

Analýzy znalostních testů: Od klasické testové teorie po IRT modely a další témata



Workshop pro odbornou veřejnost
23. 10. 2023

T A
Č R

Projekt TL05000008 Výzvy pro hodnocení znalostí: Analytická podpora tvorby znalostních testů je spolufinancován se státní podporou Technologické agentury ČR v rámci Programu ÉTA 5.

Projekt EduTest – prezentující, autoři analýz

- doc. RNDr. Patrícia Martinková, Ph.D. Pedagogická fakulta UK & Ústav informatiky AV ČR
- Mgr. Eva Potužníková, Ph.D. Pedagogická fakulta UK
- Mgr. Jan Netík Ústav informatiky AV ČR, doktorand PedF UK
- Mgr. Adéla Hladká, Ph.D. Ústav informatiky AV ČR
- Mgr. Lucie Zelená Centrum pro zjišťování výsledků ve vzdělávání (aplikační garant)

Další autoři prezentovaných analýz:

- Mgr. Michaela Vařejková Ústav informatiky AV ČR, doktorandka MFF UK
- Mgr. Jana Dlouhá Ústav informatiky AV ČR, doktorandka FF UK
- MUDr. Ing. Lubomír Štěpánek Ústav informatiky AV ČR, doktorand 1. LF UK
- Mgr. Iván Pérez Ústav informatiky AV ČR, doktorand MFF UK

Harmonogram workshopu

I. Projekt EduTest: od výzkumu k praxi (9:00–9:40)

- Představení projektu, přehled výstupů projektu

Přestávka (9:40–10:00)

II. Od tradičních metod po modelování odpovědí na položky (10:00–11:20)

- Klasická testová teorie, tradiční položková analýza
- Modely teorie odpovědí na položku (IRT), odlišné fungování položek (DIF)

Přestávka (11:20–11:40)

III. Další témata (11:40–12:35)

- Vyrovnávání více verzí testů
- Textová analýza
- Adaptivní testování

Diskuse a závěr (12:35–13:00)

Elektronické materiály:
cs.cas.cz/comps/workshop

I. Projekt EduTest

Od výzkumu k praxi

Projekt EduTest

*Výzvy pro hodnocení znalostí:
Analytická podpora tvorby znalostních testů*



- Financováno Technologickou agenturou ČR
- Řešitelé:
 - Pedagogická fakulta Univerzity Karlovy
(Ústav výzkumu a rozvoje vzdělávání,
Centrum tvorby didaktických testů a psychometrického modelování, CEMP)
 - Ústav informatiky AV ČR
(Oddělení statistického modelování,
Skupina výpočetní psychometrie, COMPS)
- Aplikační garant:
 - Centrum pro zjišťování výsledků vzdělávání (CZV)
- Doba řešení: 4/2021 – 12/2023



EduTest.cz

Projekt EduTest – Cíle

- Vytvořit soubor postupů a nástrojů pro komplexní vyhodnocování výsledků znalostních testů
- Podpořit využívání navržených postupů a nástrojů při vyhodnocování stávajících přijímacích, maturitních a jiných znalostních testů a při vývoji nových testů
- Podpořit informované rozhodování na úrovni škol, krajů a státu

Projekt EduTest – Plánované výstupy

- Soubor postupů a nástrojů pro zkvalitnění tvorby znalostních testů pomocí psychometrických modelů
 - textová část, softwarová část
- Participativní workshopy
 - pro tvůrce testů (aplikační garant, duben 2022)
 - pro uživatele testů (školy apod., prosinec 2022)
 - **pro odbornou veřejnost (2023)**

Aplikace výzkumu psychometrických metod

- EduTest má za cíl posunout náš teoretický výzkum do praxe
- Náš výzkum v oblasti psychometrie
 - Teoretické články v oblasti reliability, detekce odlišného fungování položek, aj.
 - Vývoj algoritmů v oblasti psychometrie
 - Kniha
- Naše výuka v oblasti psychometrie
 - „Vybraná témata z psychometrie“, „Statistické metody v psychometrii“
 - Probíhá na UK
 - [Seminář z psychometrie](#)
 - Počítačové adaptivní testování (19.4.2022)
 - Test equating (semináře v roce 2021)
 - Detekce podvádění, aj.

Výzkum psychometrických metod, na který navazujeme

• Modelování fungování položek

- Martinková, Drabínová et al. (2017). Checking equity: Why differential item functioning analysis should be a routine part of developing conceptual assessments. *CBE—Life Sciences Education*, 16(2), rm2. <https://doi.org/10.1187/cbe.16-10-0307>
- Drabínová & Martinková (2017). Detection of Differential Item Functioning with Nonlinear Regression: A Non-IRT Approach Accounting for Guessing. *Journal of Educational Measurement*, 54(4), 498-517. <https://doi.org/10.1111/jedm.12158>
- Martinková, Hladká, & Potužníková (2020). Is academic tracking related to gains in learning competence? Using propensity score matching and differential item change functioning analysis for better understanding of tracking implications. *Learning and Instruction*, 66, 101286. <https://doi.org/10.1016/j.learninstruc.2019.101286>
- Hladká, Martinková, & Magis (2023). Combining item purification and multiple comparison adjustment methods in detection of differential item functioning. *Multivariate Behavioral Research*. <https://doi.org/10.1080/00273171.2023.2205393>

• Reliabilita

- Martinková, Bartoš, & Brabec M (2023). Assessing inter-rater reliability with heterogeneous variance components models: Flexible approach accounting for contextual variables. *Journal of Educational and Behavioral Statistics*, 48(3), 349-383. <https://doi.org/10.3102/10769986221150517>
- Goldhaber, Grout, Wolff, & Martinková (2021). Evidence on the dimensionality and reliability of professional references' ratings of teacher applicants. *Economics of Education Review*, 83, 102130. <https://doi.org/10.1016/j.econedurev.2021.102130>
- Eroshva, Martinková, & Lee (2021). When zero may not be zero: A cautionary note on the use of inter-rater reliability in evaluating grant peer review. *Journal of the Royal Statistical Society — Series A*, 184(3), 904-919. <https://doi.org/10.1111/rssa.12681>

Výzkum psychometrických metod, na který navazujeme

• Machine learning a textová analýza pro odhad obtížnosti položek

- Štěpánek, Dlouhá, & Martinková (2023). Item difficulty prediction using item text features: Comparison of predictive performance across machine-learning algorithms. *Mathematics*, 11(19), 4104. <https://doi.org/10.3390/math11194104>

• Počítačové adaptivní testování

- Štěpánek & Martinková (2020). Feasibility of computerized adaptive testing evaluated by Monte-Carlo and post-hoc simulations. In *2020 15th Conference on Computer Science and Information Systems (FedCSIS)* (pp. 359-367). IEEE. <https://doi.org/10.15439/2020F197>

• Implementace v rámci statistického software

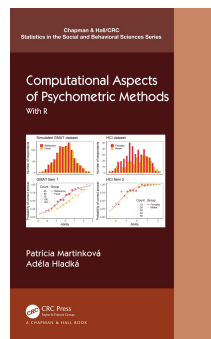
- Martinková & Hladká (2018). ShinyItemAnalysis for Teaching Psychometrics and to Enforce Routine Analysis of Educational Tests. *R Journal*, 10(2). <https://doi.org/10.32614/RJ-2018-074>
- Hladká & Martinková (2020). difNLR: Generalized Logistic Regression Models for DIF and DDF Detection. *R Journal*, 12(1). <https://doi.org/10.32614/RJ-2020-014>
- Martinková P, Drabínová A, Houdek J (2017). ShinyItemAnalysis: Analýza přijímacích a jiných znalostních či psychologických testů. *Testforum*, 9, 16-35. <https://doi.org/10.5817/TF2017-9-129>

Kniha

- Ilustrace důležitých výpočetních aspektů
- Vývoj IRT modelů krok za krokem
- Implementace metod v různých R balíčcích a v ShinyItemAnalysis
- Příklady na reálných i simulovaných datech
- Reprodukovatelný R kód pro analýzy
- Příklady pro výuku

Obsah:

1. Introduction
2. Validity
3. Internal structure of the test and factor analysis
4. Reliability
5. Traditional item analysis
6. Item functioning with regression models
7. Item response theory models
8. More complex IRT models
9. Differential item functioning
10. Outlook on applications and more advanced psychometric topics



Martinková & Hladká (2023). *Computational Aspects of Psychometric Methods: With R* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781003054313>



ShinyItemAnalysis

- Námi vyvíjená interaktivní aplikace
- Využití ve výuce i pro prezentaci výzkumu
- Možnost nahrání a analýzy vlastních dat
- Využívaná po celém světě
- Možnost přístupu offline i online



A screenshot of the ShinyItemAnalysis web application. The interface has a blue header with navigation tabs: Data, Scores, Validity, Reliability, Item analysis, Regression, IRT models, DIF/Fairness, and Modules. Below the header, there are sub-tabs for Data, Basic summary, and Data exploration. The main content area is titled "Data" and contains instructions for using the 20-item GMAT dataset. It includes a "Select dataset" dropdown menu currently set to "GMAT", a "Training datasets" section with a link to a paper by Martinkova et al. (2017), and an "Upload your own datasets" section with a "Choose data (CSV file)" button and detailed instructions on data format requirements. The footer shows the application name, version (1.5.0), and copyright information (© 2023 ShinyItemAnalysis).

