

# Jak na odlišné fungování položek?

Michaela Cichrová

11.12.2025

- Testování DIFu by mělo být běžnou součástí analýzy vícepoložkových měření
- K DIFu dochází, pokud mají respondenti se stejnou hodnotou testované (latentní) proměnné, ale z různých demografických skupin, rozdílnou pravděpodobnost správné odpovědi na položku
  - Pokud je DIF detekován, značí to, že otázka měří i nějakou sekundární proměnnou (problém s validitou, férností)

# Analýza dat maturity z češtiny

Budeme analyzovat data z jara 2024

- Použijeme proměnné
  - prvomaturant ("1" = ANO, "0" = NE)
  - smo16 (typ školy, 16 různých úrovní)
  - termin (náhradní, opravný, řádný)
  - položky začínající na "b"
- Zadání

# Cvičení 1

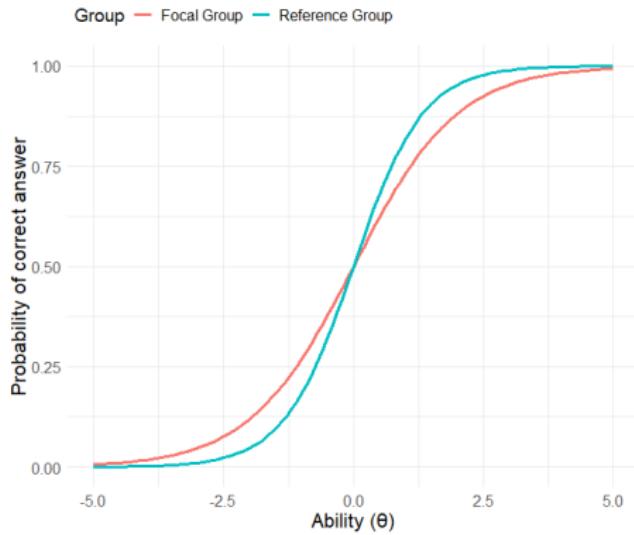
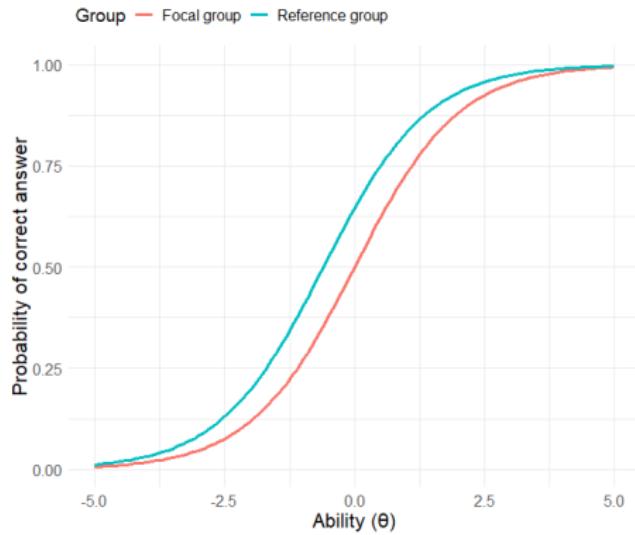
# Cvičení 1

## Shrnutí

- Pracujeme celkem s 25 položkami.
- Pro všechny 3 kategoriální proměnné jsou rozdíly v total score mezi kategoriemi

# Typy DIFu

- Uniformní vs. neuniformní DIF



## Tradiční metody

- Mantel-Haenszel test, SIBTEST

## Regresní modely

- Logistická regrese
  - Latentní proměnná odhadnuta apriori z dat
- IRT modely
  - Latentní proměnná odhadnuta spolu s koeficienty modelu

# Mantel-Haenszel test (MH)

- Rozšíření  $\chi^2$  testu nezávislosti
- Pro testování uniformního DIFu
- Testová statistika je odvozená z četnosti v kontingenční tabulce pro každou hodnotu total score ( $k = 0, \dots, K$ )

$$MH_i = \frac{\left( \left| \sum_{k=0}^K \left( n_{i01k} - \frac{n_{i0+k} n_{i+1k}}{n_{i++k}} \right) \right| - 0.5 \right)^2}{\sum_{k=0}^K \frac{n_{i0+k} n_{i1+k} n_{i+0k} n_{i+1k}}{n_{i++k}^2 (n_{i++k} - 1)}},$$

má  $\chi_1^2$  rozdělení za nulové hypotézy (no DIF)

<b>Skupina</b>	$Y_i = 1$	$Y_i = 0$	<b>Celkem</b>
Ref. skupina (0)	$n_{i01k}$	$n_{i00k}$	$n_{i0+k}$
Fok. skupina (1)	$n_{i11k}$	$n_{i10k}$	$n_{i1+k}$
<b>Celkem</b>	$n_{i+1k}$	$n_{i+0k}$	$n_{i++k}$

# Logistická regrese a IRT

- Vysvětlující proměnné: skupina, odhad latentní proměnné a jejich interakce
- Model:

$$\begin{aligned}\log \left( \frac{\text{P}(Y_{pi} = 1 | \theta_p, G_p)}{1 - \text{P}(Y_{pi} = 1 | \theta_p, G_p)} \right) &= \beta_{i0} + \beta_{i1}\theta_p + \beta_{i2}G_p + \beta_{i3}\theta_p G_p \\ &= (a_i + a_{\text{DIF}_i} G_p)(\theta_p - b_i - b_{\text{DIF}_i} G_p)\end{aligned}$$

- Testování hypotéz:

Jakýkoliv DIF  $H_0 : \beta_{i2} = 0 \text{ and } \beta_{i3} = 0, H_1 : \beta_{i2} \neq 0 \text{ or } \beta_{i3} \neq 0$

Uniformní DIF  $H_0 : \beta_{i2} = 0 \mid \beta_{i3} = 0, H_1 : \beta_{i2} \neq 0 \mid \beta_{i3} = 0$

Neuniformní DIF  $H_0 : \beta_{i3} = 0, H_1 : \beta_{i3} \neq 0$

# Cvičení 2

# Cvičení 2

## Shrnutí

- Pro GY4 vs. GY8 i proměnnou prvomaturant detekovaly DIF všechny 3 metody u většiny položek
- Důvodem může být velká sample size

# Míry velikosti účinku (effect size measures)

- Viděli jsme, že statistické testy vtipovali hodně položek jako difové položky
- Je potřeba kvantifikovat velikost DIFu vhodnou ES
  - pro MH test:  $\Delta_{\text{MH}_i} = -2.35 \log(\alpha_{\text{MH}_i})$ , kde  $\alpha_{\text{MH}_i} = \frac{\sum_{k=0}^K \frac{n_{i01k} n_{i10k}}{n_{i++k}}}{\sum_{k=0}^K \frac{n_{i00k} n_{i11k}}{n_{i++k}}}$
  - Pro logistickou regresi: Nagelkerke's  $\Delta R^2$
- Klasifikace velikosti účinku jako (A) "negligible", (B) moderate, (C) large

# Cvičení 3

# Cvičení 3

## Shrnutí

- Pouze otázka 8: Který z následujících větných celků je syntakticky bezchybný? potenciálně zvýhodňuje studenty osmiletých gymnázií oproti čtyřletým.
- Použití měr velikostí účinku by mělo být nezbytnou součástí analýzy DIFu

**Děkuji za pozornost!**