

УДК 517.538:371.26(045)

І. Ф. Бойко, д-р техн. наук, проф.,
Р. М. Дубан, асп.

ЗАСТОСУВАННЯ ІНТЕГРАЛА ВІД ЛІНІЙНОГО В-СПЛАЙНА ЯК МОДЕЛІ IRT

Інститут аерокосмічних систем управління НАУ, e-mail: romaduban@gmail.com

Запропоновано як модель профілів у тестовому контролі знань використати сплайни, побудовані з базисів, що є інтегралами від В-сплайнів. Розроблено алгоритм оптимізації моделі до емпіричних результатів тестування. Результати проілюстровано на реальних прикладах з інформаційної системи «Logit».

Ключові слова: тестовий контроль знань, сплайн функції, характеристичні криві.

Вступ. У сучасній освіті широко застосовується тестовий контроль як ефективний спосіб перевірки знань. Тестування має ряд переваг, основними з яких є економія часу та об'єктивність результатів. Системи тестового контролю є важливим елементом освітніх інформаційних комп'ютерних технологій. Питання розвитку таких технологій входять до проекту «Відкритий світ» пріоритетного напрямку соціально-економічного та культурного розвитку держави «Нова якість життя» [1].

Педагогічний тест як засіб вимірювання досягнень учнів може дати достовірний результат лише за умови його коректного застосування. Коректність застосування тесту означає розв'язання ряду питань, пов'язаних з розробленням, застосуванням та інтерпретацією результатів тестування. У вітчизняній освіті стандарти інтерпретації отриманих результатів тестового контролю знань хоча і розроблялись, але так і не набули чинності [2]. Зазвичай результати аналізуються на основі класичної теорії тестів (Classical Theory of mental tests) [3] або на IRT (Item Response Theory) [4].

Суть IRT полягає в тому, що тестові питання та респонденти характеризуються деякою функціональною залежністю від складності питання та рівня знань респондента. Ці складові є латентними, прихованими, а теорія встановлює зв'язок між множинами цих характеристик [5]. Теорія базується на математичних моделях, що дозволяють будувати профілі складності тестів та респондентів. Кількість математичних моделей постійно збільшується. Причиною цього є насамперед значний інтерес до цих питань в освіті, бажання отримати більш точну, надійну та просту для використання модель. Різні автори в IRT пропонують нові параметри та їх комбінації, обґрунтовуючи необхідність їх врахування і застосування. Питання вибору оптимальної моделі залишається актуальним і сьогодні.

Незважаючи на велику кількість моделей в IRT значного поширення і застосування набула логістична однопараметрична модель Г. Раша – Rasch measurement [6], яка і стала основою для інших моделей IRT. У цій моделі використовується поняття ймовірності P_{ij} правильного виконання i -м учасником тестування j -го тестового завдання. Значення такої ймовірності залежить від рівня знань учасника тестування Θ_i та рівня складності завдання β_j . Ця залежність, яку називають функцією успіху, має вигляд

$$P_{ij} = \frac{1}{1 + e^{-(\Theta_i - \beta_j)}}; \quad i = 1, N; \quad j = 1, M;$$

де N – кількість учасників тестування; M – кількість завдань тесту.

Правильне використання моделі Раша дозволяє відокремити оцінки рівнів знань від оцінок складності завдань і побудувати криві профілів (рис. 1).

Із наведених залежностей видно, що ймовірність правильної відповіді вища, коли рівень знань учасника тестування вищий, а рівень складності завдань тесту нижчий.