SIA moduly

- Zásuvné moduly pro ShinyItemAnalysis (SIA) napsané v {shiny}
- Oboustranná komunikace s hlavní aplikací
- Možnost využít existující infrastrukturu SIA
- Např. možnost nahrání volně dostupných dat CZVV přímo v Excelu bez potřeby data jakkoliv editovat



Moduly a aplikace projektu EduTest

1. SIA modul „CERMAT tools“
 - načtení volně dostupných položkových dat (maturita, JPZ)
 - jsou detekovány odpovědi na položky, skupinové proměnné a kovariáty
 - vyžaduje načtení metadat: soubor s popisem položek
 - možnost poslat data do hlavní ShinyItemAnalysis a využít její funkcionality shiny.cs.cas.cz/ShinyItemAnalysisCermat
 - specifická položková analýza
 - IRT model specifický pro různé typy položek

Moduly a aplikace projektu EduTest

1. SIA modul „CERMAT tools“

- načtení volně dostupných položkových dat, specifická položková analýza
- možnost poslat data do hlavní ShinyItemAnalysis a využít její funkcionality shiny.cs.cas.cz/ShinyItemAnalysisCermat

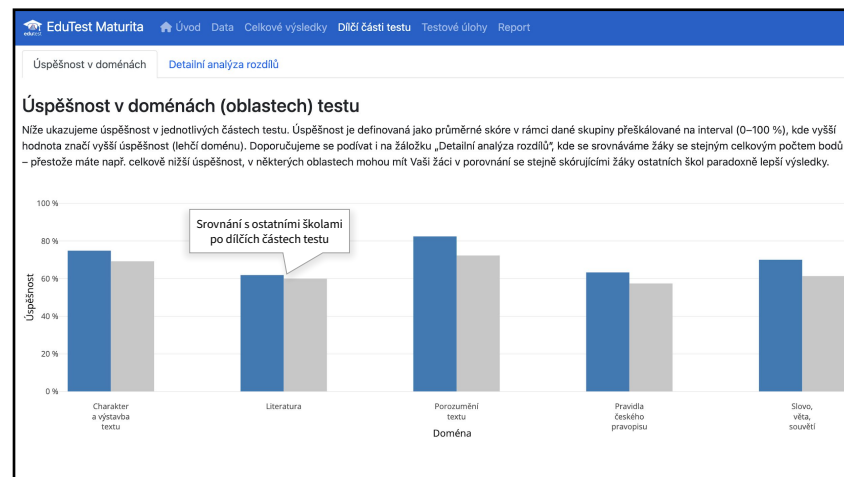
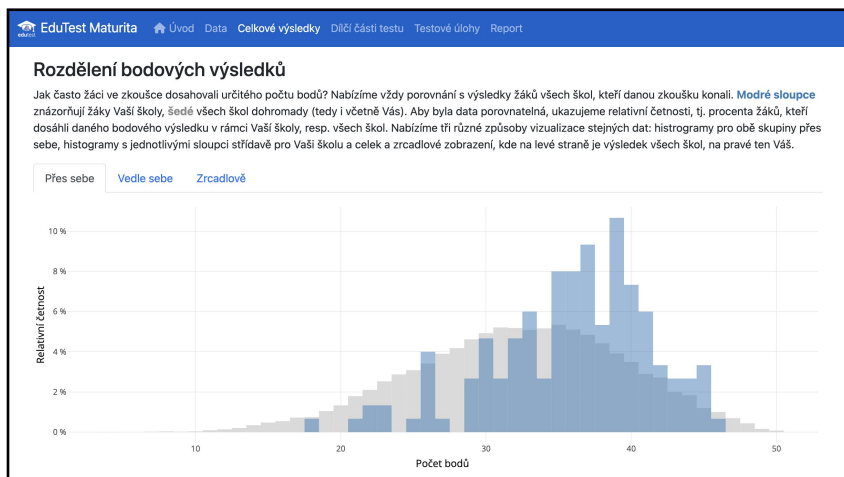
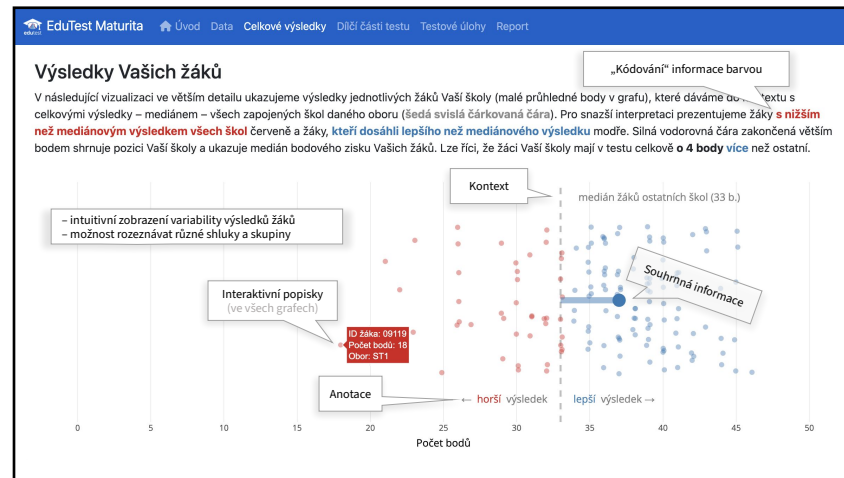
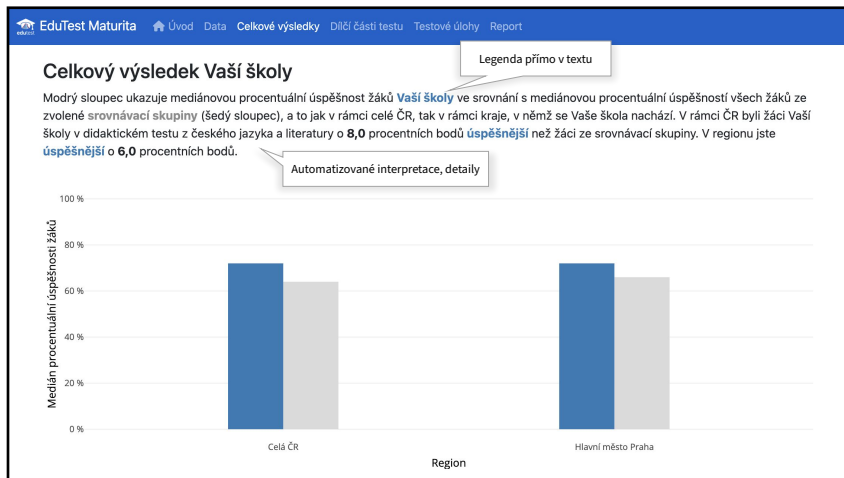
2. Aplikace pro střední školy – maturitní zkoušky

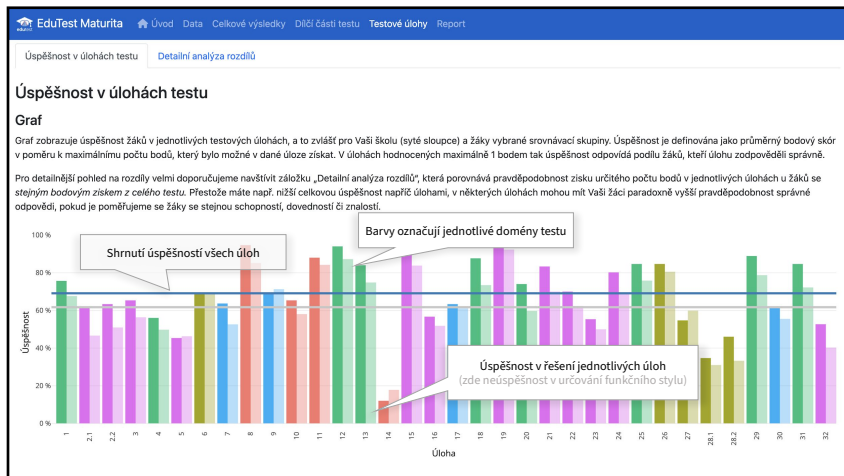
- využití dat školy (speciální export pro školy od CZVV)
- zpětná vazba pro SŠ, lze využít i v rámci přípravy na maturitu edutest.shinyapps.io/maturita_data (pouze ukázková data)

zkouška	žáků	skupina mat. oborů	žáků
Matematika	54 (34 %)	Lyceum	64 (20 %)
Anglický jazyk	115 (72 %)	SOŠ technické	281 (80 %)
Český jazyk a literatura	156 (98 %)		

skupina mat. oborů	žáků	celkový výsledek	žáků	druh termínu	žáků
Lyceum	31 (20 %)	neuspěl(a)	2 (1 %)	řádný	156 (100 %)
SOŠ technické	125 (80 %)	-	6 (4 %)		
		uspěl(a)	148 (95 %)		

prvotní zkouška	žáků	povinná zkouška	žáků	konali zkoušku	žáků
ne	1 (1 %)	ano	156 (100 %)	ne	6 (4 %)
ano	155 (98 %)			ano	150 (96 %)





EduTest Maturita Úvod Data Celkové výsledky Dílčí části testu Testové úlohy Report

Tabulka

Tabulky v záložkách níže zobrazují parametry jednotlivých položek pro Vaši školu a celek. Kromě úspěšnosti je uvedeno také průměrné skóre v položce; to je totožné s úspěšností v případě dichotomických položek (položek hodnocených 0 – nesprávná odpověď / 1 – správná odpověď). Dále je uvedena směrodatná odchylka (standard deviation; SD), minimální a maximální dosažené skóre. Následuje podíl respondentů, kteří dosáhli maximálního možného skóre, tento podíl je opět totožný s odhadem úspěšnosti v případě dichotomických položek.

Vaše škola Všechny školy Možnost řazení, označování atp.

