
**СПЛАЙН-МОДЕЛІ ПРОФІЛІВ СКЛАДНОСТІ ПИТАНЬ ТА ЗНАНЬ
РЕСПОНДЕНТІВ В ТЕСТОВОМУ КОНТРОЛІ ЗНАНЬ**

Пропонується як функція профілів питань і учасників тестування застосовувати сплайни з фіксованими краями. Описуються вирази та особливості роботи з такими моделями. Завдяки новій моделі автоматизується процес розрахунку профілів моделей IRT за даними тестування. Показуються результати оцінювання IRT-профілів.

1. Вступ

Застосування інформаційних комп'ютерних технологій у навчальному процесі - одна з найбільш важливих і стійких тенденцій розвитку світової освіти. У вітчизняній освіті питання поширення застосування інформаційних технологій входить до пріоритетних напрямів соціально-економічного та культурного розвитку держави. Зокрема створення умов для здобуття громадянами якісної освіти входить до пріоритетного напрямку "Нова якість життя"[1]. Розпочато реалізацію національного проекту «Відкритий світ», до задач якого входить стандартизація та уніфікація методик навчання, а також створення централізованої системи навчання і оцінки знань учнів.

Важливим елементом освітніх інформаційних комп'ютерних технологій є системи тестового контролю. Повноцінна система комп'ютерного тестування повинна містити модуль обробки даних, що базується на тій чи іншій теоретичній моделі тестування. Система також має вирішувати задачі визначення рівня складності тестових завдань, що в свою чергу дозволить більш об'єктивно визначити рівень знань осіб, які проходять тестування. На сьогодні відомі та використовуються для визначення характеристик тестових завдань дві основні теорії: класична теорія тестів (Classical Theory of mental tests) та сучасна теорія тестів IRT (Item Response Theory) [2].

Класична теорія тестів розвивалась та вдосконалювала свій математичний апарат із середини минулого сторіччя і дозволяє отримати статистичне обґрунтування якості тесту. Але, незважаючи на добре розроблений математичний апарат, прозорість та ясність висновків, які можна отримати, має принципові недоліки. Зокрема, тестові бали учнів, що проходили тестування, залежать від складності завдань в тесті, а складність завдань залежить від вибірки учнів. Таким чином, великий недолік класичної теорії пов'язаний з наявністю нелінійної залежності між рівнем складності тестових завдань та балами учнів, що пройшли тестування [3].

Натомість, останнім часом у вітчизняній освіті набула популярності сучасна теорія тестування, що розвивається за кордоном протягом вже декількох десятиріч - Item Response Theory[4]. В порівнянні з класичною теорією, IRT має такі переваги: оцінка складності тестових завдань мало залежить від вибірки респондентів, на яких вона була отримана; оцінка рівня підготовленості учнів не залежить від набору тестових завдань, що використовуються; неповнота даних (пропуск деяких комбінацій рівня знань учня – рівня складності завдання) не є критичним [2].

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. Теорія базується на математичних моделях, що дозволяють будувати профілі складності тестів та рівнів знань респондентів. Кількість математичних моделей постійно збільшується. Більш відомими є логістичні моделі Г. Раша та А. Бірнбаума. Основою для багатьох моделей IRT, в тому числі і моделей Бірнбаума, є однопараметрична модель Раша – Rasch measurement [5]. В ній використовується термін ймовірності P_{ij} правильного виконання i -м учасником тестування j -го тестового завдання, що залежить від їх параметрів. Ця залежність називається функцією успіху. Ймовірність визначається за різницею рівнів знань учасника тестування θ_i та рівня складності завдання β_j :

$$P_{ij} = \frac{1}{1 + e^{-(\theta_i - \beta_j)}}. \quad (1)$$

Модель Раша базується на найбільш важливому параметрі $(\theta_i - \beta_j)$ – різниці знань респондента та складності завдання. Оскільки ще в класичній теорії тестів було з'ясовано, що завдання мають різну диференційну здатність, в двопараметричну модель Бірнбаума додатково увійшла диференційна здатність завдання, і функція успіху прийняла вигляд

$$P_{ij} = \frac{1}{1 + e^{-d_j(\theta_i - \beta_j)}}, \quad (2)$$

де d_j – диференційна здатність завдання, що характеризує нахил (крутизну) його профілю. Для більшої відповідності емпіричним даним А. Бірнбаум розробив трипараметричну модель, де третій параметр відповідає ймовірності вгадування [6].