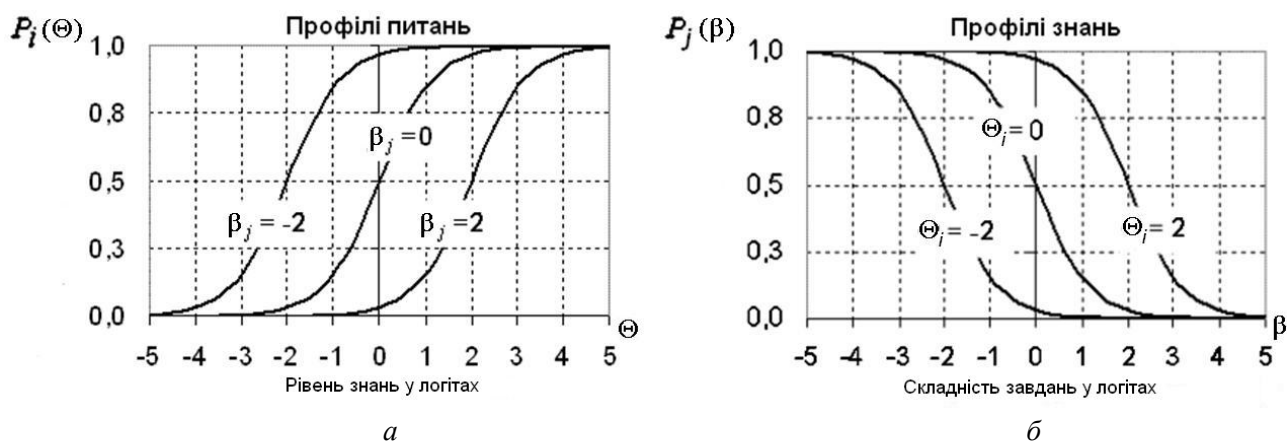


Рис. 1. Характеристичні криві: а – питань (β_j – складність питання); б – знань (Θ_i – рівень знань)

На жаль, модель Раша має істотний недолік, який намагались вирішити автори інших моделей IRT. У теорії Г. Раша не ставилась задача адекватного опису емпіричних даних. У моделі вважається, що не модель має відповідати емпіричним даним, а дані мають відповідати моделі. Відповідно до цього в тест включаються лише ті завдання, що відповідають моделі, а решта не включаються. Такий підхід є малопродуктивним й фактично реалізує принцип «якщо дані не відповідають теорії, то тим гірше для даних». У літературі можна знайти чимало критики [7] з приводу можливості застосування моделі Раша до множини тестів. По суті профілі питань є інтегральними функціями розподілу ймовірностей випадкових величин. Вони починаються з нуля, є неспадними і досягають значення одиниці. Профілі респондентів легко звести до такого ж вигляду через заміну змінної.

Постановка завдання. В роботі пропонується як моделі профілів тестів і респондентів використовувати інтеграли від сплайн-функцій, що дає змогу більш точно описувати емпіричні дані складної залежності. Застосування ермітових кубічних сплайнів з фіксованими вузлами для оброблення даних дозволило отримати відмінний результат. Але їх використання для таких цілей ускладнюється істотним недоліком: вони не враховують неспадного характеру профілів питань. Тому в роботі розглядається сплайн з базисом, який є інтегралом від базису лінійного сплайна, що є неспадною функцією. Метою дослідження є створення інформаційної технології автоматичного оцінювання профілів тестового контролю. Для досягнення мети необхідно отримати математичну інтегральну сплайн-модель профілів тестового контролю, оцінити параметри моделі за емпіричними даними результатів випробовування тесту і реалізувати алгоритми автоматичного оцінювання з оптимізацією розміщення вузлів сплайна та ітераційною процедурою уточнення. Отримані результати перевіряються на реальних даних випробовування тесту.

Виклад основного матеріалу. У загальному вигляді сплайн $S(x)$ записується у вигляді формули [8]:

$$S(x) = \sum_{i=0}^{R-1} a_i B_i(x), \quad x \in [x_1, x_2], \quad (1)$$

де $B_i(x)$ – i -та базисна функція сплайна; a_i – i -й числовий коефіцієнт; R – кількість вузлів сплайна; $[x_1, x_2]$, $x_1 < x_2$ – інтервал, на якому визначено сплайн.

Побудова інтегрального сплайн-базису на основі лінійного сплайна. Для сплайна, що пропонується використовувати, базисними функціями $B_i(x)$ є інтеграли від базисних функцій лінійного сплайна. Базиси лінійного сплайна є поліномами першого степеня, тобто відрізками прямих ліній [9]. І хоча точність інтерполяції лінійними сплайнами є невисокою і вони не забезпечують неперервності навіть похідної першого порядку, інтеграл від базисних функцій лінійного сплайна є гладкими неспадними функціями.

Для побудови у загальному вигляді B -сплайна за рекурентною формулою необхідно щонайменше три вузли [10]:

$$B_i^n(x) = \frac{x - u_i}{u_{i+n} - u_i} B_i^{n-1}(x) + \frac{u_{i+n+1} - x}{u_{i+n+1} - u_{i+1}} B_{i+1}^{n-1}(x), \quad (2)$$

де n – степінь полінома; u_i – значення сплайна в i -му вузлі.

