Groenen & Andries Van Der Ark (2006) and **Stout (2002)**

蔡介文

May 24, 2021

Papers

- Groenen, P. J., & Andries Van Der Ark, L. (2006). Visions of 70 years of psychometrics: the past, present, and future. *Statistica Neerlandica, 60*(2), 135-144.
- Stout, W. (2002). Psychometrics: From practice to theory and back. *Psychometrika, 67*(4), 485-518.

3 topics

1. Latent Dimensionality Structure
2. Test Fairness
3. Skills (Cognitive) Diagnostic Models

future research topics (2006-2016?)

- **Takane and De Boeck**: robustness studies, data mining procedures, (m)anova with random effects, (m)anova in SEM, Bayesian networks, functional estimation, longitudinal data analysis, small sample behavior, regularization techniques, and statistics freed from normality.
- **Meulman**: MCMC and bootstrapping
- **Bockenholt**: methods for investigating individual differences.
- **Van der Linden**: real-time applications (real-time IRT), computer adaptive testing.
- **Browne**: the change over time methodology.
- **Heiser**: multidimensional IRT and IRT in structural modeling.

IRT major modeling assumptions

(see Hatzinger, 2010)

1. **unidimensionality.** response probability does not depend on other variables φ.

$$P(X_{vi} = 1|\theta_v, \beta_i, \varphi) = P(X_{vi} = 1|\theta_v, \beta_i)$$

2. **sufficiency.** (Rasch) raw score $r_v = \sum_i x_{vi}$ (sum of responses) contains all information on ability, regardless which items have been solved.
3. **conditional (local) independence.** for fixed θ there is no correlation between any two items
4. **monotonicity.** response probability increases with higher values of θ

Dimensionality - define test unidimentionality

def 1 MLI1, marginal IRF (ICC)

$$P(\mathbf{U} = \mathbf{u}) = \int_{-\infty}^{\infty} P(\mathbf{U} = \mathbf{u} | \Theta = \theta) f(\theta) d(\theta)$$

def 2 strongly locally independent

$$P(\mathbf{U} = \mathbf{u} | \Theta = \theta) = \prod_{i=1}^{n} P(U_i = u_i | \Theta = \theta)$$

def 3 weak locally independent

$$\mathrm{Cov}(U_i, U_i' | \Theta = \theta) = 0$$

def 4 essential unidimensional

$$\frac{\sum_{1\leq i<i'\leq n}|\mathrm{Cov}(U_i, U_i'|\Theta=\theta)|}{\binom{n}{2}} \to 0 \quad n\to\infty$$

def 5

$$P_i(\theta)=\mathrm{H}_i(\sum_{j=1}^{d} a_{ij}\theta_j - b_i), \quad \mathrm{H}_i(-\infty)\geq 0, \quad \mathrm{H}_i(\infty)=0$$

some unidim. test indices

Section 3 discussed how to estimate the expected conditional covariance, and two types of estimators were presented there. After obtaining an estimator $\widehat{E\text{cov}}_{i_1 i_2}$, either $\widehat{E\text{cov}}_{i_1 i_2}(\alpha)$ in (18) or $\widehat{E\text{cov}}^*_{i_1 i_2}$ in (20), it is easy to construct an estimator of the theoretical DETECT by substituting the expected conditional covariances with their corresponding estimators, that is,

$$\widehat{D}(\mathcal{P}) = \frac{2}{n(n-1)} \sum_{1 \le i_1 < i_2 \le n} \delta_{i_1 i_2}(\mathcal{P})\, \widehat{E\text{cov}}_{i_1 i_2},$$

where $\delta_{i_1 i_2}(\mathcal{P})$ is defined by (23). Similarly, one can construct estimators for ASSI($\mathcal{P}$) and $R(\mathcal{P})$, that is,

$$\widehat{\text{ASSI}}(\mathcal{P}) = \frac{2}{n(n-1)} \sum_{1 \le i_1 < i_2 \le n} \delta_{i_1 i_2}(\mathcal{P})\, \text{Sgn}(\widehat{E\text{cov}}_{i_1 i_2})$$

and

$$\widehat{R}(\mathcal{P}) = \frac{\sum_{1 \le i_1 < i_2 \le n} \delta_{i_1 i_2}(\mathcal{P})\, \widehat{E\text{cov}}_{i_1 i_2}}{\sum_{1 \le i_1 < i_2 \le n} \left| \widehat{E\text{cov}}_{i_1 i_2} \right|}.$$

