

# Teorie odpovědi na položku

---

PSY028 | JARO 2018 | BLOK 6  
DIF a multigroup IRT modely  
Další témata

# Počítačové adaptivní testování

---

Balíček mirtCAT, který používá Shiny rozhraní a MIRT package pro CAT.

- <https://github.com/philchalmers/mirtCAT>
- V případě potřeby lze využít RStudio nainstalované na webu Testfora.

Software Concerto od Cambridge Psychometrics Centre

- <https://www.psychometrics.cam.ac.uk/newconcerto>

# Multigroup IRT

---

Stejná logika jako multigroup konfirmační faktorová analýza.

Lze použít pro více různých účelů:

- DIF analýza, DTF analýza.
- Testová/faktorová invariance.

Jedna analýza se spustí nad více skupinami najednou.

- Některé parametry mohou být různé, a nebo shodné (omezené na stejnou hodnotu) napříč těmito skupinami.
- Sleduje se, zda a nakolik se model zhoršuje při omezování parametrů.

# IRT invariance: binární položky

	slopes	intercepts	lat. means	lat. variances
konfigurální	free	free	0	1
metrická	constrained	free	0	free*
skalární	constrained	constrained	free*	free*
(striktní)	constrained	constrained	0	free*
(paralelní)	constrained	constrained	0	1

\* V první skupině fixované, v dalších volně odhadnuté.

Nelze testovat reziduální invariance jako v CFA, protože chyba měření (informační funkce položky) je přímo funkcí jednotlivých parametrů.

# IRT invariance: polytomní položky

---

Není úplně jednoduché odhadovat zvlášť skalární a metrickou invarianci.

- Metrická invariance vyžaduje „stejný“ průběh charakteristické křivky položky.
- Charakteristická funkce položky je definována jak diskriminačním parametrem, tak prahy.

Metrická invariance lze testovat jen u modelů s tradiční IRT parametrizací, které obsahují zvlášť parametr obtížnosti a zvlášť parametry jednotlivých prahů.

- Tedy modified graded response model, případně PCM rodina modelů.
- Např. gpcmIRT, grsmIRT

# IRT invariance: polytomní položky

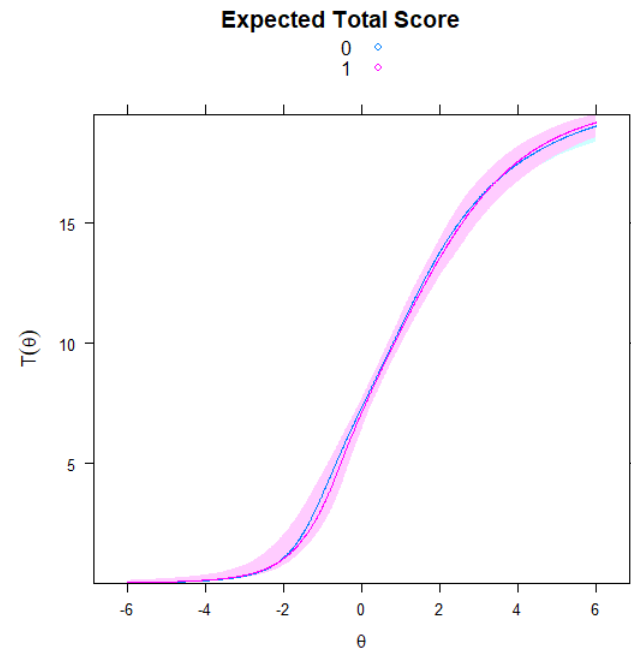
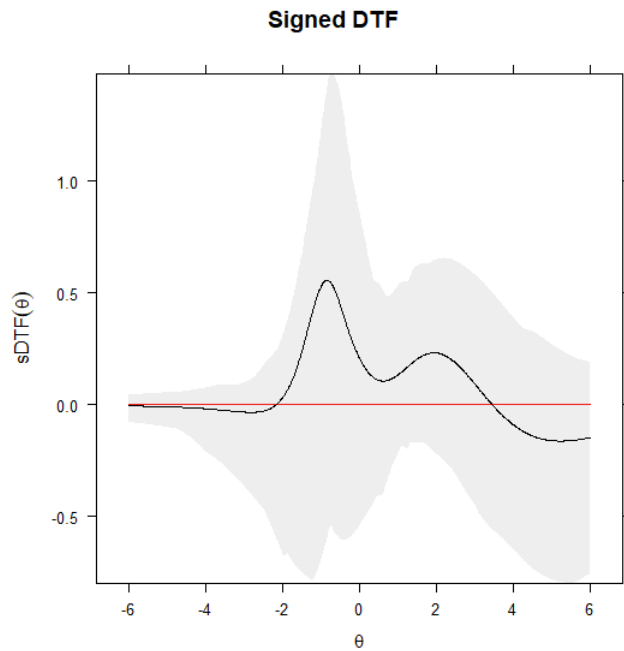
	a (diskr.)	b (difficulty)	$\tau$ thresholds	lat. means	lat. variances
konfigurální	free	free	free	0	1
metrická	constrained	free	constrained	0	free*
skalární	constrained	constrained	constrained	free*	free*
(striktní)	constrained	constrained	constrained	0	free*
(paralelní)	constrained	constrained	constrained	0	1

\* V první skupině fixované, v dalších volně odhadnuté.

Nelze testovat reziduální invariance jako v CFA, protože chyba měření (informační funkce položky) je přímo funkcí jednotlivých parametrů.

# Difference test functioning

Jak velký vliv má noninvariance na skórování testu – tedy vztah odhadovaných a skutečných latentních rysů/hrubých skórů?



