

## ІНФОРМАЦІЙНА СИСТЕМА “LOGIT” – ІНСТРУМЕНТ ПІДТРИМКИ ТЕСТОВОГО КОНТРОЛЮ ЗНАНЬ

© Дубан Р., Шелевицький І., 2012

Наведено інформаційну систему “Logit”, яка розроблена для забезпечення підтримки тестового контролю знань на всіх стадіях його життєвого циклу. У представленій інформаційній системі використовуються авторські сплайн-моделі для побудови профілів складності питань та рівня знань.

**Ключові слова:** інформаційна система, тестовий контроль знань, сплайн-модель.

This paper describes an information system “Logit”, which is designed to support test control of knowledge in all stages of its life cycle. This information system use authors spline-model to build profiles of the issues and the level of knowledge.

**Key words:** information system, test control knowledge, spline-model.

### Вступ

Важливою складовою навчального процесу є контроль знань учнів. В українських навчальних закладах активно впроваджують сучасні методи контролю та оцінювання, що дають змогу зробити навчання ефективним та наблизити його до світових стандартів. Все більшої популярності та поширення набувають тестові технології контролю знань. З 2008 року в Україні запроваджено обов'язкове зовнішнє незалежне оцінювання для випускників шкіл, а більшість вищих навчальних закладів впроваджують тестування як один із основних засобів проведення проміжного та підсумкового контролю знань студентів. В умовах достатнього розвитку комп'ютерних технологій процес складання тестів і тестових завдань доцільно виконувати з використанням комп'ютерної техніки та баз даних. Отже, тестовий контроль знань займає важливе місце серед інформаційних освітніх технологій, а питання розроблення інформаційних систем, що забезпечують процес підготовки тестів та проведення тестування, є актуальним.

Хоча тестовий контроль знань має низку переваг, його правильне застосування є доволі складним процесом і потребує багато ресурсів. Основні вимоги до тестів та процедуру їх створення і перевірки викладено в комплексі нормативних документів МОН [1], монографіях [2, 3] та навчальних посібниках [4, 5]. Однак, під час розроблення освітніх тестів як на загальнодержавному рівні (тести ЗНО), так і на рівні окремих навчальних закладів критерії якості тестів або не перевіряють взагалі, або за результатами їх перевірки не вживають заходів із вдосконалення тестів. Зазвичай це пояснюють необхідністю дотримання режиму таємності (стосовно тестів ЗНО), нестачею часу та інших ресурсів. У цьому випадку постає питання про можливість і доцільність практичного використання результатів такого тестування [6].

Можливою причиною такого стану є відсутність практично зорієнтованих методів, алгоритмів і програмного забезпечення для оцінювання тестів. Ціль статті – опис особливостей розробленої інформаційної системи “Logit”, яка розроблялась як інструмент для підтримки тестового контролю знань на етапах створення тесту, його наповнення, визначення характеристик якості, практичного застосування та подальшої верифікації.

### Найважливіший інструментарій розроблення тестів

Для розроблення тестів та їх проведення розроблено чималу кількість програмного забезпечення. Раніше це були комп'ютерні програми, які або зберігали питання тесту всередині, або

поширювались разом із окремим файлом, що містив тест із питаннями. Розвиток технологій дав можливість використовувати клієнт-серверні та web-орієнтовані системи – це підвищило надійність зберігання даних разом із спрощенням процедури оновлення тесту.

Найвідоміше сучасне вільне програмне забезпечення для тестового контролю – це Moodle, iTest, та OpenTEST2. Ці програми дають змогу створювати бази тестів та проводити тестування з використанням комп'ютерної техніки. OpenTEST2 і спеціальний пакет в Moodle дають змогу здійснювати статистичний аналіз результатів. Також існують комерційні програмні продукти, що мають схожий функціонал, але їх перевагами є надання технічної підтримки та покращений вигляд інтерфейсів. Сам підхід до створення тестів у цих та інших програмних засобах має закритий характер, автор чи група авторів розробляють тест і використовують його для вимірювання знань без попереднього рецензування, випробовування та визначення характеристик питань тесту.