Stejně informace podrobněji v podobě tabulky

Popis	Doména	Úspěšnost %	Průměr %	SD %	Min. bodů	Max. bodů	Podíl max. bodů %
14 charakter textu: určit funkční styl	Charakter a výstavba textu	0,12	0,12	0,33	0	1	0,12
28.1 literární historie: identifikovat literární směr, který je charakterizován ve výchozím textu (naturalismus)	Literatura	0,35	0,35	0,48	0	1	0,35
5 syntax: rozlišit větu jednoduchou x souvětí, identifikovat počet vět v souvětí	Slovo, věta, souvětí	0,45	0,45	0,50	0	1	0,45
28.2 literární historie: identifikovat spisovatele, o němž se píše ve výchozím textu (Zola)	Literatura	0,46	0,46	0,50	0	1	0,46
32 význam slov: posoudit, v kterých kontextech jsou určité dvojice slov antonymy (chut x nechuť; dostatek x nedostatek)	Slovo, věta, souvětí	0,53	0,53	0,50	0	1	0,53
27 literární teorie: identifikovat literární žánr (sonet)	Literatura	0,55	0,55	0,50	0	1	0,55
23 syntax: identifikovat větné členy (podmět, předmět)	Slovo, věta, souvětí	0,55	0,55	0,50	0	1	0,55
4 porozumění textu (práce s definicí – vokalizovaná x nevokalizovaná podoba předložek): najít v úsecích předložku, která může mít vokalizovanou podobu (pod)	Porozumění textu	0,56	0,56	0,50	0	1	0,56
16 syntax: identifikace příslušných určení	Slovo, věta, souvětí	0,57	0,57	0,50	0	1	0,57
2.1 ověřené	Slovo, věta, souvětí	0,61	0,61	0,49	0	1	0,61

EduTest Maturita Úvod Data Celkové výsledky Dílčí části testu Testové úlohy Report

Report

Pro nahraná data a zvolenou zkoušku si můžete vygenerovat a následně vytisknout report, který nabízíme ve formátu PDF.

V současnosti používá report pro srovnání Vaší školy s ostatními hodnoty vybrané v předchozích záložkách. Pokud jste s výběrem skupin v aplikaci nemanipulovali, report použije výchozí hodnoty, tj. srovnávají se žáci všech oborů Vaší školy s žáky ostatních škol odpovídajících oborů.

Sestrojit grafy jen pro různé fugující položky Přizpůsobení

Uložte si a vytiskněte to, co jste viděli v aplikaci

edutest

Český jazyk a literatura – jaro 2022

Výsledky z didaktického testu v rámci maturitní zkoušky

Střední škola Uhltázková, Testová 600/1, Testov 1, příspěvková organizace (IČEDIZO: da3296532.)

Vygenerováno 6. prosince 2022

© 2022
Skupina výpočetní psychometrie,
Ústav informatiky, Akademie věd ČR

Název smyšlené školy pro ukázkou

Obsah

Úvod	3
Celkové výsledky	4
Srovnání úspěšnosti Vaší školy a ostatních	4
Výsledky žáků Vaší školy v kontextu mediánu celku	5
Rozdělení dosažených bodů v testu	6
Výsledky v dílčích částech testu	7
Výsledky v testových úlohách	8
Detailní analýza rozdílů	10

Moduly a aplikace projektu EduTest

1. SIA modul „CERMAT tools“

- načtení volně dostupných položkových dat, specifická položková analýza
- možnost poslat data do hlavní ShinyItemAnalysis a využít její funkcionality shiny.cs.cas.cz/ShinyItemAnalysisCermat

2. Aplikace pro střední školy - maturitní zkoušky

- využití dat školy (speciální export CZVV)
- zpětná vazba pro SŠ, lze využít i v rámci přípravy na maturitu edutest.shinyapps.io/maturita_data (pouze ukázková data)

3. Aplikace pro základní školy – jednotné přijímací zkoušky

- přímé nahrání tzv. školní zprávy (Excel běžně zasílaný CZVV do škol)
- zpětná vazba pro ZŠ, lze využít i v rámci přípravy na JPZ edutest.shinyapps.io/prijimacky/

The screenshot shows the EduTest application interface. At the top, there is a navigation bar with links: EduTest, Úvod, Data, Celkové skóry, Položková analýza, DIF analýza, Report. Below the navigation bar, there are three main sections:

- Nahrání dat:** A section for uploading data. It includes a text box for the file name (JPZ_skolni_zprava_ZS_2l) and a 'Upload complete' button.
- Volba zkoušky:** A section for selecting the exam type. It includes a dropdown menu with options: MAJ, MA, ČJL. Below the dropdown, there is a note: 'Aut. rozpoznání dostupných položkových dat'.
- Náhled dat:** A table showing the results of the data upload. The table has columns for 'žák', 'celkem bodů', and various subject scores (u1 to u5.1). The data is as follows:

žák	celkem bodů	u1	u2.1	u2.2	u3.1	u3.2	u4.1	u4.2	u4.3	u5.1	u5.1
1	50	1	1	1	2	2	1	1	2	2	
2	50	1	1	1	2	2	1	1	2	2	
3	50	1	1	1	2	2	1	1	2	2	
4	49	1	1	1	2	2	1	1	1	2	
5	49	1	1	0	2	2	1	1	2	2	
6	48	1	1	1	2	2	1	1	2	2	
7	48	1	1	1	2	2	1	1	2	2	
8	48	1	1	1	2	2	1	1	2	2	
9	48	1	1	0	2	1	1	1	2	2	
10	48	1	1	1	2	2	1	1	2	2	

At the bottom of the table, it says 'Zobrazují 1 až 10 z celkem 91 záznamů'. There are also navigation buttons: 'Předchozí', '1', '2', '3', '4', '5', '...', '10', 'Další'.

Praktická část

Praktická část: výstupy projektu EduTest

- ? Která prezentovaná aplikace nebo modul jsou z vaší perspektivy nejvíce zajímavé, jak byste je využili?

Zatím představené interaktivní aplikace a moduly:

0. [ShinyItemAnalysis](#)
1. SIA modul [CERMAT tools](#) - pro tvůrce testů
2. [Aplikace pro střední školy](#) s cvičnými daty z maturity
3. [Aplikace pro základní školy](#) (JPZ)

Online materiály a odkazy: cs.cas.cz/comps/workshop

II. Od tradičních metod k modelování odpovědí na položky

CTT, IRT, DIF

Obsah

Od tradiční položkové analýzy po modelování odpovědí na položky

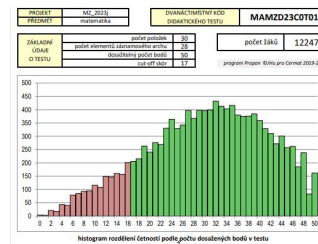
- Tradiční položková analýza
- Regresní modely pro popis položek
- Modely teorie odpovědi na položku (*Item Response Theory, IRT*)
- Odlišné fungování položek (*Differential Item Functioning, DIF*)
- Praktická část



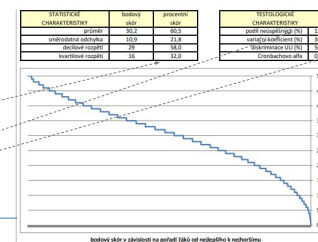
POLOŽKOVÁ ANALÝZA UKÁZKA VÝSTUPŮ

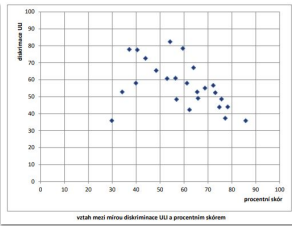
Zpracoval: Centrum pro zjišťování výsledků vzdělávání – Oddělení analýz
ŘÍJEN 2023

ZÁKLADNÍ INFORMACE O TESTU CELKEM



- Četnost dosažených bodů
- Průměrný skór
- Míry variability
- Podíl neúspěšných
- Reliabilita testu





PŘEHLED ÚLOH PRŮMĚRNÝ SKÓR VS. DISKRIMINACE

vtah mezi mírou diskriminace ULI a procentním skórem

průměrný skóre	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
90-100								
80-90								
70-80		9	10	14	7			
60-70			8	21	3			
50-60			1	12	19	13	11	
40-50			24		15	6	13	18
30-40			16		11	2	4	5
20-30								
10-20								

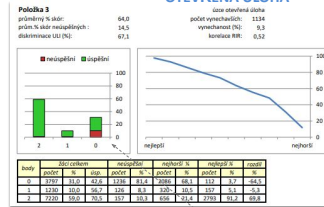
vtah mezi ULI a procentním skórem - uvedeno číslo položky

- Celkový pohled na jednotlivé úlohy v kombinaci % skóre vs. diskriminace

42

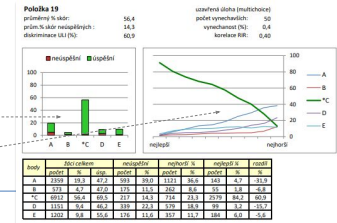
ÚDAJE O TESTOVÝCH POLOŽKÁCH

OTEVŘENÁ ÚLOHA



- Průměrný skóre
- Diskriminace ULI
- Korelace RIR
- Vynechanost úlohy

UZAVŘENÁ ÚLOHA



- Četnost dosažených bodů, resp. četnost volby alternativ
- Průměrný % skóre, resp. míra volby alternativ podle výsledku v testu

43

ShinyItemAnalysis Test and item analysis

Data Scores Validity Reliability Item analysis Regression IRT models DIF/Fairness Modules Reports

Data CTT IRT

Data

Item data
 Browse... mz_2022_spring_ma_>
 Upload complete

Options
 Include only the cases with "termin = fádny", "prvomaturant", and "zkouska_p"
 Display preview of the data

Item metadata
 Browse... mz_2022_spring_ma_>
 Upload complete

Use data in the main app
 Here you can pass your data you have just loaded in this module to the main SIA app. You can chose which column of the uploaded dataset will be considered a group or criterion variable.