Зі співвідношення (2) отримаємо вираз для полінома першого степеня:

$$B_i^1(x) = \frac{x - u_i}{u_{i+1} - u_i} B_i^0(x) + \frac{u_{i+2} - x}{u_{i+2} - u_{i+1}} B_{i+1}^0(x), \quad (3)$$

де $B_i^0(x)$ набуває значення

$$B_i^0(x) = \begin{cases} 1, & \text{якщо } x \in [u_i, u_{i+1}); \\ 0, & \text{якщо } x \notin [u_i, u_{i+1}). \end{cases} \quad (4)$$

Графічно $B_i^0(x)$ і $B_i^1(x)$ зображено на рис. 2.

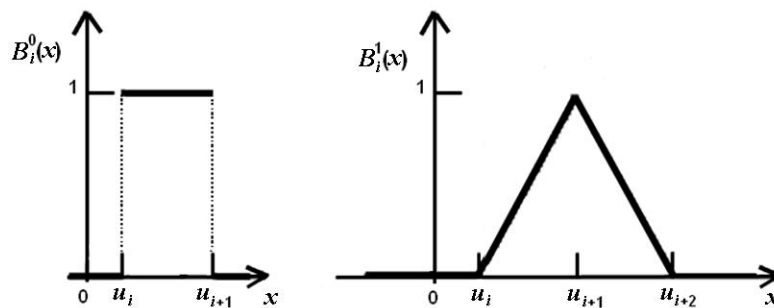


Рис. 2. B -сплайни степені в 0 та 1

Підставивши у співвідношення (3) значення із (4), для $i = 0$ отримаємо

$$B^1(x) = \begin{cases} \frac{x - u_0}{u_1 - u_0}, & \text{якщо } x \in [u_0, u_1); \\ \frac{u_2 - x}{u_2 - u_1}, & \text{якщо } x \in [u_1, u_2); \\ 0, & \text{якщо } x \notin [u_0, u_2). \end{cases} \quad (5)$$

Для отримання інтеграла від базисної функції лінійного сплайна проінтегруємо праву частину співвідношення (5) і отримаємо вирази для першого $[u_0, u_1)$ та другого $[u_1, u_2)$ інтервалів:

$$\begin{aligned} \int_{u_0}^x \frac{y - u_0}{u_1 - u_0} dy &= -\frac{(x - u_0)^2}{2(u_0 - u_1)}, & \text{якщо } x \in [u_0, u_1); \\ \int_{u_1}^x \frac{u_2 - y}{u_2 - u_1} dy &= -\frac{(x - u_1)(x + u_1 - 2u_2)}{2(u_1 - u_2)}, & \text{якщо } x \in [u_1, u_2). \end{aligned} \quad (6)$$

З виразу (6) отримаємо інтегральний базис. Оскільки інтеграл є нагромадженою площею базису лінійного сплайна, кожний наступний фрагмент буде починатись з останнього вузла попереднього. Отже, інтегральний базис для профілів питань $Iq(x)$ запишеться так:

$$Iq(x) = \left\{ \begin{array}{ll} 0, & \text{якщо } x \in (-\infty, u_0), \\ -\frac{(x-u_0)^2}{2(u_0-u_1)}, & \text{якщо } x \in [u_0, u_1), \\ -\frac{(u_1-u_0)^2}{2(u_0-u_1)} + \frac{(x-u_1)(x+u_1-2u_2)}{2(u_1-u_2)}, & \text{якщо } x \in [u_1, u_2), \\ -\frac{(u_1-u_0)^2}{2(u_0-u_1)} + \frac{(u_2-u_1)(u_2+u_1-2u_2)}{2(u_1-u_2)}, & \text{якщо } x \in [u_2, \infty), \end{array} \right\} =$$

$$= \left\{ \begin{array}{ll} 0, & \text{якщо } x \in (-\infty, u_0), \\ -\frac{(x-u_0)^2}{2(u_0-u_1)}, & \text{якщо } x \in [u_0, u_1), \\ \frac{u_2-u_0}{2} + \frac{(x-u_2)^2}{2(u_1-u_2)}, & \text{якщо } x \in [u_0, u_1), \\ \frac{u_2-u_0}{2}, & \text{якщо } x \in [u_2, \infty). \end{array} \right\}. \quad (7)$$

Його графік показано на рис. 3.