# DIF: Differential item functioning (analysis)

---

DIF popisuje situaci, kdy položka „funguje“ rozdílně napříč různými skupinami respondentů.

Protože „fungování“ položky je definované prostřednictvím charakteristické funkce položky, lze testovat rozdílnost jednotlivých parametrů.

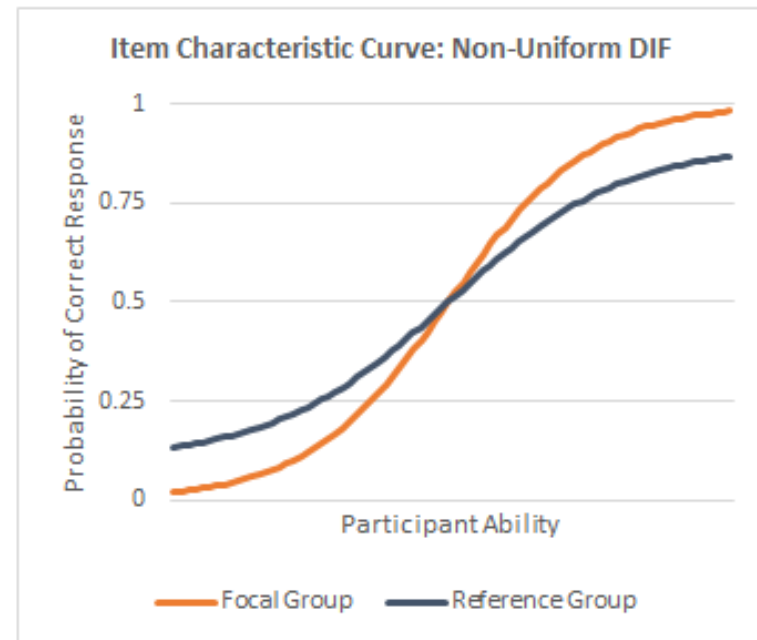
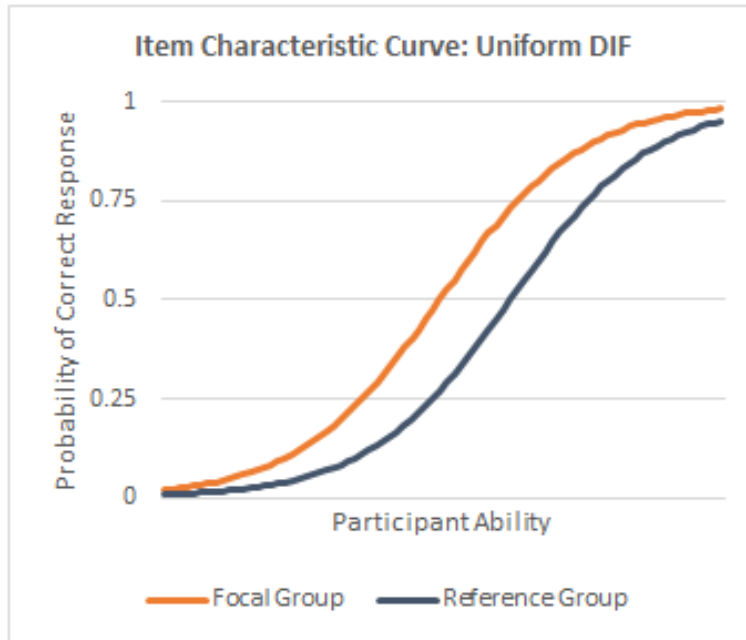
Uniform DIF: položky se liší obtížností.

Non-uniform DIF: položky různě diferencují (liší se diskriminačním parametrem).



# DIF: Differential item functioning (analysis)

---



# Postupy ověření DIF: Raschův model a Winsteps

---

## t-test

- Prosté srovnání parametrů obtížností položek napříč skupinami.
- Signifikantní t-test: obtížnosti položek se liší a tedy uniformní DIF.
- Non-uniformní DIF nedává u Raschova modelu příliš smysl, protože všechny položky  $a=1$ .
  - Lze nicméně stratifikovat respondenty do skupin podle úrovně latentního rysu a (s pomocí interakce s příslušností ke skupině) testovat non-uniform DIF opět t-testem.

## Mantel-Haenszel chi-squared test

- Postup pro ověřování DIF převzatý z CTT.
- Respondenti jsou stratifikováni do více (typicky pěti) skupin podle latentního rysu.
- Pomocí testu dobré shody se ověřuje stejná kumulativní distribuce těchto frekvencí.

# DIF v IRT

---

Lordův (1980) Waldův test.

- Ve dvou skupinách pro  $n$  parametrů naráz má chí-kvadrát rozdělení s  $n$  stupni volnosti a statistikou:

$$\chi^2 = \frac{(v_R - v_F)^T (v_R - v_F)}{(\Sigma_R + \Sigma_F)}$$

- kde  $v_R$  a  $v_F$  jsou vektory maximum likelihood odhady parametrů ve skupině R a F, a  $\Sigma_R$  a  $\Sigma_F$  jsou varianční a kovarianční matice chyb těchto odhadů.
- Lze testovat omezování, nebo naopak uvolňování parametrů napříč modely.

Logistická regrese:

- Logistická regrese s pozorovanou odpovědí (závislá proměnná) a odhadem latentního rysu, členství ve skupině (uniform) a interakcí (nonuniform) jako prediktory.
- Signifikance skupiny či interakce signalizuje příslušný DIF.

Řada dalších.

# Odhad v R

---

DIF funkce v mirt package (Waldův test).

Další, specializované balíčky, jako např.:

- difR
- deltaPlotR
- řada dalších