Окрім вказаних двох основних моделей в IRT використовуються ще й інші моделі і їх кількість збільшується [7]. Причиною цього є насамперед значний інтерес до цих питань в освіті і бажання отримати більш точну, надійну та просту у використанні 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.

Значного поширення і застосування серед моделей IRT набула саме однопараметрична модель Г.Раша, хоча і її застосування є досить складним процесом [8]. Нелінійна залежність моделі від параметрів значно ускладнює оцінювання функцій за емпіричними даними. Так, якщо модель не може адекватно описати емпіричні дані, то вони вибраковуються і не використовуються. Такий підхід є малопродуктивним й фактично реалізує принцип “якщо дані не відповідають теорії, то тим гірше для даних”.

Існують математичні функції з лінійною залежністю від параметрів, які дозволяють описувати функціональні залежності достатньо складної форми – це сплайни. В роботі пропонується як моделі профілів тестів та респондентів використовувати сплайн-функції, що дозволяють більш точно описати емпіричні дані складної залежності. Розглядається ермітів кубічний сплайн з фіксованими краями. Метою є створення інформаційної технології автоматичного оцінювання профілів тестового контролю. Для досягнення мети необхідно адаптувати сплайн-модель до апріорних умов профілів тестового контролю, оцінити параметри моделі за емпіричними даними результатів випробовування тесту та реалізувати алгоритми автоматичного оцінювання з оптимізацією розміщення вузлів сплайна. Отримані результати перевіряються на реальних даних випробовування тесту.

2. Виклад основного матеріалу

В практиці обробки даних найбільш часто застосовують кубічні сплайни з двома неперервними похідними [9]. Для їх розрахунку необхідно вирішити систему інтерполяційних рівнянь. Зручніше користуватись кубічними ермітовими сплайнами, що мають неперервність лише першої похідної і локальні розв'язки інтерполяційних рівнянь.

В загальному вигляді сплайни розраховуються за формулою:

$$P_n(x) = \sum_{i=0}^{R-1} a_i H_i(x), \quad (3)$$

де $H_i(x)$ – базисна функція сплайна; a_i – числові коефіцієнти; R – кількість вузлів сплайна.

В ермітового сплайна базисні функції побудовані так, що числові коефіцієнти є значеннями сплайна та його похідними у вузлах:

$$S(x) = \sum_{i=0}^{R-1} f(u_i) H_i(x) + \sum_{i=0}^{R-1} f'(u_i) \hat{H}_i(x), \quad (4)$$

тут u – вузли сплайна.

Для точки x_k , що належить фрагменту, і вираз (4) можна записати так:

$$S(x_k) = f(u_i)^0 h_i(x_k) + f(u_{i+1})^1 h_{i+1}(x_k) + f'(u_i)^0 \hat{h}_i(x_k) + f'(u_{i+1})^1 \hat{h}_{i+1}(x_k), \quad (5)$$

де h – складові базисних сплайнів $H_i(x)$, $H_{i+1}(x)$, $\hat{H}_i(x)$, $\hat{H}_{i+1}(x)$ (рис. 1).

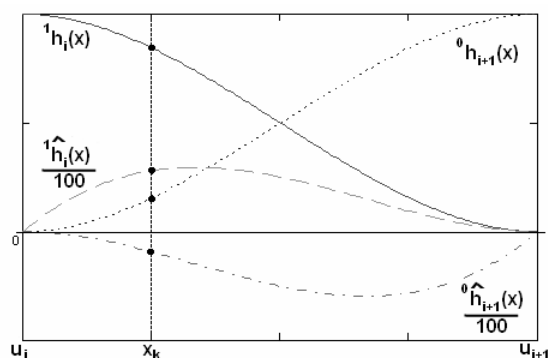


Рис. 1. Складові ермітових базисних сплайнів на інтервалі

Аби не мати справу із похідними та двома видами базисних функцій $H_i(x)$ та $\hat{H}_i(x)$, похідні замінюють центральними розділеними різницями у вузлах сплайна:

$$f'(u_i) \approx \frac{u_{i+1} - u_i}{u_{i+1} - u_{i-1}} \frac{f(u_i) - f(u_{i-1})}{u_i - u_{i-1}} + \frac{u_{i-1} - u_i}{u_{i-1} - u_{i+1}} \frac{f(u_{i+1}) - f(u_i)}{u_{i+1} - u_i}. \quad (6)$$