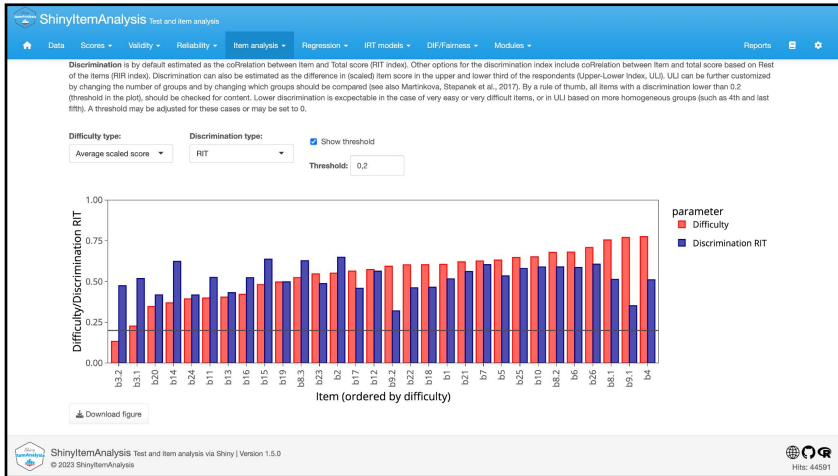
Group
 none

Criterion
 none

Pass data to SIA

Modules menu:
 Description
 About the modules
 Modules
 CERMAT tools
 CEFR level analysis
 Computerized Adaptive Tests

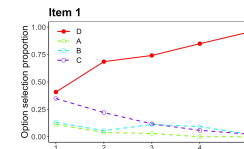




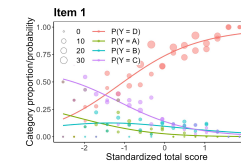
Od tradiční položkové analýzy k IRT

- Tradiční položková analýza využívá podílů, procent a korelací k popisu vlastností položky
- Regresní modely modelují pravděpodobnost dané (např. správné) odpovědi hladkou křivkou, odhadují se parametry křivky (obtížnost, diskriminace)
- Modely teorie odpovědi na položku (IRT) navíc předpokládají, že úroveň znalosti respondentů je neznámá a je třeba ji odhadnout společně s parametry položek

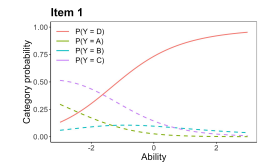
Tradiční položková analýza



Regresní modely



Modely IRT



VÝCHOZÍ TEXT K ÚLOZE 3

Vlak má tři vagony, všechny se stejným počtem míst. V každém vagonu je o 20 míst k stání více než k sezení.

Při odjezdu z Roztok byl vlak zaplněn přesně do poloviny své kapacity. V prvním a posledním vagonu byla všechna místa k sezení obsazena, ale ve druhém vagonu zůstalo 25 % míst k sezení volných.

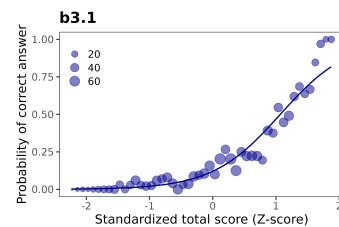
(Kapacita vlaku je součet počtu všech míst k stání a sezení. Každý cestující obsadí buď jedno místo k stání, nebo jedno místo k sezení.)

max. 2 body

- 3 Počet míst k sezení v jednom vagonu označme n .
Vyjádřete v závislosti na veličině n počet všech cestujících, kteří při odjezdu z Roztok
- 3.1 byli ve vlaku;
- 3.2 ve vlaku stáli.

Obtížné položky, dobře diskriminují Položka b3.1

$a = 1.86 \dots$ diskriminační schopnost (sklon křivky v inflexním bodě)
 $b = 1.07 \dots$ obtížnost (poloha inflexního bodu křivky)
Celkové skóre 1SD nad průměrem jsou potřeba k tomu aby položka byla zodpovězena správně s pravděpodobností 0,5



- 4 Pro $a \in \mathbb{R} \setminus \{-3; 0; 3\}$ zjednodušte:

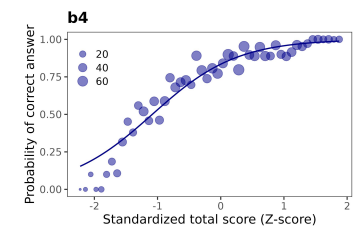
$$\frac{1 + \frac{3}{a}}{\frac{a^2}{3} - 3}$$

V záznamovém archu uveďte celý postup řešení.

max. 2 body

Lehká položka, dobře diskriminuje

$a = 1.49 \dots$ diskriminační schopnost (sklon křivky v inflexním bodě)
 $b = -1.08 \dots$ obtížnost (poloha inflexního bodu křivky)
Celkové skóre 1SD pod průměrem stačí k tomu aby položka byla zodpovězena správně s pravděpodobností 0,5

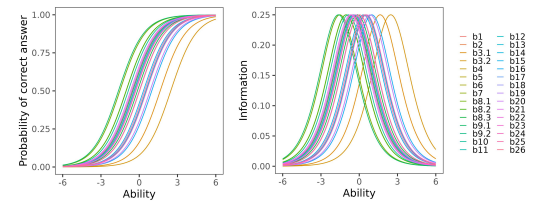


Modely teorie odpovědi na položku (IRT)

- Odhad znalosti respondentů společně s parametry položek
- Nutnost dostatečného množství pozorování (respondentů)
- 1–4PL modely pro binární data, modely pro ordinální i nominální položky

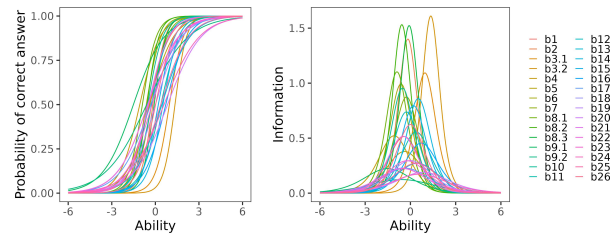
Raschův/1PL IRT model

- Různé obtížnosti položek
- Model předpokládá stejný sklon pro všechny položky
- Pro 1PL IRT model je odhad obtížnosti ekvivalentní s celkovým skóre (pokud jsou položky binární)



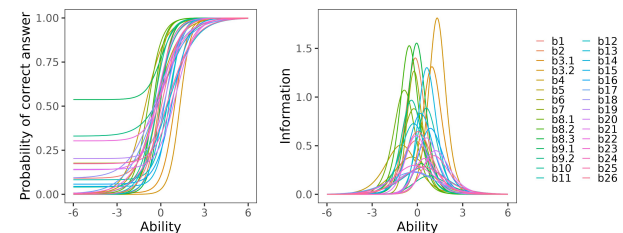
2PL IRT model

- Různé obtížnosti i sklony pro položky
- Každá položka je různě informativní



3PL IRT model

- Různé obtížnosti i sklony pro položky
- Modeluje možné hádání správné odpovědi (pseudo-guessing)



22 V geometrické posloupnosti platí:

$$a_2 = \sqrt[3]{3}$$

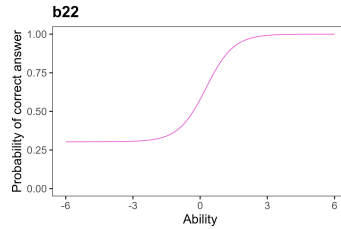
$$a_3 = -\sqrt[3]{9}$$

Jaká je hodnota součtu $a_1 + a_4$?

- A) 2
- B) 1
- C) 0
- D) -1
- E) jiná hodnota

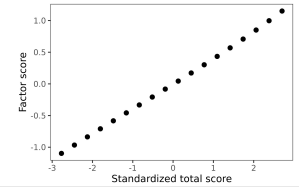
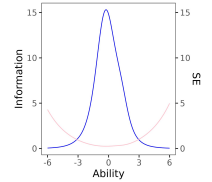
Položka s možným hádáním správné odpovědi

Odhadnuté parametry položky:
 $a = 1.63$... diskriminační schopnost (sklon křivky v inflexním bodě)
 $b = 0.26$... obtížnost (poloha inflexního bodu křivky)
 $c = 0.30$... uhádnutelnost (dolní asymptota křivky)



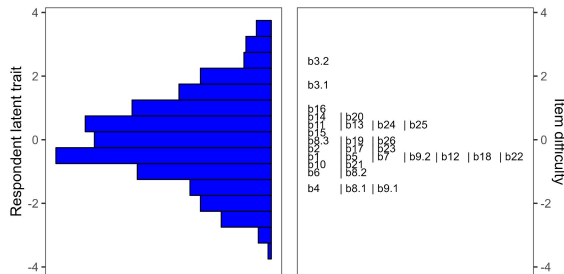
Využití modelů teorie odpovědi na položku

- IRT modely poskytnou konfidenční interval pro odhad znalosti žáka
- Reliabilita testu není konstantní, ale je funkcí znalosti žáka
- Pro IPL model je odhadnutá znalost ekvivalentní s celkovým skóre
- Odhad znalosti v případě více verzí testu
- Informační funkce jednotlivých položek
 - Lze dále využít: počítačové adaptivní testování, vyvažování (equating), ...



Využití modelů teorie odpovědi na položku

- Interpretace obtížnosti položek vůči znalostem žáků (Wright Map)



ShinyItemAnalysis Test and Item analysis

Data CTT IRT

IRT

By pressing the button below, you can fit an IRT model based on raw, unscored responses for "I*" items. Unlike the IRT section of the main app, this module allows you to define individual IRT "models" for each item, as specified in the metadata file (which you must upload before fitting the model). Below, you can review the model that is going to be fitted:

Model	Items
3PL	k15.1, k15.2, and k15.3
gpcm	k1, k2, k3, k4, k5.1, k5.2, k5.3, k6, k7, k8.1, k8.2, k9, k10, k11.1, k11.2, k12.1, k12.2, k13, and k14
nominal	k16, k17, k18, k19, k20, k21, k22, k23, k24, k25.1, and k25.2