Існують наукові дослідження, присвячені теорії тестування з погляду теорії ймовірності та статистики. Однією з центральних проблем досліджень є побудова адекватної моделі умовної ймовірності правильної відповіді на запитання. Найвідомішими моделями є однопараметрична функція Г. Раша [7], двопараметрична та трипараметрична функції А. Бірнбаума [8]. Ці моделі є основою Item Response Theory, за допомогою якої оцінюють характеристики тесту. Практично важливою є побудова цих моделей за результатами тестування. Однак серйозною складністю є нелінійна залежність вказаних моделей від параметрів. Також значним недоліком цих моделей IRT є те, що вони вимагають вилучати з вибірки емпіричних даних ті дані, які не задовольняють теорію. Така вимога порушує цілісність вхідних даних і цим спотворює результат розрахунків.

Існує велика кількість програмного забезпечення – як комерційного, так і безкоштовного, що виконують розрахунки одно-, дво- та трипараметричних функцій, але здебільшого це консольні та призначені для моделювання програми [9]. Їх метою здебільшого є доведення теоретичних розробок, та ці програмні засоби майже не пристосовані для використання на емпіричних даних у зв'язку із складністю інтерфейсів та введеними обмеженнями на кількість вхідних даних.

Оскільки інформаційних систем, які б охоплювали весь життєвий цикл тесту – від створення до використання – не було знайдено, то розроблення нового програмного засобу для вирішення цієї проблеми стане доцільним.

### **Розроблення інформаційної системи підтримки тестового контролю знань**

До розробленої інформаційної системи закладено поняття життєвого циклу тесту (за аналогією із життєвим циклом інформаційних систем). Тест обов'язково до його використання як вимірювального інструмента має пройти стадії створення та наповнення, рецензування, багаторазового випробування та вдосконалення. Лише після отримання тесту з нормованими характеристиками якості його можна рекомендувати до використання. З огляду на те, що інформаційна система підтримки тестового контролю знань планувалась як максимально відкрита та безкоштовна, технології, що використовувались для її створення, було обрано серед вільного програмного забезпечення. Для надання спрощеного доступу до системи використовуються web-технології. Як http-сервер обрано популярний Apache2, що забезпечує стабільну роботу та має велику кількість добре задокументованих налаштувань, підтримується майже усіма операційними системами. Для створення серверного програмного коду використано мову програмування PHP5, що є однією із лідерів серед мов web-програмування та використовує моделі об'єктно-орієнтованого програмування. Дані, використовувані системою, зберігаються в базі даних. Як сервер баз даних використовується MySQL5, обмін даними виконується через запити мовою SQL. Отже, за допомогою обраних технологій реалізується трирівнева клієнт-серверна архітектура інформаційної системи. На стороні користувача використовується довільний web-браузер.

У зв'язку з тим, що до процесу розроблення тесту має залучатись група користувачів системи з різними правами доступу, система обов'язково повинна мати можливість авторизації користувачів, розподілу прав та надання ролей. Передбачені рівні доступу: адміністратор, менеджер, користувач, гість. Адміністратор інформаційної системи – технічний спеціаліст, що володіє знаннями про