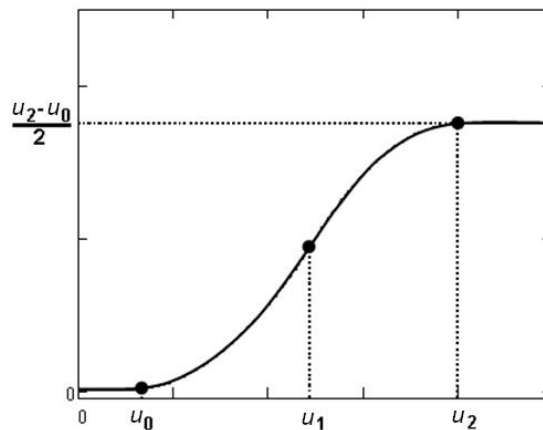


Рис. 3. Інтегральний базис, отриманий з лінійного базису сплайна

Як видно з рис. 3, функція відповідає основним характеристикам профілю питання, але не може застосовуватись для профілів знань. Із формули (7) отримаємо функцію $Ir(x)$ шляхом дзеркального відображення відносно осі абсцис та зміщення на ординату $\frac{u_2-u_0}{2}$:

$$Ir(x) = \frac{u_2-u_0}{2} - Iq(x) = \left\{ \begin{array}{ll} \frac{u_2-u_0}{2}, & \text{якщо } x \in (-\infty, u_0], \\ \frac{u_2-u_0}{2} + \frac{(x-u_0)^2}{2(u_0-u_1)}, & \text{якщо } x \in (u_0, u_1], \\ -\frac{(x-u_2)^2}{2(u_1-u_2)}, & \text{якщо } x \in (u_1, u_2], \\ 0, & \text{якщо } x \in (u_2, \infty). \end{array} \right\} \quad (8)$$

Вона має подібні характеристики і її можна використовувати для побудови профілів знань учасників тестування. Інтегральний базис, побудований за формулою (8), показано на рис. 4.

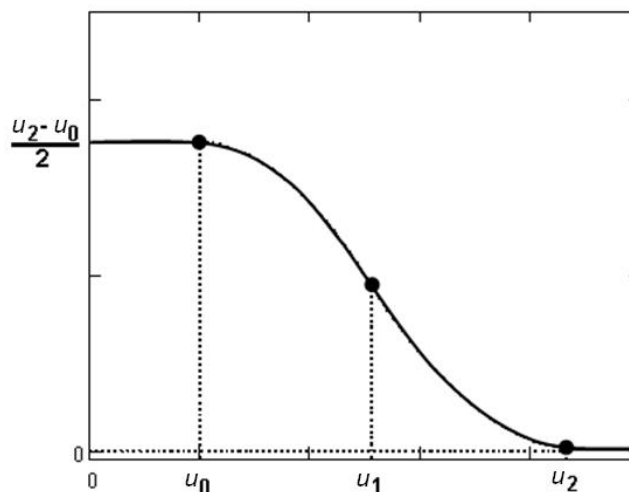


Рис. 4. Дзеркальний інтегральний базис для профілів знань

Побудова інтегрального сплайна та його оптимізація. Отримані інтегральні базиси підставляються в загальну формулу сплайна (1). При цьому спочатку задаються чотири вузли сплайна $u_j : \{0, 1/3, 2/3, 1\}$. Це дозволяє побудувати сплайн, що складається з двох фрагментів інтегрального базису (рис. 5). Шляхом апроксимації емпіричних даних за методом найменших квадратів знаходимо оцінки параметрів a_i [11]. Для закріплення крайніх вузлів використовується зважений метод найменших квадратів (МНК), це дозволяє задовольнити характерні особливості профілів: $f(0)=0$ і $f(1)=1$ для профілів складності питань тесту та $f(0)=1$ і $f(1)=0$ для профілів знань учасників тестування.