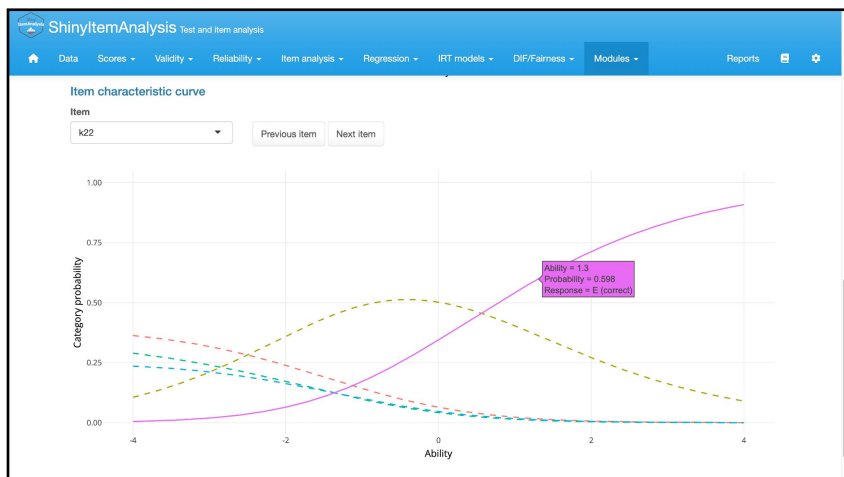
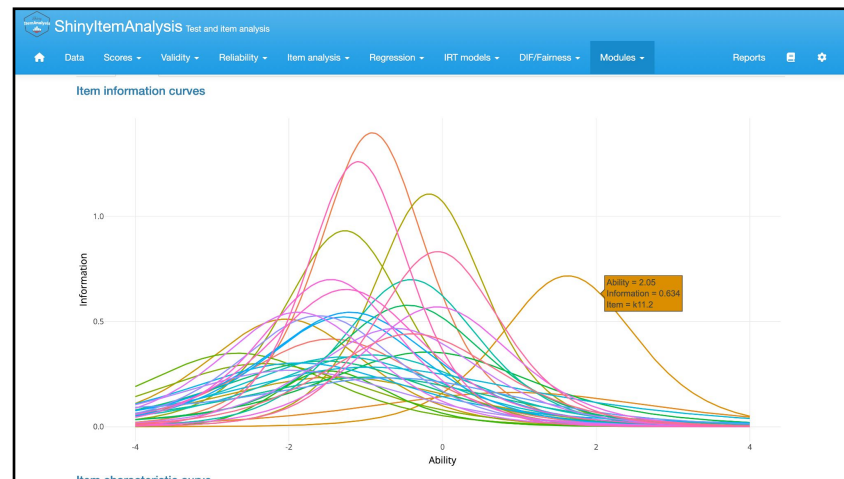
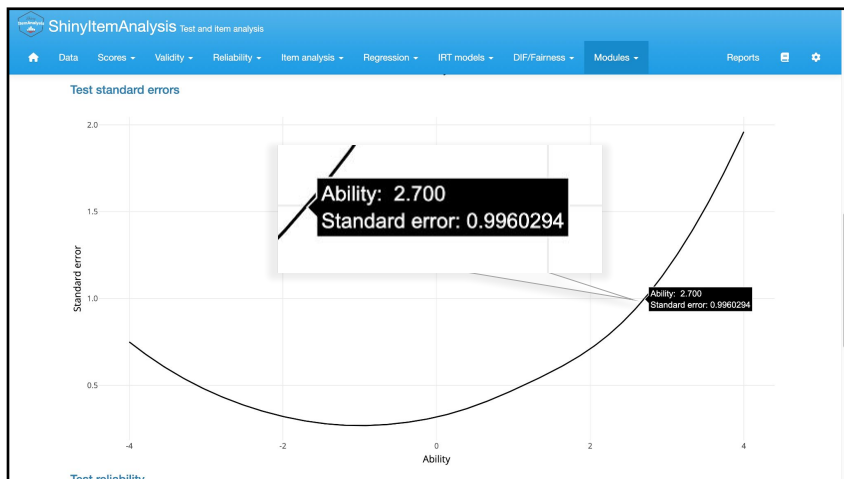
Fit the model

Results

Test Items

Test information curve

Fitting the model This may take a while...



Odlíšné fungování položek

DIF, z angl. Differential Item Functioning
= respondenti z různých skupin se stejnou znalostí mají odlišné pravděpodobnosti správné odpovědi na položku

The three plots show the probability of a correct response (Pravděpodobnost správné odpovědi) versus latent knowledge (Latentní znalost) for two groups: Referenční (blue solid line) and Fokální (yellow dashed line). The first plot shows no DIF (parallel curves). The second plot shows positive DIF (Fokální curve is higher). The third plot shows negative DIF (Fokální curve is lower).

Odlíšné fungování položek

- Může poukazovat na neférovost položek (položka měří něco, co nemá)
- Ne všechny položky, které fungují odlišně, jsou neférové (položka měří, co má; rozdíl není způsoben neférovostí)
- DIF analýza
 - přináší hlubší poznání o fungování položek, miskonceptech, senzitivitě na instrukci
 - pomáhá informovat tvůrce testu o charakteristikách položek
 - pomáhá informovat o silných a slabých stránkách skupin

Odlíšné fungování položek – příklady

DIF položka v testu SAT z 60. let – označena jako neférová

„Běžec je k maratonu jako“

- A. vyslanec k velvyslanectví
- B. mučedník k masakru
- C. veslař k regatě**
- D. rozhodčí k turnaji
- E. koně ke stájím

častěji správně odpověděli bílí než afroameričtí studenti se stejnými celkovými znalostmi

- ? **Co způsobilo rozdíly mezi skupinami v odpovídání na tuto položku?**
- ? **Byl rozdíl způsobený znalostí něčeho, co test neměl testovat? – ANO**

Odlíšné fungování položek – příklady

DIF položka z přijímacího testu na lékařské fakulty – NEoznačena jako neférová

„Nedostatek vitamínu D v dětství může způsobit“

- A. křivici**
- B. kurděje
- C. nanismus
- D. mentální retardaci

častěji správně odpověděly studentky než studenti se stejnými celkovými znalostmi

- ? **Co způsobilo rozdíly mezi skupinami v odpovídání na tuto položku?**
- ? **Byl rozdíl způsobený znalostí něčeho, co test neměl testovat? – NE**

Příklady využití DIF analýzy

- Pro tvůrce testů
 - ověření férovosti, vybalancování položek napříč skupinami
- Pro tvůrce státních politik
 - porozumění vzdělávacím výsledkům různých skupin
- Pro školy a vyučující
 - odhalení silných a slabých stránek

DIF analýza: Příklad 1

- Využití DIF analýzy pro tvůrce testů, pro vyučující
 - JPZ 2023, matematika
 - UA vs. CZ žáci
-
- Položky fungují odlišně, ale nejsou neférové (testují znalost matematiky)

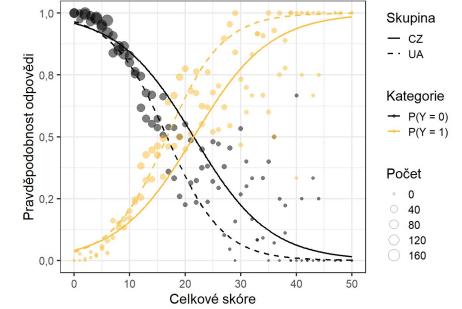
4.1 Upravte a rozložte na součin vytknutím:

$$2 \cdot (x^2 - x) + x =$$

Ukrajínští žáci mají **větší** pravděpodobnost správné odpovědi při stejném celkovém skóre

Možné důvody:

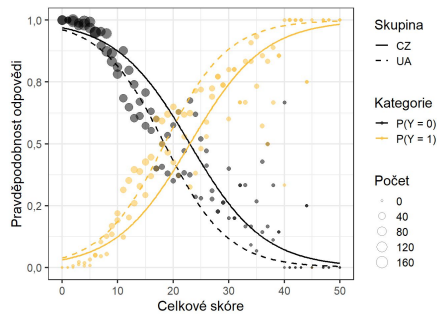
- jiné kurikulum
- málo textu



4.2 Umocněte a zjednodušte (výsledný výraz nesmí obsahovat závorky):

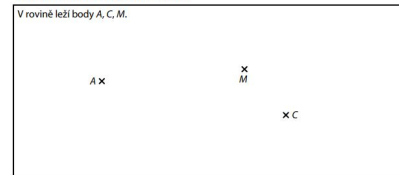
$$\left(\frac{2}{3}a - 3\right)^2 =$$

Ukrajínští žáci mají **větší** pravděpodobnost správné odpovědi při stejném celkovém skóre



VÝCHOZÍ TEXT A OBRÁZEK K ÚLOZE 9

V rovině leží body A, C, M.



(C20)

max. 2 body

9 Body A, C jsou vrcholy obdélníku ABCD. Bod M leží na úhlopříčce BD tohoto obdélníku.

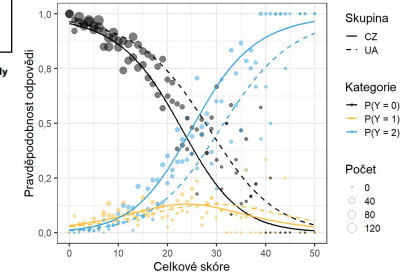
Sestrojte vrcholy B, D obdélníku ABCD, označte je písmeny a obdélník narysujte.

V záznamovém archu obtněte celou konstrukci propisovací tužkou (čáry i písmena).

Ukrajínští žáci mají **menší** pravděpodobnost zisku plného počtu bodů

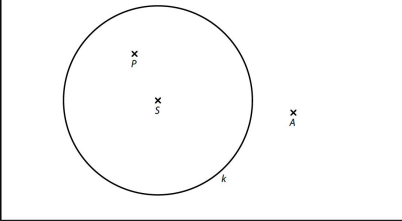
Možné důvody:

- jiné kurikulum
- více textu



VÝCHOZÍ TEXT A OBRÁZEK K ÚLOZE 10

V rovině leží body A, P a kružnice k se středem S.



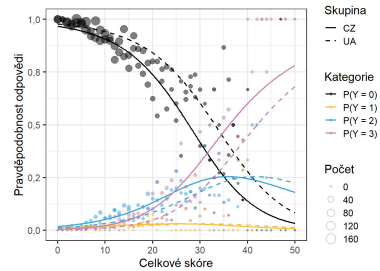
(CZVI)

10 Bod A je vrchol **rovnoramenného** trojúhelníku ABC, jehož **základna** leží na přímce AP. **max. 3 body**

Vrcholy B, C tohoto trojúhelníku leží na kružnici k.
Sestrojte vrcholy B, C trojúhelníku ABC, **označte** je písmeny a trojúhelník **narysujte**. Najděte všechna řešení.
 V **záznamovém archu** obtáhněte celou konstrukci **propisovací tužkou** (čáry i písmena).

