Додаток Л

до Звіту про результати першого циклу загальнодержавного моніторингового дослідження якості початкової освіти

«Стан сформованості читацької та математичної компетентностей випускників початкової школи закладів загальної середньої освіти». Частина І. Методологія та технологія. 2018 р.

Характеристика моделей IRT, використаних під час обробки даних моніторингового дослідження

Двопараметрична модель

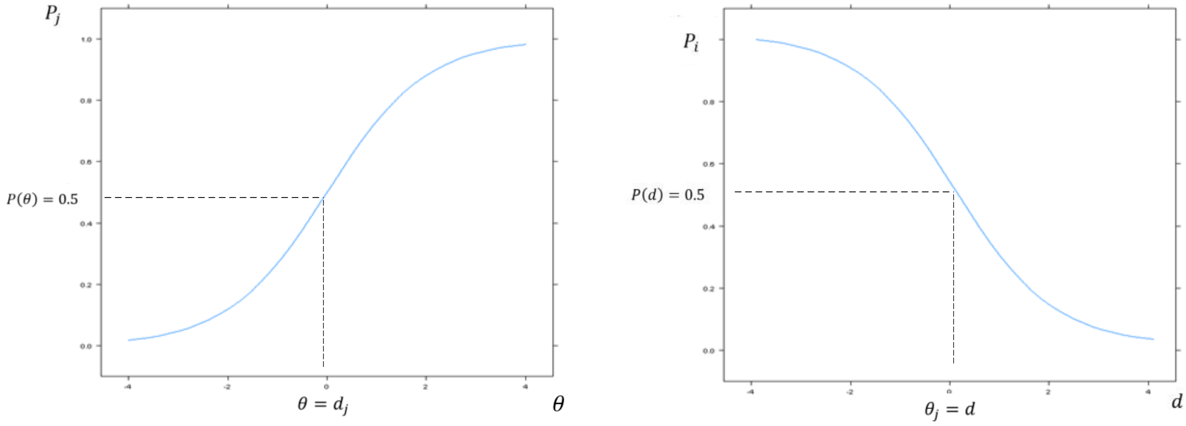
У межах моніторингового дослідження двопараметрична модель IRT використовувалася для оцінки в шкалі логітів параметрів тестових завдань, які мають відповіді в дихотомічній шкалі {0,1}, де 0 – неправильна відповідь, 1 – правильна відповідь.

Відповідно до цієї моделі ймовірність p_{ij} того, що i-й учасник тестування надасть правильну відповідь чи розв'яже j-е тестове завдання (тобто отримає один бал), якщо рівень підготовки цього i-го учасника тестування θ_i, а складність j-го тестового завдання d_j, буде дорівнювати:

$$p_{ij} = P\{x_{ij} = 1 \middle| \theta_i, d_j\} = \frac{1}{1+e^{-G\cdot g_j(\theta_i - d_j)}} \qquad (1)$$

де x_{ij} – відповідь i-го учасника тестування на j-е тестове завдання; $G = 1.7$ – сталий множинник, який узгоджує модель тестового завдання з моделлю нормальної огіви; g_j – параметр розподільної здатності тестового завдання.

