise: o item
escalabilidade de um conjunto de itens
- * George Rasch, 1950: pessoas e itens na mesma escala
Birnbaum, 1960: novas formas paramétricas
Mokken, 1970: teoria não paramétrica
- * Análise de Fatores + função de ligação logística
= TRI paramétrica (Mellenberg, 1994)
- * informação do teste
os resultados independem da população

Variáveis Latentes: Sociologia

- * As pesquisas de levantamento, *survey*
- * A análise de caminhos e os modelos estruturais
- * A análise multinível

as pesquisas de survey

- * Até início de 1900: levantamentos de problemas sociais
quantificação no sentido amplo
Le Play, 1864, estudar o sentimento religioso
Medição de atitudes
- * Década de 30: recenseamentos em grande escala nos EUA
técnicas de construção de questionários
técnicas de amostragem e coleta
nascem as empresas de pesquisa de opinião
- * A partir do anos 50: avanços teóricos da psicometria
incorporados á pesquisa sociológica

análise de caminhos e modelos estruturais

- * Sewall Wright, 1921: relações de causalidade estudadas a partir da estrutura de covariância: análise de caminhos
- * Redescoberto na década de 60 por econométristas e sociólogos
Variáveis medidas com erros podiam ser usados nos modelos de regressão
- * A síntese conceitual unindo as relações causais com variáveis latentes medidas por indicadores:
Joreskog, 1973: modelo LISREL

análise multinível

- * A sociologia inclui os efeitos de contexto
- * Robinson, 1950: falácia ecológica
Coleman, 1966: equidade escolar
- * Goldstein, 1986; Raudenbush, 1996; Snijders, 1999
- * Síntese: TRI como modelos multinível
(Mislevy Bock, 1989; Adam, Wilson Wu, 1997)
TRI como modelo generalizado de análise fatores
(Mellenberg, 1994; Muthèn, 2001)

Os modelos

- * Primeira hipótese: dimensionalidade do traço latente
 - θ contínua ou ordinal: variável latente
 - θ discreta: classes latentes
- * Segunda hipótese: independência condicional
 - A probabilidade de resposta fica explicada por θ
- * Terceira hipótese: função resposta ao item
 - função com forma definida: modelo paramétrico
 - função sem forma definida: modelo não paramétrico

Formulação geral

- * Estimar um vector de parametros β
a partir dos dados observados $Y = y$
 Y é distribuído segundo a densidade $p(y; \beta)$
 $p(y; \beta)$ é difícil de calcular
- * Hipótese: o modelo pode ser aumentado
com uma variável latente U
produzindo uma densidade condicional $p(y; \beta | U)$
 $p(y; \beta | U)$ é fácil de calcular

exemplo: modelo de regressão logística

- * nível 1: modelo amostral de Bernoulli onde

$Y_{ij} = 1$ com probabilidade φ_{ij}

- * função de ligação logit $\eta_{ij} = \log[\varphi_{ij}/(1 - \varphi_{ij})]$

- * nível 2: linear

$$\eta_{ij} = \sum_{q=1}^Q \beta_q x_{ijq} + \sum_{r=1}^R u_{jr} z_{ijr} \quad \text{ou} \quad \eta_j = X_j \beta + Z_j u_j$$

onde β é um vetor Q -dimensional de efeitos fixos,

$u_j \sim N(0, \Gamma)$ é um vetor R -dimensional de efeitos aleatórios

X_j e Z_j são matrizes $n_{1j} \times Q$ e $n_{1j} \times R$

exemplo: modelo de Rasch como modelo multinível

- * é uma regressão logística de dois níveis
variáveis explicativas são indicadores das respostas
os efeitos aleatórios $u_j \sim N(0, \tau^2)$ são os traços latentes
- * as equações do model

$$\eta_{ij} = \sum_{q=1}^Q \delta_q d_{ijq} + \tau w_j$$

- * o vetor de parâmetros é $\beta = (\delta', \tau)'$

Estimação de Modelos Não Lineares

- * Aproximando a verossimilhança
aproximação quadrática da log-verossimilhança
equivalent to MQL-1
Longford 87
- * Aproximando as equações do modelo
MQL 1 e 2, PQL 1 e 2
série de Taylor de primeira e segunda ordem
Goldstein 91, Breslow & Clayton 93, Goldstein & Rasbash 96