Загальні вирази (7.1)-(7.4) для таких сплайнів отримано в [10]:

$${}^0h_i(x) = \frac{2x^3 - 3x^2(u_i + u_{i+1}) + 6u_i u_{i+1} x - u_{i+1}^2(3u_i - u_{i+1})}{(u_i^2 - 2u_i u_{i+1} + u_{i+1}^2)(u_{i+1} - u_i)}, \quad (7.1)$$

$${}^1h_{i+1}(x) = \frac{(x - u_i)^2(2x + u_i - 3u_{i+1})}{(u_i - u_{i+1})(u_i^2 - 2u_i u_{i+1} + u_{i+1}^2)}, \quad (7.2)$$

$${}^0\hat{h}_i(x) = \frac{(x - u_i)(x^2 - 2u_{i+1}x + u_{i+1}^2)}{(u_i - u_{i+1})^2}, \quad (7.3)$$

$${}^1\hat{h}_{i+1}(x) = \frac{(x - u_i)^2(x - u_{i+1})}{(u_i - u_{i+1})^2}. \quad (7.4)$$

Важливим аспектом моделі є компактне представлення профілів тестових питань та учасників тестування. Вся інформація про модель міститься в значеннях вузлових точок сплайна або пов'язаних з ними параметрах. Однак модель-профіль має важливі особливості, які явно повинна врахувати сплайн-модель. Для профілів питань сплайн-модель має бути “зафіксованою” на краях згідно з умовами: $f(0) = 0$; $f(1) = 1$. Для профілів респондентів властиві зворотні умови: $f(0) = 1$; $f(1) = 0$. Крайові умови для профілів питань та учасників тестування схематично зображені на рис. 2. Якщо не враховувати ці особливості, то сплайн модель може суттєво порушувати ці умови й буде неадекватною процесу. Будемо називати далі такі ермітові кубічні сплайни фіксованими $F^+(x)$ для першого випадку і $F^-(x)$ – для другого випадку.

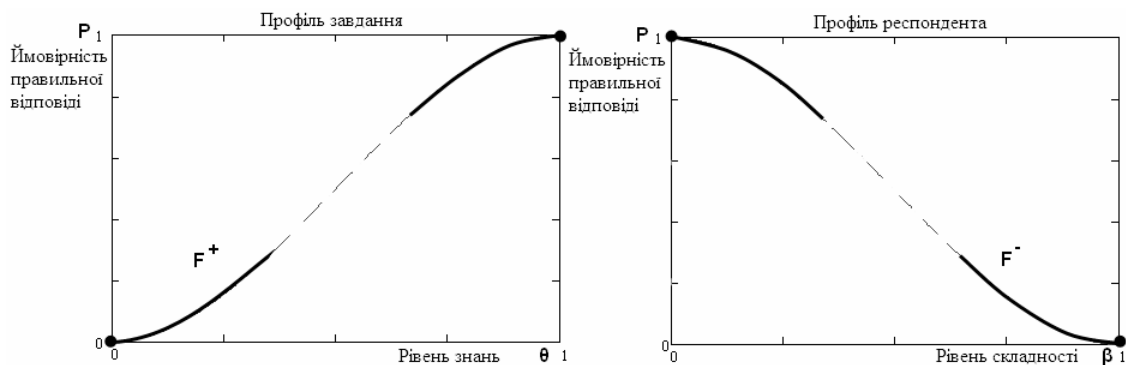


Рис.2. Фіксовані на краях значення профілів

Отримаємо розрахункові вирази для $F^+(x)$ на проміжку $x \in [u_0, u_{R-1}]$. У цьому випадку з виразу (5) отримаємо для першого фрагмента, де $x \in [u_0, u_1]$, вираз:

$$F_1^+(x) = f(u_1)^1 h_1(x) + f'(u_1)^1 \hat{h}_1(x). \quad (8)$$

Для побудови сплайна кількість вузлів R має бути не менше трьох. Якщо вузлів сплайна лише три, тобто проміжок лише два, то другий проміжок є останнім, а вузол u_1 збігається з вузлом u_{R-2} . Після підстановки значень крайніх вузлів у (5) функція для останнього фрагмента прийме вигляд (9), де $x \in [u_{R-2}, u_{R-1}]$:

$$F_{R-1}^+(x) = f(u_{R-2})^0 h_{R-2}(x) + f'(u_{R-2})^0 \hat{h}_{R-2}(x). \quad (9)$$

Графічно в загальному вигляді функції першого (8) та останнього (9) фрагментів сплайна зображено на рис. 3.