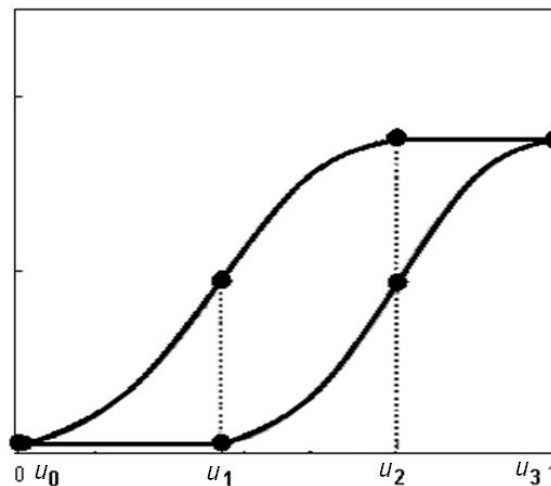


Рис. 5. Чотири вузли і два інтерполяційних базиси

Після першого розрахунку параметрів починається процес оптимізації моделі. Він здійснюється методом покоординатного підбору взаємного розміщення вузлів. Якщо це не дає задовільних результатів, то в середину фрагмента з найбільшою похибкою додається ще один вузол і процес оптимізації починається з початку. Алгоритм оптимізації виконується за методом золотого перерізу, що прискорює процес пошуку оптимального розміщення вузлів сплайна. Оптимальність розміщення вузлів визначається мінімальним значенням критерію оптимізації для відповідного профілю.

Критерій оптимізації K є сумою трьох складових: W_1 – середньоквадратична похибка наближення до емпіричних даних; W_2 – штрафне значення за вихід значень побудованого сплайна за межі інтервалу $[0, 1]$; W_3 – штрафна функція за спадний характер сплайна. До кожного параметра емпіричним шляхом підбирається відповідне значення ваги v_1, v_2, v_3 :

$$K = v_1 W_1 + v_2 W_2 + v_3 W_3,$$

де $W_1 = \sum_{i=1}^N (y_i - S_i)^2$; N – кількість емпіричних даних; y_i – емпіричні значення; S_i – розраховані значення сплайна;

$$W_2 = \sum_{i=1}^M \begin{cases} (S_i - 1)^2, & \text{якщо } S_i > 1, \\ S_i^2, & \text{якщо } S_i < 0, \\ 0, & \text{якщо } S_i \in [0, 1]; \end{cases}$$

$$W_3 = \sum_{i=1}^{M-1} \begin{cases} (S_{i+1} - S_i)^2, & \text{якщо } (S_{i+1} - S_i) < 0, \\ 0, & \text{якщо } (S_{i+1} - S_i) \geq 0, \end{cases}$$

де M – кількість точок інтерполяції побудованого сплайна; S_i – значення точок інтерполяції побудованого сплайна.

У роботі запропоновано ітераційно процедуру уточнення оцінок складності питань та рівня знань респондентів. На першій ітерації емпіричні дані отримуємо як оцінки ймовірності за відносною частотою. На другій та наступних ітераціях емпіричні значення замінюємо на оцінки, отримані за допомогою сплайн-моделі. Ітераційний процес дозволяє уточнювати оцінки рівнів складності питань та знань респондентів, а також підвищує диференційну здатність профілів. Для отримання задовільних результатів виявилось достатнім провести дві-три ітерації. Завдяки задовільним наближенням властивостям сплайна та лінійній залежності від параметрів оператору-експерту не потрібно брати участь у побудові кожного профілю.

Реалізація в інформаційній системі «Logit». Розроблена модель входить у модуль розрахунку профілів складності тестових завдань та рівня знань респондентів, що проходять тестування в авторській інформаційній системі «Logit». Приклад профілів на другій ітерації уточнення для питань показано на рис. 6, а для респондентів – на рис. 7.