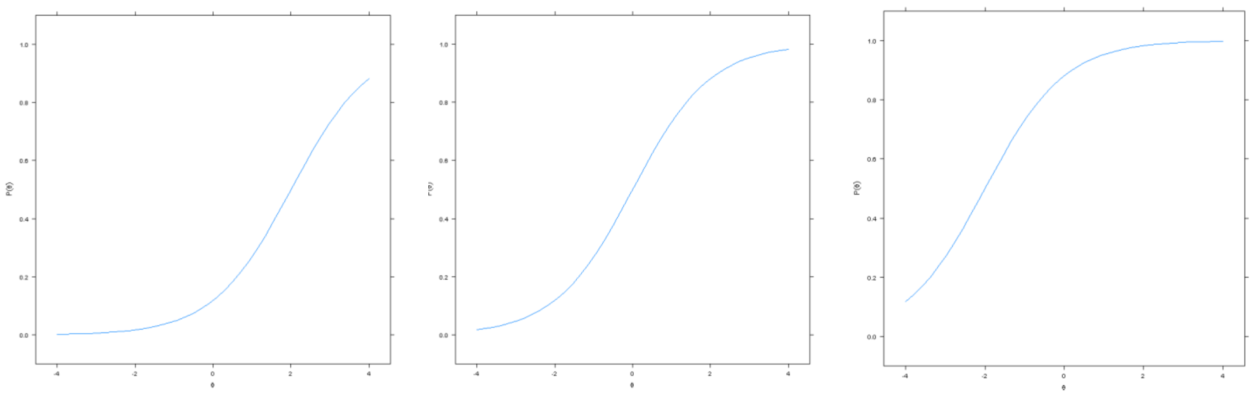
Якщо зафіксувати складність тестового завдання $d = d_j$, то визначається умовна ймовірність того, що учасники тестування з різними рівнями підготовки θ будуть виконувати тестові завдання зі складністю d_j, а якщо зафіксувати $\theta = \theta_i$, то визначається умовна ймовірність того, що учасник із певним рівнем підготовки θ_i буде виконувати тестове завдання з різною складністю d. У першому випадку залежність називається характеристичною функцією тестового завдання складності d_j, а її графік – характеристичною кривою (ICC) j-го тестового завдання (***Рисунок 1а***). У другому випадку функція буде характеристичною функцією рівня підготовленості θ_i, а її графік називається індивідуальною кривою *i*-го учасника тестування (PCC) (***Рисунок 1б***).



а) б)

Рисунок 1 – **Приклад характеристичної кривої тестового завдання (а) і характеристичної функції рівня підготовленості учасника тестування (б)**

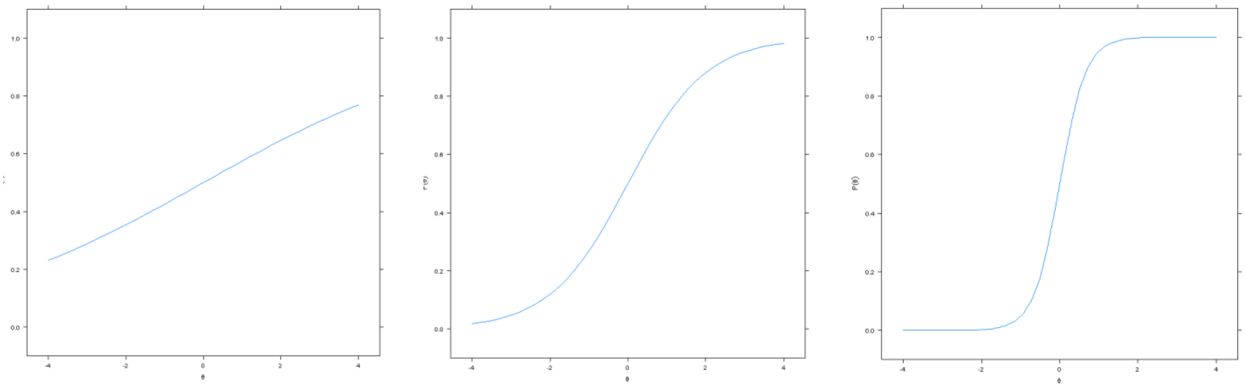
У двопараметричній моделі параметри тестового завдання, які враховуються в процесі обчислень, – це складність d_j і розподільна здатність g_j. На ***Рисунку 2*** наведено графіки характеристичних кривих для тестових завдань різної складності. Графік тестового завдання, яке є оптимальним за складністю, подано на ***Рисунку 2б***.