- Zhang, J. (2007). Conditional Covariance Theory and Detect for Polytomous Items. *Psychometrika, 72*(1), 69–91. doi:10.1007/s11336-004-1257-7

conditional covariances

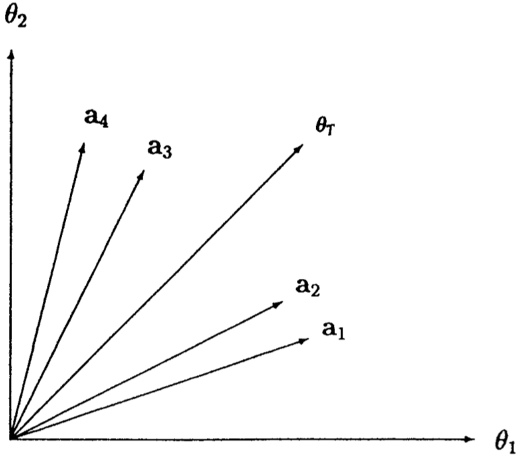


FIGURE 1.
Geometric representation of a four item two-dimensional test.

conditional covariances

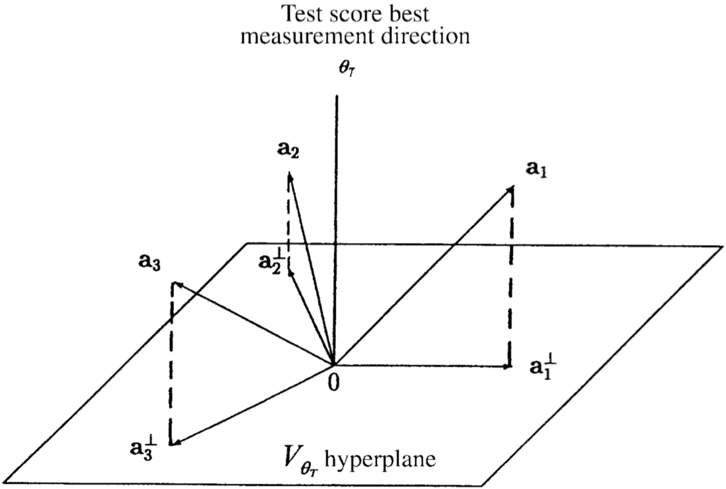


FIGURE 2.
A three dimensional test with projections of item discrimination vectors onto V_{θ_T} hyperplane.

conditional covariances

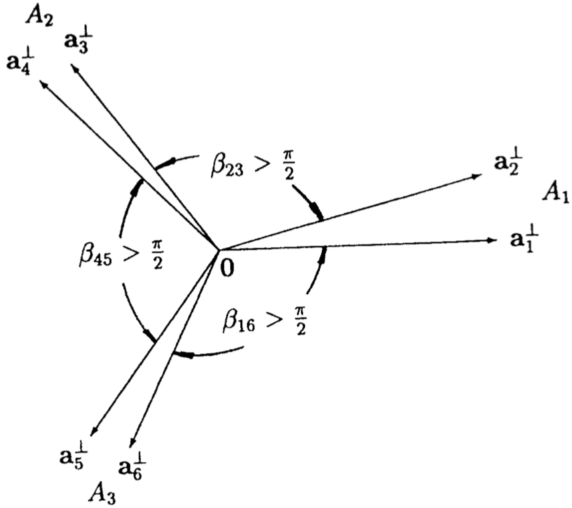


FIGURE 3.
Projection of item discrimination vectors onto V_{θ_T} hyperplance for a six item three-dimensional approximate sample structure.

sirt

- **conf.detect()** Confirmatory DETECT and polyDETECT Analysis
- `expl.detect()`
- `unidim.test.csn()`

The classification scheme of these indices are as follows (Jang & Roussos, 2007; Zhang, 2007):

Strong multidimensionality	DETECT > 1.00
Moderate multidimensionality	.40 < DETECT < 1.00
Weak multidimensionality	.20 < DETECT < .40
Essential unidimensionality	DETECT < .20

Maximum value under simple structure	ASSI=1	RATIO=1
Essential deviation from unidimensionality	ASSI > .25	RATIO > .36
Essential unidimensionality	ASSI < .25	RATIO < .36

DIF - MMD model

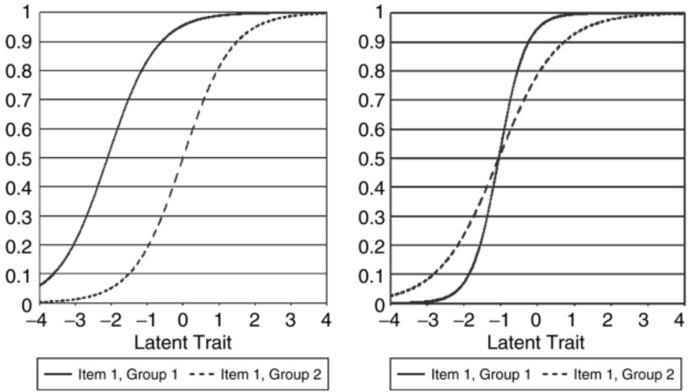
Table 1
Effect on DIF of Two Factors From the Shealy-Stout MMD

Conditional Probability Density Functions	Probability Correct $P(\theta,\eta)$	Probability Correct $P(\theta)$
$f_R(\eta\|\theta) \neq f_F(\eta\|\theta)$	DIF	No DIF
$f_R(\eta\|\theta) = f_F(\eta\|\theta)$	No DIF	No DIF

- $P(\theta)$: indicates that the probability of a correct response depends only on θ and not on η. There can, in principle, be a different amount of DIF at different values of 0. For example, this occurs in **crossing DIF**.