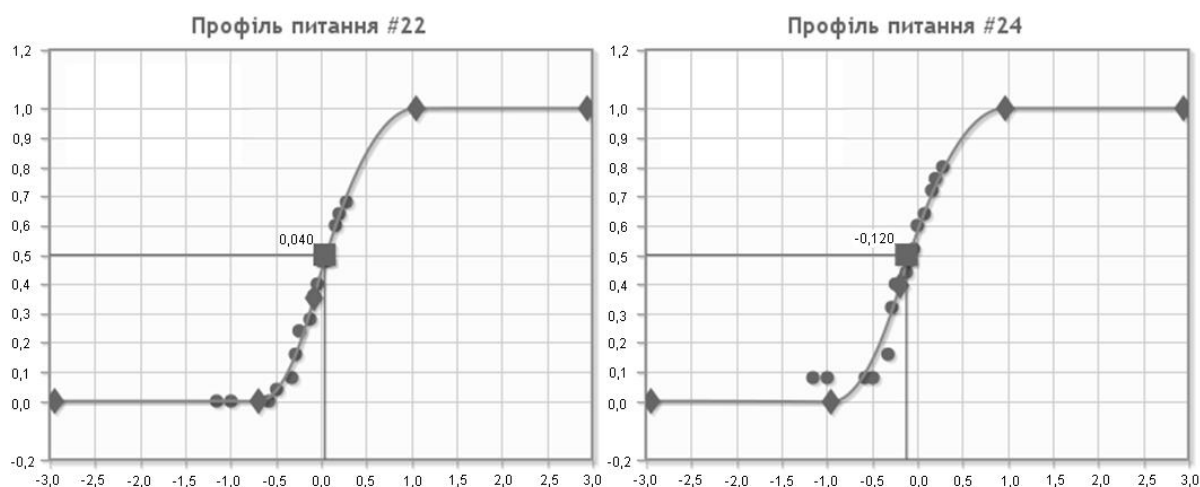


Рис. 6. Профілі питань, побудовані в системі «Logit»: • – емпіричні дані; ♦ – вузли сплайна; ■ – рівень складності

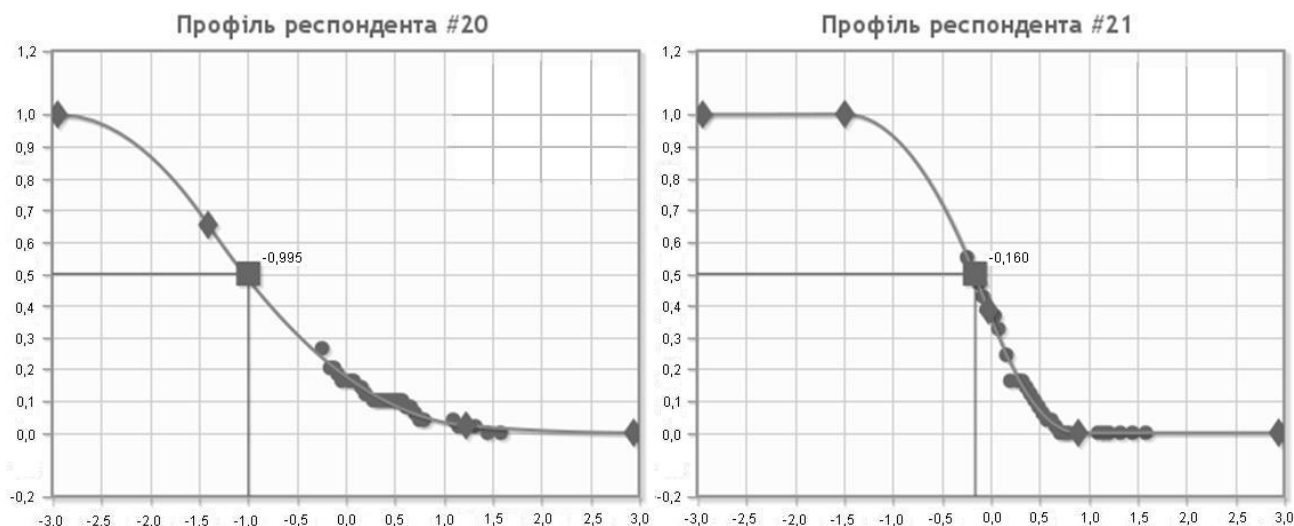


Рис. 7. Профілі респондентів, побудовані в системі «Logit»: • – емпіричні дані; ♦ – вузли сплайна; ■ – рівень знань

Профілі довільних питань з тесту, за яким проводилось тестування на контрольній групі, зображено на рис. 6, вибіркові профілі респондентів – на рис. 7.

Рівень складності питань та рівень знань учасників тестування визначається за перетином профілем рівня $P = 0,5$.

Висновки. Порівняно з відомими моделями IRT отримана модель є загалом більш надійною, універсальною і простою, не потребує участі експертів у побудові профілів. Модель, побудована з використанням як базисів інтеграла від лінійного сплайна має такі переваги:

- модель точніше як інші моделі враховує апріорні посилки;
- модель має лінійну залежність від параметрів;
- оцінки параметрів моделі можна отримати за МНК;
- модель дає можливість оптимізувати профіль.