а) $(\boldsymbol{d = -2})$ б) $(\boldsymbol{d = 0})$ с) $(\boldsymbol{d = 2})$)

Рисунок 2 – **Приклади характеристичних кривих тестових завдань із різними складностями**

У двопараметричній моделі тестові завдання з різними розподільними здатностями мають різні кути нахилу відносно осі θ. Чим ближче до прямого кута, тим розподільна здатність завдання вища, тобто цим тестовим завданням можна розрізняти слабко підготовлених і сильних учнів. Чим ближчий нахил характеристичної кривої до осі θ, тим гірше завдання розподіляє учасників тестування, тобто цим завданням не можна розрізнити добре й слабко підготовлених. На ***Рисунку 3*** представлені графіки характеристичних кривих тестових завдань із різними розподільними здатностями. Характеристична крива тестового завдання, яке дає найбільше інформації щодо індивідуальних відмінностей тестованих, наведена на ***Рисунку 3б***.



а) ($g = 0.3$) б) ($g = 1$) с) ($g = 3$)

Рисунок 3 – **Приклади характеристичних кривих тестових завдань із різними розподільними здатностями**

Адаптована модель відповідей GRM

Для політомічних тестових завдань, тобто тестових завдань, які мають частково правильні відповіді, тобто результати яких можуть бути виміряні в тестових балах {0,1,2}, де 0 – неправильна відповідь, 1 – часткова правильна відповідь, 2 – правильна відповідь, використовувалася адаптована модель відповідей – Graded Response Model (GRM).

Адаптована модель відповідей GRM будується на основі таких припущеннях: нехай тестований i отримує тестове завдання j, у якому можливо подолати m рівнів (категорій). Тоді опитуваний може отримати бал w_{ij}. Ймовірність того, що цей бал набуде вищого значення, ніж l-та категорія, залежить від рівня підготовки тестованого θ, складності подолання цього порогу λ_{jl} та розподільної здатності тестового завдання g_j. Якщо припустити, що w_{ij} має логістичний розподіл, то ймовірність того, що i-й тестований у j-му тестовому завданні подолає l-й або вищий поріг матиме вигляд:

$$p_{lij}^{*} = \frac{e^{g_j(\theta_i - \lambda_{jl})}}{1 + e^{g_j(\theta_i - \lambda_{jl})}}, (l = 0,1,2, \dots, m)$$

У цій моделі λ_{jl} – складність подолання l-го порогу, g_j – параметр розподільної здатності тестового завдання. Варто зазначити, що в GRM $\lambda_{j1} < \lambda_{j2} < \lambda_{j3} < \dots < \lambda_{jm}$, тобто вищому значення порогу, який подолав тестований, завжди відповідає вищий рівень підготовки.

У загальному випадку модель GRM має такий вигляд. Ймовірність того, що тестований подолає l порогів дорівнює:

$$\begin{aligned} p_{lij} &= P\{x_{ij} = l \big| \theta_i, g_j, \lambda_j\} = p_{lij}^{*} - p_{(l+1)ij}^{*} = \\ &= \frac{1}{1 + e^{-g_j(\theta_i - \lambda_{jl})}} - \frac{1}{1 + e^{-g_j(\theta_i - \lambda_{j(l+1)})}} \end{aligned} \qquad (2)$$

Застосовуючи формулу 2 для $l = \overline{0, m}$, необхідно пам'ятати, що $p_{0ij}^{*} = 1$, а $p_{(m+1)ij}^{*} = 0$, тобто кожний тестований може отримати 0, не подолавши жодного порогу, і жоден тестований не подолає $m + 1$ поріг.

На ***Рисунку 4*** представлені функції категорій відповідей політомічного завдання з чотирма категоріями (трьома пороговими значеннями), де значення складності категорій підпорядковані за зростанням: $\lambda_{j1} < \lambda_{j2} < \lambda_{j3}$. Складність категорій така: $\lambda_{j1} = -1$, $\lambda_{j2} = 0$, $\lambda_{j3} = 1$. Значення розподільної здатності для цього тестового завдання дорівнює 1.