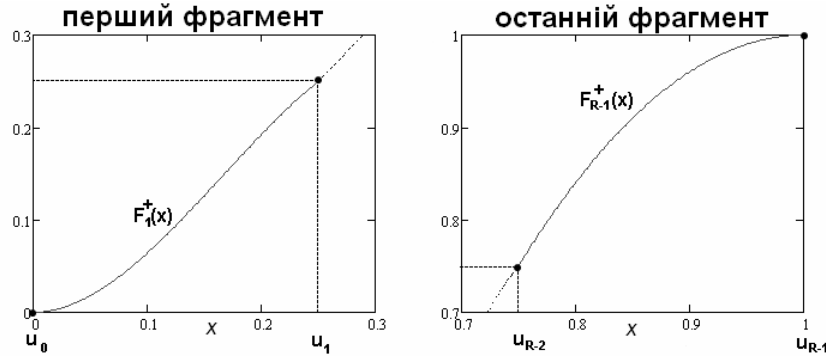


Рис.3. Перший $F_1^+(x)$ та останній $F_{R-1}^+(x)$ фрагменти фіксованого сплайна $F^+(x)$

У випадку $R=4$ середній фрагмент фіксованого сплайна матиме вигляд:

$$F_2^+(x) = f(u_1)^0 h_1(x) + f(u_2)^1 h_2(x), \quad (10)$$

де $x \in [u_1, u_{R-2}]$.

Якщо вузлів сплайна більше трьох, то другий фрагмент прийме вигляд

$$F_2^+(x) = f(u_1)^0 h_1(x) + f(u_2)^1 h_2(x) + f'(u_2)^1 \hat{h}_2(x), \quad (11)$$

тут $x \in [u_1, u_2]$, а передостанній –

$$F_{R-2}^+(x) = f(u_{R-3})^0 h_{R-3}(x) + f(u_{R-2})^1 h_{R-2}(x) + f'(u_{R-3})^0 \hat{h}_{R-3}(x), \quad (12)$$

де $x \in [u_{R-3}, u_{R-2}]$.

Якщо ж вузлів більше чотирьох, то усі фрагменти з другого до передостаннього розраховуються за загальною формулою ермітового кубічного сплайна (5).

В загальному вигляді сплайн-модель $F^+(x)$ профілів тестових завдань на проміжку $x \in [u_0, u_{R-1}]$ для $R=3$ матиме вигляд

$$F^+(x) = \begin{cases} F_1^+(x), & \text{якщо } x \in [u_0, u_1]; \\ F_{R-1}^+(x), & \text{якщо } x \in [u_1, u_{R-1}]; \end{cases} \quad (13)$$

для $R=4$:

$$F^+(x) = \begin{cases} F_1^+(x), & \text{якщо } x \in [u_0, u_1]; \\ F_2^+(x), & \text{якщо } x \in [u_1, u_2]; \\ F_{R-1}^+(x), & \text{якщо } x \in [u_2, u_{R-1}]; \end{cases} \quad (14)$$

для $R \geq 5$:

$$F^+(x) = \begin{cases} F_1^+(x), & \text{якщо } x \in [u_0, u_1]; \\ F_2^+(x), & \text{якщо } x \in [u_1, u_2]; \\ F_{i+1}^+(x), & \text{якщо } x \in [u_i, u_{i+1}], i \in [2, R-3], R > 5; \\ F_{R-2}^+(x), & \text{якщо } x \in [u_{R-3}, u_{R-2}]; \\ F_{R-1}^+(x), & \text{якщо } x \in [u_{R-2}, u_{R-1}]. \end{cases} \quad (15)$$

Це графічно зображено на рис. 4.