Ukrajínští žáci mají **menší** pravděpodobnost zisku plného počtu bodů

Možné důvody:
 - jiné kurikulum
 - více textu, specifické pojmy



DIF analýza: Příklad 2

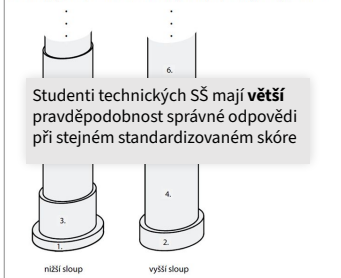
- Využití DIF analýzy pro tvůrce státních politik
- Maturitní test, matematika, jaro 2019
- Gymnázia vs. technické střední školy (ST1, ST2)
- Položky fungují odlišně, ale nejsou neférové (testují znalost matematiky)
- DIF analýza poukazuje na silné stránky různých typů škol

VÝCHOZÍ TEXT A OBRÁZEK K ÚLOZE 9

V Kocourkově navrhli nereálný plán stavby dvou sloupů sahajících do nebe. Na stavbu se má použít celkem 20 válců. Jednotlivé válce jsou podle výšky označeny pořadovými čísly od 1 do 20.

Například je 1. válec s výškou 1 m, 2. válec má výšku 2 m a rovněž každý další válec je dvojnásobek výšky předcházejícího válece s pořadovým číslem o 1 nižším. (Tedy 3. válec má výšku 4 m, 4. válec 8 m atd.)

Nižší sloup bude postaven ze všech válců označených lichými pořadovými čísly od 1 do 19, vyšší sloup ze všech válců označených sudými pořadovými čísly od 2 do 20.



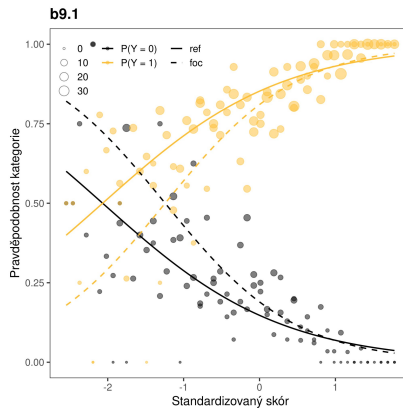
(CZVI)

9 Určete v metrech

9.1 výšku 20. válce.

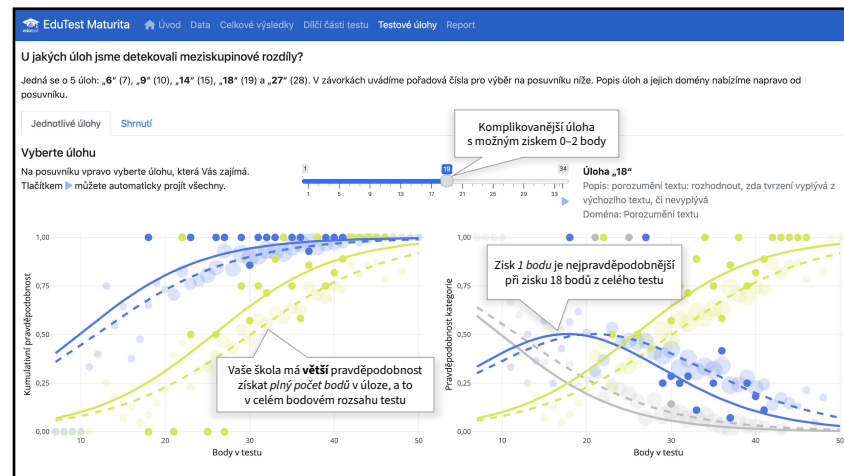
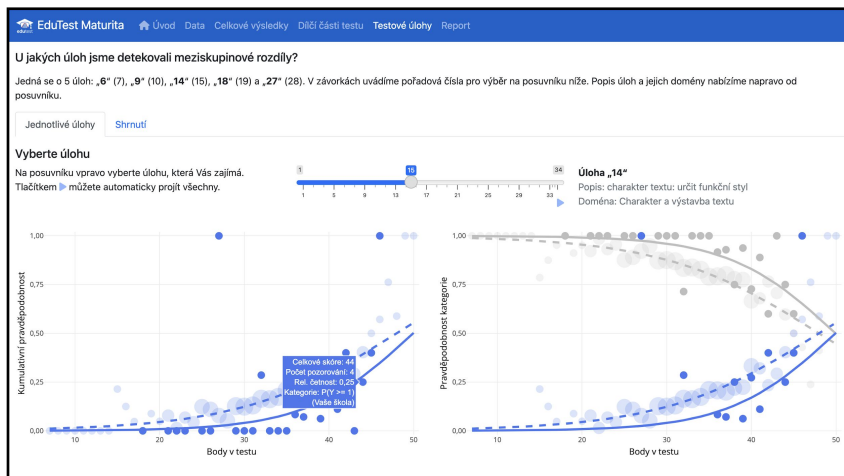
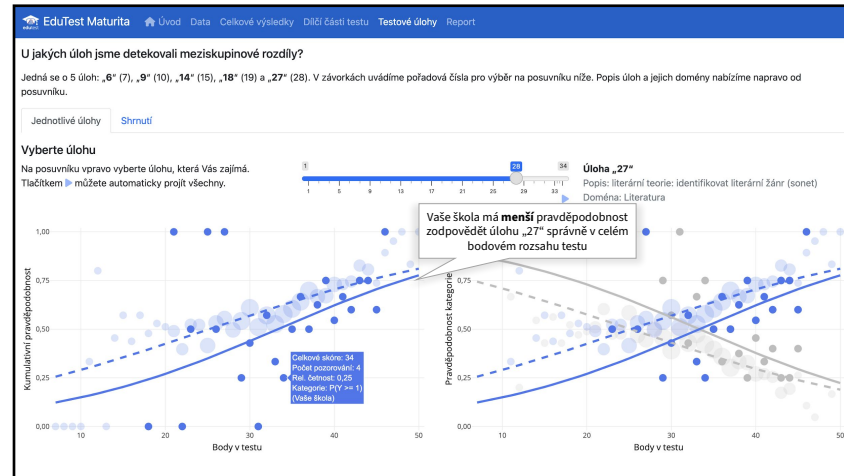
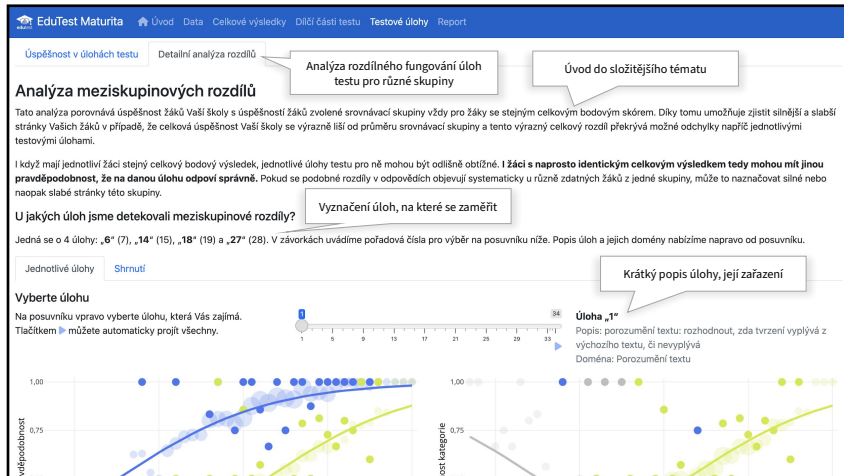
max. 2 body

Studenti technických SŠ mají **větší** pravděpodobnost správné odpovědi při stejném standardizovaném skóre



DIF analýza: Příklad 3

- Využití DIF analýzy pro školy
- Maturitní test, český jazyk a literatura, jaro 2022
- Vybraná škola vs. jiné školy stejného typu
- Položky fungují odlišně, ale nejsou neférové (testují znalost českého jazyka a literatury)
- DIF analýza poukazuje na silné stránky dané školy



EduTest Maturita Úvod Data Celkové výsledky Dílčí části testů Testové úlohy Report

Jednotlivé úlohy Shrnoutí Sekce pro zvidavé

Testové statistiky a odhady parametrů

V první části tabulky naleznete číslo, název úlohy, testovou statistiku, p-hodnotu a značku stat. významnosti (viz poznámky pod tabulkou). V pravé části poskytujeme odhady všech parametrů modelu, který byl pro DIF analýzu použit.