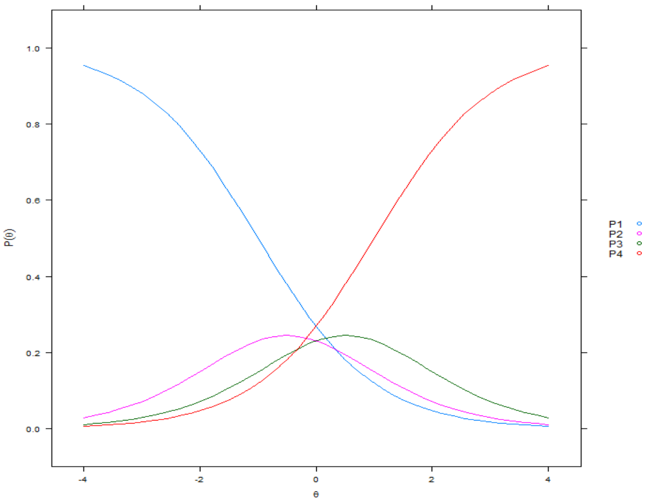


Рисунок 4 – **Приклад функцій категорій відповідей політомічного завдання за моделлю RGM зі складностями категорій $\lambda_{j1} = -1$, $\lambda_{j2} = 0$, $\lambda_{j3} = 1$**

Як видно з ***Рисунка 4*** функції категорій відповідей перетинаються дуже близько одна до одної. Це зумовлено тим, що порогові значення, які відповідають рівням підготовки λ_{j1} та λ_{j3} знаходяться недалеко одне від одного. Тому ймовірність p_{0ij} та p_{3ij} приблизно однакова й дорівнює 0,25. Якщо відстань між значеннями складності категорій буде більше, наприклад, $\lambda_{j1} = -3$, $\lambda_{j2} = 0$, $\lambda_{j3} = 3$, то ймовірність подолання 1-го та 3-го порогу буде дорівнювати 0,5 (***Рисунок 5***).

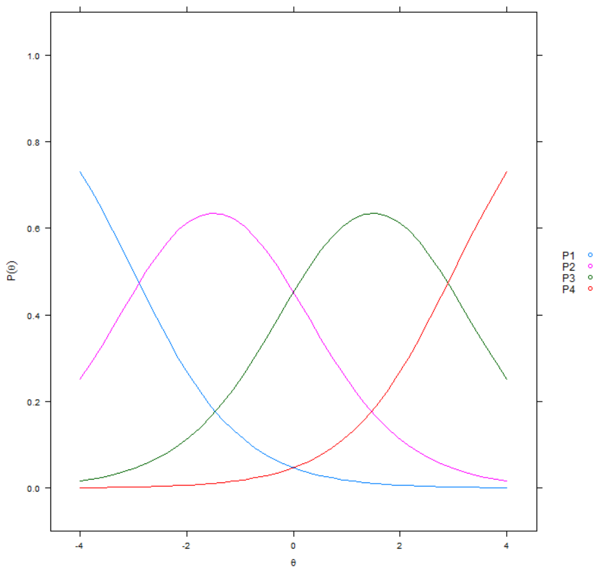


Рисунок 5 – **Приклад функцій категорій відповідей політомічного завдання за моделлю RGM зі складностями категорій $\boldsymbol{\lambda_{j1} = -3,\ \lambda_{j2} = 0,\ \lambda_{j3} = 3}$**

Якщо складність категорій тестового завдання різна, то і ймовірність різна для подолання різних порогів. На ***Рисунку 6*** наведено приклад функцій тестового завдання з такими значеннями складностей категорій: $\lambda_{j1} = -1.5$, $\lambda_{j2} = 1$, $\lambda_{j3} = 2.5$ та розподільною здатністю $g_j = 1.6$. Загальний вигляд інформаційної кривої наведено на ***Рисунку 7***.