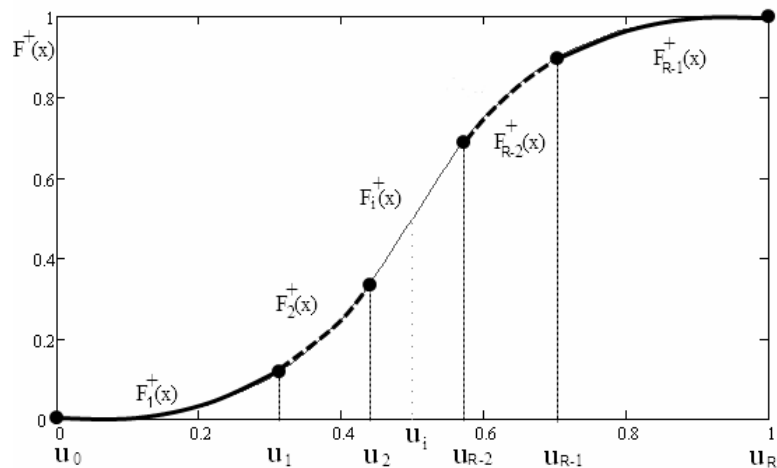


Рис.4. Сплайн-модель профілів тестових завдань

Для побудови профілю рівня знань особи, що проходить тестування, використовується сплайн-функція $F^-(x)$, отримана з виведеної функції $F^+(x)$:

$$F^-(x) = F^+(1-x). \quad (16)$$

Отримана сплайн-модель враховує специфіку профілів, лінійно залежить від параметрів, що є значеннями сплайна у точках стикування – вузлах. Конкретні значення визначаємо за даними тестового контролю в контрольних групах, де отримуємо емпіричні оцінки профілів у точках. Оцінки параметрів сплайна шукаємо за методом найменших квадратів на фіксованій сітці вузлів, так що для профілів питань досягається:

$$D = \sum_{i=1}^N (d_i - F^+(\theta, U))^2 \rightarrow \min.$$

Зменшення похибки наближення досягається також за рахунок оптимізації числа та схеми розміщення вузлів сплайна. Завдяки хорошим наближувочим властивостям сплайна та лінійній залежності від параметрів відпадає необхідність участі оператора-експерта в побудові кожного профілю.

Розроблена модель входить в модуль розрахунку профілів складності тестових завдань та рівня знань учнів, що проходять тестування, в авторській інформаційній системі “Logit”. Приклад профілів питань, що надає система, зображено для питань на рис. 5, для респондентів – на рис. 6.

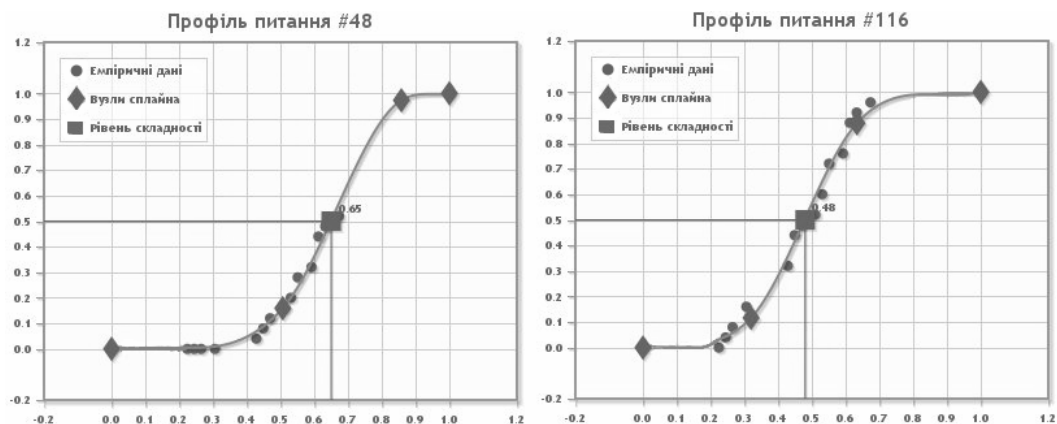


Рис.5. Профілі питань, побудовані в системі “Логіт”

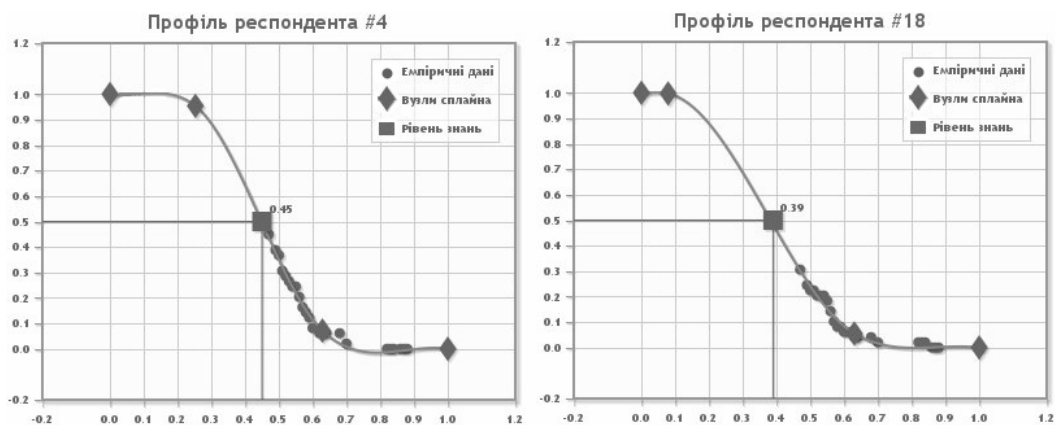


Рис .6. Профілі респондентів, побудовані в системі “Логіт”