DIF - crossing-DIF (i.e. non-uniform DIF)

Graphical Representation of Uniform and Nonuniform (Crossing) Differential Item Functioning



(see Holmes & French, 2007).

DIF - MMD model

marginal IRF

$$P_g(\theta) = \int_{-\infty}^{\infty} P(\theta, \eta) f_g(\eta|\theta) d\eta$$

- g = R (reference) or F (focal) group
- $f_g(\eta|\theta)$: the density of H_g given $\Theta_g = \theta$

amount of DIF at θ

$$B(\theta) = P_R(\theta) - P_F(\theta)$$

average amount of unidirectional DIF

$$\beta_{UNI} = \int_{-\infty}^{\infty} B(\theta) f_F(\theta) d(\theta)$$

- $f_F(\theta)$: density of Θ_F.
- β_{UNI}: the fundamental DIF index of MMD.

expected diff in the means of η given a fixed θ

$$E_R(\eta|\theta) - E_F(\eta|\theta) = (\mu_{\eta_R} - \mu_{\eta_F}) - \rho(\mu_{\theta_R} - \mu_{\theta_F})$$

(see Roussos & Stout, 1996)

Diagnostic Models

- LLTM as an item explanatory model under the **Explanatory IRT** framework (De Boeck & Wilson, 2004)

TABLE 2.2. Summary of the four models.

	$\eta_{pi} =$			
Model	Person part	Item part	Random effect	Model type
Rasch model	θ_p	$-\beta_i$	$\theta_p \sim N(0, \sigma_\theta^2)$	Doubly descriptive
Latent reg Rasch model	$\sum_{j=1}^{J} \vartheta_j Z_{pj} + \varepsilon_p$	$-\beta_i$	$\varepsilon_p \sim N(0, \sigma_\varepsilon^2)$	Person explanatory
LLTM	θ_p	$-\sum_{k=0}^{K} \beta_k X_{ik}$	$\theta_p \sim N(0, \sigma_\theta^2)$	Item explanatory
Latent reg LLTM	$\sum_{j=1}^{J} \vartheta_j Z_{pj} + \varepsilon_p$	$-\sum_{k=0}^{K} \beta_k X_{ik}$	$\varepsilon_p \sim N(0, \sigma_\varepsilon^2)$	Doubly explanatory

Diagnostic Models - RM

$$\eta = \theta - \beta$$

$$\eta = \log(\frac{\pi}{1-\pi}) = \log(\frac{\exp(\theta)}{\exp(\beta)})$$

Rasch Model

$$\pi = \frac{\exp(\theta - \beta)}{1 + \exp(\theta - \beta)}$$

Diagnostic Models - LLTM

- Fischer's (1973) linear logistic trait/test model (LLTM) [notations from De Boeck]

$$\eta_{pi} = \theta_p - \sum_{k=0}^{K} \beta_k X_{ik}$$

  - η_{pi}: examinee p's prob.of success on item i.
  - θ_p: latent trait for examinee p.
  - β_k: regression weight of item property k.
  - X_{ik}: the value of item i on item property k $(k = 0, ..., K)$

discrimination vector

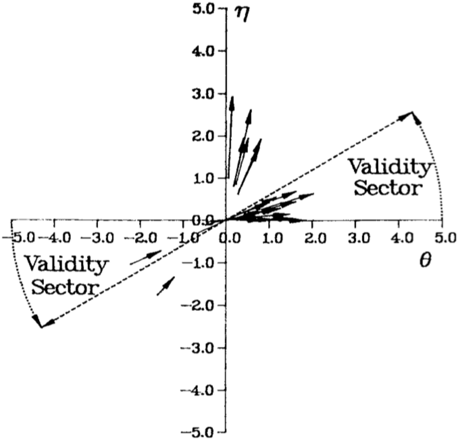


FIGURE 5.
Item discrimination vectors of a 22 item validity sector.

Diagnostic Models - Q matrix approach

a simplified ver. of **Unified Model (UM)** (DiBello, Stout, & Roussos, 1995) [notations form **reparameterized unified model (RUM)** (Henson, Roussos, & Templin, 2005)

$$P(X_{ij} = 1|\alpha_j, \eta_j) = \pi_i^* \prod_{k=1}^{K} r_{ik}^{*(1-\alpha_{jk})\times q_{ik}} P_{ci}(\eta_j)$$

- Henson, R. A., Roussos, L., & Templin, J. L. (2005). *Fusion model "fit" indices.* Unpublished ETS project report, Princeton, NJ.

Software

1. Dimensionality

DIMTEST, **HCA/CCPROX**, **DETECT**, **CONCOV**

2. DIF

SIBTEST, **MH DIF**

3. Diagnostic model

MCMC Bayes UM