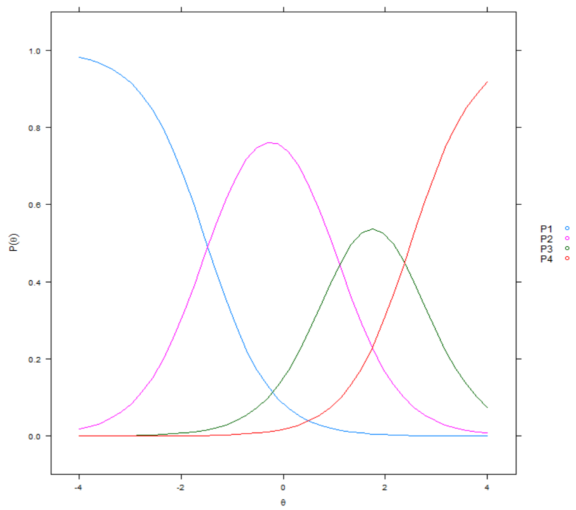


Рисунок 6 – **Приклад функцій категорій відповідей політомічного тестового завдання за моделлю RGM зі складностями категорій $\boldsymbol{\lambda_{j1} = -1,5,\ \lambda_{j2} = 1,\ \lambda_{j3} = 2.5}$ та розподільною здатністю $\boldsymbol{g_j = 1.6}$**

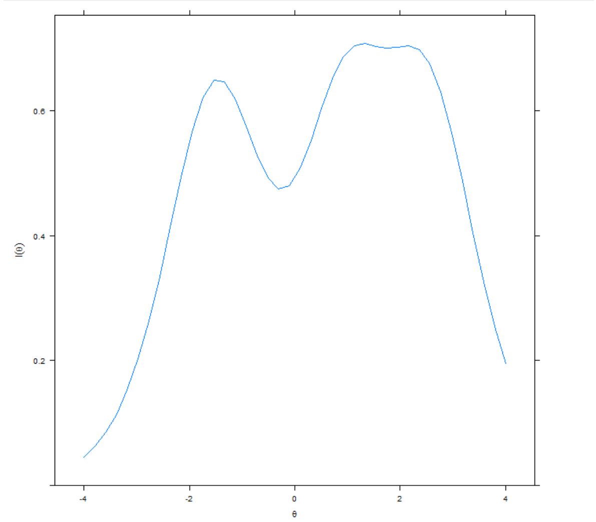


Рисунок 7 – **Приклад інформаційної кривої політомічного тестового завдання за моделлю RGM зі складностями категорій $\lambda_{j1} = -1,5,\ \lambda_{j2} = 1,\ \lambda_{j3} = 2.5$ та розподільною здатністю $g_j = 1.6$**

Якщо розподільна здатність тестових завдань низька, то функції категорій відповідей політомічного завдання, незалежно від складності категорій, будуть мати вигляд прямих. На ***Рисунку 8*** наведено приклад функцій категорій відповідей для тестового завдання, у якого параметр $g_j = 0.2$. Варто зазначити, що в межах одного тестового завдання коефіцієнт розподільної здатності не змінюється, що дає змогу уникнути зайвих перетинів кривих і від'ємних ймовірностей.

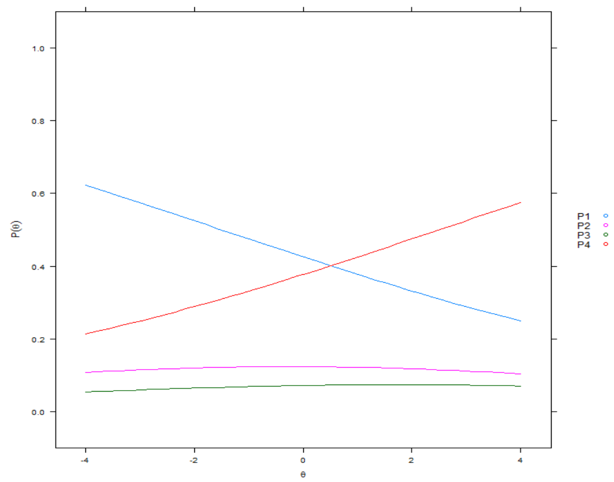


Рисунок 8 – **Приклад функцій категорій відповідей політомічного тестового завдання за моделлю RGM із розподільною здатністю $g_j = 0.2$.**