- * Integração numérica da equação de verossimilhança
quadratura de Gauss-Hermite, transformada de Laplace
Hedeker & Gibbons 97, Raudenbush & al. 00
- * Outros métodos
bootstrap paramétrico, amostrador de Gibbs
Kuk 95, Spiegelhalter & al. 96, Browne 98, & Draper 00
- * Problemas: viés para baixo, computação pesada,
dificuldades na convergência
- * minha tese: aproximação estocástica, uma MCMC frequentista

Programas de Equações Estruturais

- * LISREL, o primeiro e ainda o mais usado
usa matriz tetracórica para variáveis ordinais
- * EQS (de P. Bentler), AMOS (de Arbuckle, agora no SPSS)
- * Mx (de Neale): free
- * MPlus (de B. Muthén): o mais moderno e flexível
estima modelos multinível

Programas de Teoria da Resposta ao Item

- * BILOG-MG (Zimovky et al.) usado no SAEB
MULTILOG (Mislevy et al.)
estimação bayesiana, TRI de 3 parâmetros
- * OPLM (Verhelst et al.): itens politômicos
estimação condicional em TRI de dois parâmetros
- * WINSTEPS e BIGSTEPS, modelo de Rasch 30 mil itens, 10 milhões de pessoas
- * ConQuest: modelo de Rasch multidimensional
usado no PISA (OECD)

Programas de Modelos Multinível

- * MLwiN, estimação MV, bootstraap e bayesiana
HLM, transformada de Laplace...
- * BUGS e WinBUGS, bayesiano
- * STATA 9: gllmm de Rabe-Hesketh e Skrondal
- * Multinível em R: [stat.columbia.edu/ gelman/arm](http://stat.columbia.edu/gelman/arm)
- * Psicometria em R: cuddyvalley.org/psychoR

Medindo a mobilidade física

- 1 Consigo me levantar da cama sozinho
- 2 Consigo ir ao banheiro sozinho
- 3 Consigo ir ao supermercado sozinho
- 4 Consigo caminhar três quilômetros
- 5 Consigo correr dez quilômetros
- 6 Consigo correr a Maratona

Medindo a motivação

- * O indivíduo realiza uma atividade
 - motivação intrínseca: o porque está na própria atividade
 - motivação extrínseca: o porque está fora da atividade
- * Motivação: operacionalizada no "porque" ela é realizada
- * 60's: behaviorismo: recompensa para controlar comportamento
 - 70's: recompensa funcionava enquanto estava operante
 - Deci, 1971: algumas atividades têm sua própria recompensa inerente

- * Harter, 1981: motivações como opostos de um contínuo
M. Extr. \longleftrightarrow M. Intr.
declínio gradual na MI em estudantes de 3a à 8a série
- * Pajares & Grahan, 1999: valor atribuído à disciplina
(= importância percebida + interesse associado)
decrece para matemática e ciências de 1a à 12a série
- * Gottfried et al., 2001: MI declina pouco para leitura
e substancialmente para matemática e ciências

- * O resultado de Harter tem que ser interpretado com cuidado conceitualizando as duas motivações no mesmo contínuo não se tem avaliação independente de cada uma
- * Lepper et al., 2005: usa o instrumento de Harter (1981) mas reconhece MI e ME como dimensões diferentes
MI decresce da 3a à 8a série e ME não muda sem efeito de gênero e etnia
Preocupante: baixo nível de MI em estudantes mais velhos
- * Outra solução: teoria da auto determinação

a teoria da auto determinação

- * Deci & Ryan, 1999 e 2002: 3 formas de motivação extrínseca
 - 1 regulação externa (estudo para agradar...)
 - 2 regulação introjetada (me obrigo a estudar...)
 - 3 regulação identificada (estudar é bom pra mim...)

- * a teoria propõe ainda uma forma extrema de motivação a desmotivação (estudar não vale a pena...)

- * tipologia mais complexa
explica melhor os resultados experimentais

- * Duas outras características da teoria da auto determinação
- * Cada forma de motivação pode ser colocada em um contínuo crescente de auto determinação
motivação não auto determinada (desmot., r.ext. e r.introj.)
motivação auto determinada (r.identif. e mot.intr)
- * A motivação é maleável: mudanças ocorrem naturalmente
processo de internalização tendência inata que serve para promover o desenvolvimento da motivação intrínseca