Але, як і всі моделі, що входять до IRT, сплайн-модель має свої недоліки, зокрема за високої диференційної здатності профіля модель порушує умови невиходу за межі інтервалу ймовірностей $[0, 1]$. Утім такі випадки трапляються відносно рідко, а порушення умови не надто значне. На більшому експериментальному матеріалі це можна спробувати виправити підбором коефіцієнтів критерію оптимізації, або застосувати МНК з обмеженнями. Перед використанням моделі як вимірювальної системи в тестовому контролі знань необхідне провести її апробацію на великій кількості емпіричних даних. Для спрощення процесу апробації модель інтегровано у web-систему «Logit» (<http://kdpu.edu.ua/logit/>), а процес створення профілів автоматизовано.

Наукова новизна роботи полягає в тому, що вперше запропоновано як профілі у тестовому контролі використовувати сплайн, базисом якого є інтеграл від лінійного сплайна.

Практичне значення роботи: за допомогою нової моделі вдалося реалізувати автоматичне оцінювання профілів питань та знань респондентів. Завдяки цьому істотно скорочується час на оцінювання характеристик тесту, немає потреби залучати до процесу оцінювання фахівця-статистика, підвищується точність і достовірність оцінювання.

Предметом подальших досліджень є завершення роботи над інформаційною системою, що втілює продемонстровані результати.

Список літератури

1. Указ Президента України від 08.09.2010 року № 895 «Про заходи щодо визначення і реалізації проектів із пріоритетних напрямів соціально-економічного та культурного розвитку» // Офіційний вісник Президента України. – 2010. – № 26. – 15 с.
2. Комплекс нормативних документів для розробки складових системи вищої освіти. Додаток 1 до Наказу Міносвіти № 285 від 31 липня 1998 р. / Г. Я. Антоненко, І. Є.Булах, В. Л. Петренко та ін. – К.: Ін-т змісту і методів навчання, 1998. – 124 с.
3. Аванесов В. С. Основы научной организации педагогического контроля в высшей школе / В. С. Аванесов – М.: МИСИС, 1989. – 167 с.
4. DeMars C. Item Response Theory / C. DeMars – Oxford University Press, 2010. – 144 p.
5. Ким В. С. Тестирование учебных достижений: монография / В. С. Ким. – Уссурийск: УГПИ, 2007. – 214 с.
6. Rasch G. Probabilistic Model for Some Intelligence and Attainment Tests / G. Rasch – Chicago: Univ. of Chicago Press, 1981. – 199 p.
7. Аванесов В. С. Знания как предмет педагогического измерения / В. С. Аванесов // Педагогические измерения. – 2005. – №3. – С. 3–31.
8. Шелевицький І. В. Слайни в цифровій обробці даних і сигналів / І. В. Шелевицький, М. О. Шутко, В. М. Шутко, О. О. Колганова. – Кривий Ріг: Видавничий дім, 2008. – 232 с.
9. Бойко И. Ф. Статистическая идентификация линейных динамических систем с помощью полиномиальных сплайнов / И. Ф. Бойко, И. В. Шелевицкий, Н. А. Шутко // Статистические методы в теории передачи и преобразования информационных сигналов: тез. докл. междунар. НТК. – К.: КИИГА, 1992. – С. 94.
10. Роджерс Д. Математические основы машинной графики / Д. Роджерс, Дж. Адамс. – М.: Мир, 2001. – 604 с.
11. Шелевицький І. В. Методи та засоби сплайн технології обробки сигналів складної форми / М. О.Шутко. – Кривий Ріг: Європ. ун-т, 2002. – 304 с.

И. Ф. Бойко, Р. М. Дубан

Использование интеграла от линейного *B*-сплайна в качестве модели IRT

Предлагается в качестве модели профилей в тестовом контроле знаний использовать сплайны, построенные из базисов, которые являются интегралами от *B*-сплайнов. Разработан алгоритм оптимизации модели к эмпирическим результатам тестирования. Результаты проиллюстрированы на реальных примерах из информационной системы «Logit».

I. F. Boiko, R. M. Duban

Applying integral linear *B*-spline as a IRT-model

This article studies the usage of splines constructed on the bases of the *B*-splines integrals as a knowledge testing profiles model. Also it represents the model optimization algorithm developed for empirical test results and demonstrates actual results/examples of its implementation into the “Logit” information system.