На рис. 5 зображені довільні питання з тесту, за яким проводилось тестування на контрольній групі. Вибіркові профілі респондентів зображені на рис. 6. Рівень складності питань та рівень знань учасників тестування визначається за IRT на рівні $P = 0,5$.

3. Висновки

Незначним ускладненням моделі зменшується кількість вхідних параметрів та вирішується проблема фіксації крайніх точок, але швидкість розрахунків комп'ютером підвищується. У порівнянні з моделями IRT отримана модель є загалом більш надійною, універсальною та простою, не потребує участі експертів у побудові профілів.

Модель, що побудована з використанням складеної сплайн-функції з фіксованими кінцями, має такі переваги:

- враховує апріорні посилки,
- має лінійну залежність від параметрів,
- оцінку за МНК,
- дає можливість оптимізувати профіль.

Але, як і всі моделі, що входять до IRT, сплайн-модель має свої недоліки. Одним із найбільших недоліків є те, що функція не враховує неспадаючий характер при побудові профілів питань. Цю проблему можна вирішити, підібравши ще більш вдалий базис сплайна. Перед тим, як рекомендувати використовувати модель як вимірювальну систему в тестовому контролі знань, необхідне проведення її апробації 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 полягає в тому, що вперше запропоновано як профіль у тестовому контролі використовувати ермітові кубічні сплайни з фіксованими кінцями.

Практичне значення роботи в тому, що з допомогою нової моделі вдалося реалізувати автоматичне оцінювання профілів питань та респондентів тесту. Завдяки цьому суттєво скорочується час на оцінювання якісних характеристик тесту, немає необхідності у залученні до процесу оцінювання фахівця-статистика, зростає точність та достовірність оцінювання.

Предметом подальших досліджень є вдосконалення сплайн-моделі для більш точного врахування особливості профілів (а саме їх неспадаючий характер) та 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”. 2. Ким В. С. Тестирование учебных достижений. Монография. Усурийск. УГПИ, 2007. 214 с. 3. Аванесов В. С. Основы научной организации педагогического контроля в высшей школе. М., 1989. 167 с. 4. C. DeMars. Item Response Theory. Oxford University Press: 2010. 144 p. 5. Rasch G. Probabilistic Model for Some Intelligence and Attainment Tests. Chicago: Univ. of Chicago Press, 1981. 199 p. 6. Birnbaum A. Some Latent Trait Models and Their Use in Inferring and Examinee’s Ability. In Lord F.M., Novick M. Statistical Theories of Mental Test Scores. Addison-Wesley Publ. Co. Reading, Mass, 1968. P.397-479. 7. Baker F.B. The Basics of Item Response Theory. ERIC, 2001. 172p. 8. Wilson M. Constructing Measures: An Item

Response Modeling Approach. Mahwah, New Jersey: Lawrence Erlbaum associates, 2005. 228 p. **9. Шелевицький І. В.** Методи та засоби сплайн-технології обробки сигналів складної форми /М.О.Шутко. Кривий Ріг: Європейський університет, 2002. 304 с. **10. Шелевицький І.В.** Сплайни в цифровій обробці даних і сигналів /І.В. Шелевицький, М.О.Шутко, В.М.Шутко, О.О. Колганова. Кривий Ріг, 2008. 232 с.

Надійшла до редколегії 25.08.2011

Дубан Роман Миколайович, аспірант Національного авіаційного університету. Наукові інтереси: web-технології. Адреса: Україна, 50093, Кривий Ріг, вул. Гутовського 27-1, тел. (067)9018237.

Шелевицький Ігор Володимирович, д-р техн. наук, доцент, заступник директора з наукової роботи Криворізького педагогічного інституту ДВНЗ “Криворізький національний університет”. Наукові інтереси: сплайни і їх застосування. Адреса: Україна, 50086, Кривий Ріг, пр. Гагаріна 54, тел. (096)5320143.