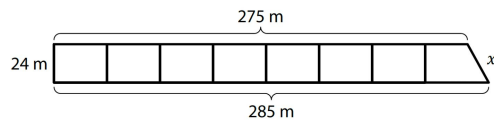
úloha ^{1%}	LR (χ^2) ^{2%}	p-hodnota ^{3%}	sig. ^{4%}	b_1 ^{5%}	SE(b_1) ^{6%}	b_2 ^{7%}	SE(b_2) ^{8%}	a ^{9%}	SE(a) ^{10%}	b_3 ^{11%}	SE(b_3) ^{12%}	b_4 ^{13%}	SE(b_4) ^{14%}	b_{DIF} ^{15%}	SE(b_{DIF}) ^{16%}	b_{DIF} ^{17%}	SE(b_{DIF}) ^{18%}
7	6	13,301	0,001 **	5,218	0,764	19,191	0,348	0,135	0,003	28,798	0,210	36,014	0,232	10,396	1,925	6,943	0
19	18	12,005	0,002 **	11,781	0,775	30,054	0,304	0,121	0,005	0	0	0	0	-1,984	5,469	-4,360	0
15	14	11,799	0,003 **	47,996	0,972	0	0	0,109	0,008	0	0	0	0	1,949	4,728	0	0
28	27	6,653	0,036 *	25,205	0,851	0	0	0,059	0,005	0	0	0	0	8,101	2,620	0	0
10	9	6,468	0,039 *	20,916	0,775	0	0	0,090	0,006	0	0	0	0	5,417	3,858	0	0
32	30	5,158	0,076 .	20,235	0,266	25,311	0,195	0,168	0,003	33,878	0,183	42,211	0,259	0	0	0	0
22	21	5,115	0,078 .	23,424	0,586	0	0	0,099	0,006	0	0	0	0	0	0	0	0
16	15	4,488	0,106	18,487	0,670	0	0	0,134	0,007	0	0	0	0	0	0	0	0
13	12	4,246	0,120	13,403	1,127	0	0	0,114	0,008	0	0	0	0	0	0	0	0
18	17	4,216	0,121	27,354	0,456	0	0	0,096	0,006	0	0	0	0	0	0	0	0
30	28.2	3,970	0,137	38,373	0,364	0	0	0,130	0,006	0	0	0	0	0	0	0	0
26	25	3,858	0,145	20,588	0,658	0	0	0,111	0,006	0	0	0	0	0	0	0	0
21	20	3,720	0,156	28,463	0,395	0	0	0,105	0,006	0	0	0	0	0	0	0	0
9	8	3,334	0,189	16,653	0,791	0	0	0,132	0,008	0	0	0	0	0	0	0	0

DIF analýza: Příklad 4

- didaktický test z matematiky, maturitní zkouška jaro 2022
- chlapci vs. dívky
- položky fungují odlišně, ale nejsou neférové (testují znalost matematiky)

VÝCHOZÍ TEXT A OBRÁZEK K ÚLOZE 5

Pozemek má tvar pravoúhlého lichoběžníku s výškou 24 m a základnami délek 285 m a 275 m. Pozemek je rozdělen na 8 parcel o stejné výměře. Prvních sedm parcel tvoří shodné obdélníky, poslední parcela má tvar pravoúhlého lichoběžníku.



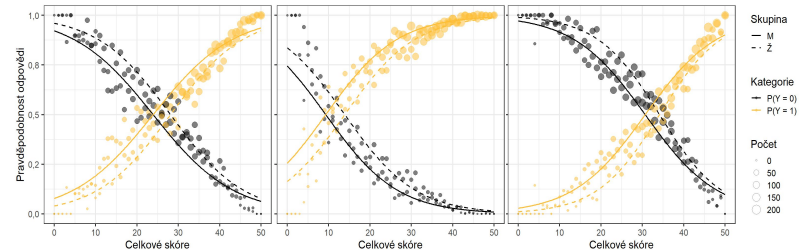
(CZVV)

max. 3 body

5 Vypočítejte

- 5.1 v m² výměru jedné parcely,
- 5.2 v m chybějící délku x strany pozemku,
- 5.3 v m obvod poslední parcely.

„Podpoložky“ b5.1, b5.2 a b5.3



Dívky odpovídají správně s **menší** pravděpodobností než chlapci.
Možné důvody: nároky na prostorovou orientaci?

7 V oboru R řešte:

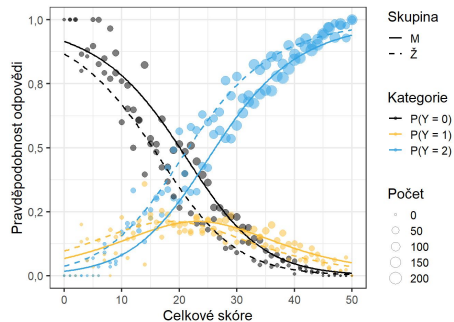
$$\frac{x-2}{x+2} + \frac{3}{x} + \frac{16}{x^2+2x} = \frac{x}{x+2}$$

V záznamovém archu uveďte celý postup řešení.

max. 2 body

Dívky mají větší pravděpodobnost získu plného počtu bodů.

Možné důvody: nároky na pečlivost/preciznost?



Praktická část

Praktická část a diskuse

- V aplikaci [ShinyItemAnalysis](#) (záložka Data) vyberte dataset CZmaturaS
 - (MZ z matematiky, jaro 2019; skupina: gymnázia vs. ostatní; vzorek 2 000 žáků)
 - Vyzkoušejte si dostupné analýzy v záložkách: Item Analysis, Regression, IRT models, DIF/Fairness
- Prohlédněte si modul CERMAT tools (záložka Modules)
 - Nahrajte položková data M 2022 a příslušná metadata
 - Vyzkoušejte si dostupné analýzy

- ? Které analýzy využíváte při tvorbě a vyhodnocování testů?
- ? Které analýzy považujete za klíčové?

Online materiály a odkazy: cs.cas.cz/comps/workshop

III. Další témata

Vyvažování verzí testu,
odhadování obtížnosti z analýzy textu,
počítačové adaptivní testování (CAT)

Vyvažování verzí testu

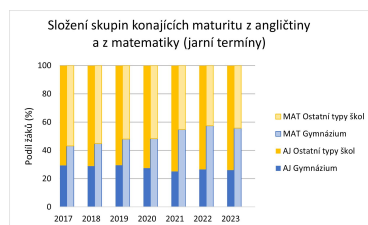
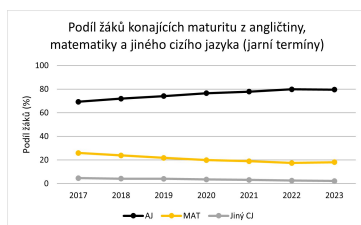
- Dle Standardů pro pedagogické a psychologické testování (AERA, APA, NCME, 2014) je třeba zajistit srovnatelnost škál všech paralelních forem testu, aby testové skóry byly tzv. „zaměnitelné“
- Paralelní formy testu – určené k měření stejného konstruktů, vytvořené podle stejné specifikace, klíčová je zde expertní role tvůrců testu
- Test equating – skupina statistických metod používaných k vyvažování (vyrovnávání) nominálně paralelních forem testu, které ve skutečnosti nemusí být perfektně paralelní
- Cílem je odlišit rozdíly v obtížnosti forem testu od rozdílů ve znalostech respondentů

Vyvažování verzí testu

- 2 hlavní přístupy k equatingu:
 - **stejně (ekvivalentní) skupiny** respondentů pro všechny formy testu – pozorované rozdíly jsou způsobeny rozdílnou obtížností testových forem
 - **stejně položky** do **všech** forem testu (tzv. *anchor test*) – pozorované rozdíly ve výsledcích *anchor testu* jsou způsobeny rozdílnými znalostmi skupin respondentů
 - zlatý standard v situacích, kdy nelze předpokládat ekvivalenci skupin

Vyvažování verzí testu – případ maturity

- možnost volby cizího jazyka nebo matematiky
 - složení žáků se může meziročně lišit



Vyvažování verzí testu – případ maturity

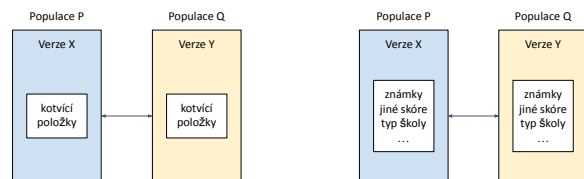
- jarní a podzimní termín
 - složení žáků se značně liší pro všechny předměty

Termín	ČJL	AJ	MAT
jaro 2017	32,14 (8,47)	71,40 (17,58)	21,89 (13,74)
podzim 2017	24,06 (6,56)	49,14 (18,25)	13,33 (8,77)
jaro 2018	31,81 (8,37)	70,22 (17,51)	21,15 (13,00)
podzim 2018	24,80 (5,70)	47,29 (18,59)	13,79 (8,22)
jaro 2019	31,64 (8,37)	77,42 (17,49)	24,15 (13,21)
podzim 2019	22,91 (6,20)	52,46 (20,06)	11,94 (7,84)

Pozn. Maximální počet bodů: ČJL = 50, AJ = 95, MAT = 50;
hranice úspěšnosti: ČJL = 22, AJ = 42, MAT = 17

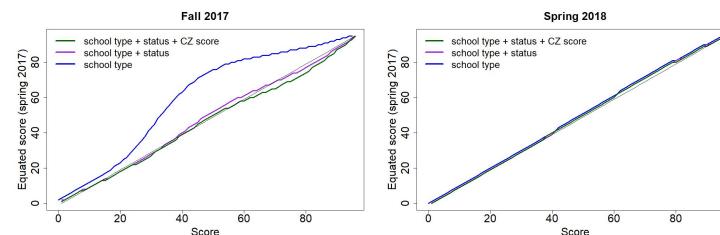
Vyvažování verzí testu – případ maturity

- Skupiny nejsou ekvivalentní a testy neobsahují společné položky
- Společné položky lze využít v pilotáži, ale znění se může změnit, navíc může být složité zajistit reprezentativní vzorek (malé cizí jazyky)
- Místo společných položek lze využít kovariáty - informace o respondentech, které mohou vysvětlit rozdíl ve znalostech mezi skupinami (př. SweSAT)



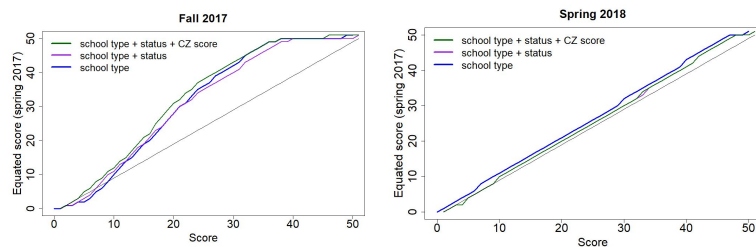
Vyvažování maturitních testů – anglický jazyk

- jaro 2017 vs. podzim 2017 a jaro 2018, různé kovariáty



Vyvažování maturitních testů – matematika

- jaro 2017 vs. podzim 2017 a jaro 2018, různé kovariáty



Vyvažování maturitních testů – závěry

- Při vyvažování testů pomocí kovariát je důležité zahrnout všechny relevantní kovariáty
- Při slabé korelaci kovariát s testovým skórem nejsme schopni s jejich pomocí dostatečně dobře vysvětlit rozdíly mezi skupinami – testové verze se mohou jevit různě obtížné, i když jsou fakticky srovnatelné
- Covid-19???
- Pokud chceme zaručit srovnatelnost testů, bez kotvících položek se do budoucna neobejdeme

Textová analýza

- Navazujeme na teoretický výzkum v oblasti využití machine learning a textové analýzy pro odhad obtížnosti položek (Štěpánek et al., 2023)
- Vycházíme ze srovnání textových charakteristik již administrovaných testových položek a jejich obtížnosti odhadnuté z žákovských odpovědí
- Pomocí machine learning algoritmů se snažíme
 - určit, které textové charakteristiky nejlépe predikují obtížnost položky
 - predikovat obtížnost nových položek
- Využití: doplnění a zpřesnění expertních odhadů tvůrců na nových položkách
- Položky na porozumění textu v cizím jazyce (části 5 a 7 maturitních testů)

Štěpánek, L., Dlouhá, J., & Martinková, P. (2023). Item difficulty prediction using item text features: Comparison of predictive performance across machine-learning algorithms. *Mathematics*, 11(19), 4104. <https://doi.org/10.3390/math11194104>

Textová analýza

- 4 typy textových charakteristik
 - **počty** (vět, slov, slabik, znaků) a odvozené ukazatele (např. průměrná délka slov)
 - **běžnost/vzácnost slov** v textu s oporou o frekvenční slovníky a CEFR úroveň
 - **indexy čitelnosti textu** – většinou vycházejí z údajů o počtech a frekvencích
 - **lexikální podobnost** různých částí testové položky (např. výchozí text vs. správná odpověď, výchozí text vs. distraktory)
- Nejsilnější prediktory obtížnosti: celkový počet znaků, SD délky slova, čitelnost textu, podobnost mezi výchozím textem a distraktory
- V rámci SIA modulů ukázka určení CEFR úrovně jednotlivých slov a výpočtu základních textových charakteristik

ShinyItemAnalysis Test and item analysis

Data Scores Validity Reliability Item analysis Regression IRT models DIF/Fairness Modules Reports

Text analysis

Item passage:

In 2003, a 106-year-old woman from Os in Norway received an offer for free bus rides to school. Town officials sent it to her attend school in autumn 2003. Ingeborg Thuen, born in 1897, is the oldest citizen of Os and actually started school in 1903 the town hall of Os read the 97 of her birth year as 1997 and added her to the list of children starting the first grade in autumn she laughed and said. Free rides are a very good idea, although now I live near the school. When I started school back then

Question:

Why was the woman offered free bus rides? She was offered free bus rides:

Correct option:

Because of a computer mistake.

Incorrect option 1:

Because she was the oldest person in the town.

Incorrect option 2:

Because she wanted to start school.

Incorrect option 3:

Because of her hour-long walk to school every morning.

Analyze

ShinyItemAnalysis Test and item analysis

Data Scores Validity Reliability Item analysis Regression IRT models DIF/Fairness Modules Reports

CEFR level analysis

Item passage:

In 2003, a 106-year-old woman from Os in Norway received an offer for free bus rides to school. Town officials sent it to her attend school in autumn 2003. Ingeborg Thuen, born in 1897, is the oldest citizen of Os and actually started school in 1903 when she was six years old. Computers in the town hall of Os read the 97 of her birth year as 1997 and added her to the list of children starting the first grade in autumn 2003. When Ingeborg received the offer, she laughed and said, Free rides are a very good idea, although now I live near the school. When I started school back then, I had to walk for an hour every morning, which was really hard. However, as I can already read, write and count, I will skip school this time, she joked.

Question:

Why was the woman offered free bus rides? She was offered free bus rides:

Correct option:

Because of a computer mistake.

Incorrect option 1:

Because she was the oldest person in the town.

Incorrect option 2:

Because she wanted to start school.

Incorrect option 3:

Because of her hour-long walk to school every morning.

CEFR legend: A1 A2 B1 B2 C1 C2

Item features

Item feature	Combined text	Item passage	Question	Correct option	Incorrect option 1	Incorrect option 2
Number of characters	667.0	502.0	49.0	22.0	26.0	22.0

ShinyItemAnalysis Test and Item analysis

Data Scores Validity Reliability Item analysis Regression IRT models DIF/Fairness Modules Reports

Item features

Item feature	Combined text	Item passage	Question	Correct option	Incorrect option 1	Incorrect option 2
Number of characters	667.0	502.0	49.0	22.0	26.0	27.0
SD of word length (chars)	1.8	1.8	1.5	0.6	1.6	1.5
Longest word length (chars)	12.0	12.0	7.0	8.0	7.0	7.0
Text readability - FOG index	8.9	11.2	2.8	10.0	3.6	2.4
Text readability - Dale-Chall index	6.1	6.7	0.3	7.0	0.4	0.3
Text readability - Traenkle-Bailer index	-323.9	-383.3	-191.7	-287.4	-228.0	-256.2
Text readability - SMOG index	8.8	10.0	3.1	8.8	3.1	3.1
Number of tokens (words, numbers, interpunction)	133.0	102.0	11.0	3.0	5.0	5.0
Number of multi-syllable words (>= 4)	0.8	1.0	0.0	0.0	0.0	0.0
Number of monosyllabic words	59.4	61.8	72.7	0.0	40.0	60.0

Please see the article below for more details and explanations of each item feature. Excluded words (i.e. "stopwords"): I, a, about, an, are, as, at, be, by, com, for, from, how, in, is, it, of, on, or, that, the, this, to, was, what, when, where, who, will, with, www.

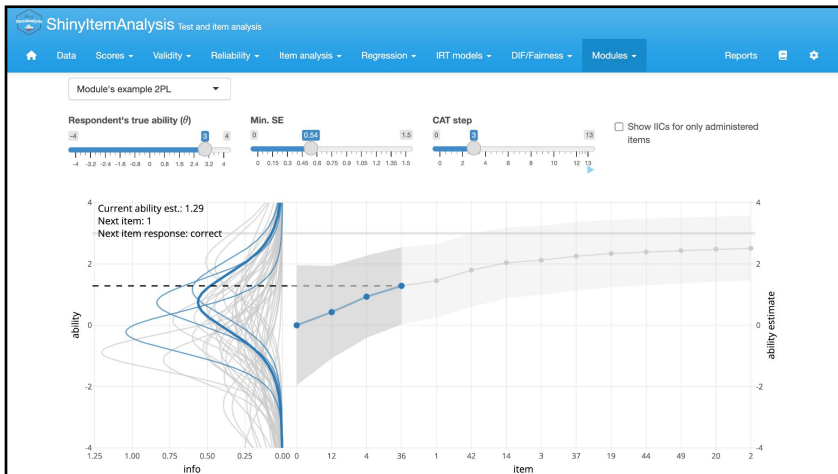
References

CEFR levels based on The Oxford 3000™ by CEFR level and The Oxford 5000™ by CEFR level

Stěpánek, L., Dlouhá, J., & Martinková, P. (2023). Item Difficulty Prediction Using Item Text Features: Comparison of Predictive Performance across Machine-Learning Algorithms. Mathematics, 11(19), 4104. <https://doi.org/10.3390/math11194104>

Počítačové adaptivní testování

- Využití informačních funkcí položek k adaptivní administraci
- Položka je žákovi zadaná dle aktuálního odhadu znalosti (a znalosti parametrů položky)
- SIA modul „Computerized Adaptive Tests“
 - Tzv. post-hoc analýza, vhodná pro nastavení parametrů CATu
 - Simulované odpovědi na základě latentní znalosti a IRT modelu
- Adaptivní verze maturitního testu z matematiky
 - Využití např. pro cvičné testy (viz procvicprijimacky.cermat.cz)



9

Lichoběžník $ABCD$ je rozštěn úhlopříčkou na dva posobné trojúhelníky ABD a BDC . V trojúhelníku BCD jsou vyznačeny dvě další úhlopříčky CH a BI .

Pat: $|AB| = 5,6\text{cm}$, $|BD| = 6,4\text{cm}$, $|CD| = 8\text{cm}$.

Rozhodněte, zda je pravdivé (A), či nepravdivé (N).

$|AB| : |BD| = |BD| : |CD|$

ANO

Obvod trojúhelníku BCD je 1,25krát větší než obvod trojúhelníku ABD .

NE

$|AB| = 5,13\text{cm}$

$|BC| = 7\text{cm}$

ANO

Odešlat

Atuální B: 1.1008, Aktuální SE: 0.42470

Praktická část

Praktická část a diskuse

- Vyzkoušejte si modul k textové analýze
 - Využijte zadání testů z angličtiny (úlohy z části 5) v online materiálech
- Vyzkoušejte si adaptivní test z matematiky
 - Modul adaptivní testování
 - Ukázková aplikace s položkami z maturitních testů
 - Vyzkoušejte různé kombinace správných a nesprávných odpovědí a sledujte, jak se mění odhad latentní znalosti

Online materiály a odkazy: cs.cas.cz/comps/workshop

- ? Řešíte v praxi vyvažování testů? Jak?
- ? Jak stanovujete obtížnost položek? Využíváte analýzu textu?
- ? Využíváte v praxi adaptivní testování?

Diskuse a závěr

Od klasické testové teorie (CTT)
po modely teorie odpovědi na položku (IRT)

Diskuse

- ? Co vše by mělo být součástí standardních výstupů z analýzy znalostních testů (např. pro odbornou komisi MŠMT)?
- ? Jaké výstupy by bylo vhodné zveřejňovat a proč?
- ? Jaké výstupy z dat CZVV byste využili a jak?
- ? Která metoda nebo nástroj Vám připadá nejdůležitější, resp. bylo by dobré ji dále rozvinout či rozšířit její využití?
- ? ...

Děkujeme za pozornost!
www.EduTest.cz