функціонування системи та обслуговує її, допомагає у вирішенні проблем всім користувачам, виправляє помилки, виявлені в програмі, та розвиває її функціонал. Менеджери – найдосвідченіші та поважні користувачі системи, які мають досвід розроблення тестів та здебільшого мають науковий ступінь. Уся відповідальність за просування тесту на наступну стадію лежить на них. Звичайні зареєстровані користувачі інформаційної системи можуть мати певні ролі або їхні комбінації, а саме: тестувальник, рецензент, автор. Менеджер має всі можливі з перерахованих ролей користувачів та може надавати їх будь-якому користувачу. Гості є незареєстрованими користувачами системи, що можуть лише переглядати списки тестів та, можливо, проходити тестування. Роль тестувальника надає користувачу проводити тестування з групою респондентів на етапі випробування тесту, після чого він має ввести опис групи та результати. Роль рецензента надає користувачу можливість переглядати цікаві йому тести, їхні розділи і питання та залишати рецензії на них у вигляді зауважень або коментарів. Автор має змогу створювати нові тести, визначати їхні розділи та наповнювати їх питаннями, публікувати чи знімати з публікації питання, відповідати на зауваження в рецензіях. Автор має змогу запропонувати в тест, що йому не належить, своє питання, але опублікувати це питання має змогу лише автор-власник тесту.

Новий тест, створений автором, має містити тему та опис. Будь-який менеджер (краще, якщо він є спеціалістом за темою тесту), вирішує, чи потрібен такий тест і публікує його, призначивши відповідальним за контроль над тестом менеджера, що фактично стає адміністратором цього тесту. Задача адміністратора тесту полягає у відповідальності за переведення тесту на наступну стадію “життєвого циклу”. Першою стадією є розроблення, на неї тест переходить відразу після його публікації і призначення менеджером адміністратора. На першій стадії автор тесту визначає розділи та наповнює їх питаннями, інші автори можуть запропонувати свої питання до певних розділів тесту, автор тесту може їх або прийняти із збереженням авторства, або відхилити, рецензенти мають переглядати питання та робити зауваження і коментарі до тесту загалом і до окремих його питань. Після того, як тест має достатню кількість питань (не менше тридцяти на кожний розділ) та взято до уваги усі зауваження рецензентів або подано пояснення до них, автор може зробити запит до адміністратора тесту на перехід до наступної стадії. Адміністратор тесту, перевірявши виконання необхідних умов, переводить тест на наступну стадію – випробування, або висловлює свої зауваження. На етапі випробування проводиться пробне тестування з групами респондентів за розділами тесту. Випробування проводять користувачі-тестувальники, які мають фіксувати, за якою саме групою респондентів проводилось тестування.

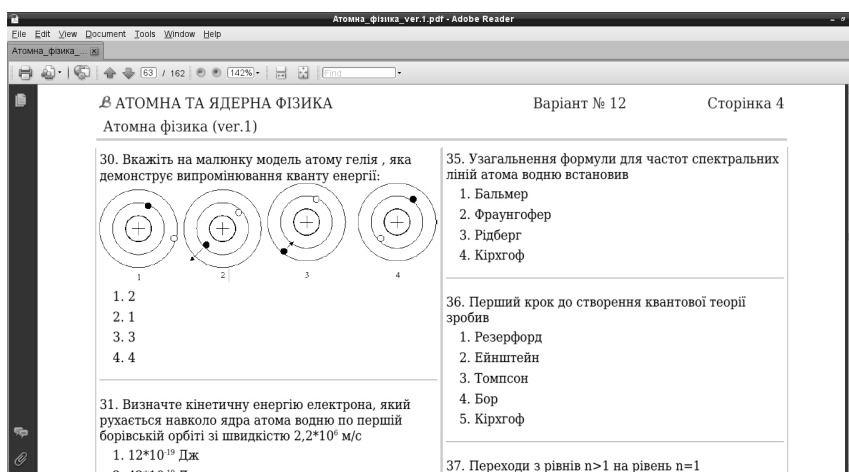


Рис. 1. Файл з варіантами за розділом “Атомна фізика” тест “Атомна та ядерна фізика”

Інформаційною системою передбачено два можливі варіанти проведення тестування: на комп'ютері або в звичайному паперовому вигляді. До проведення тестування необхідно згенерувати потрібну кількість варіантів за обраним розділом. На етапі випробування кожен варіант містить усі питання розділу, але позиція питання та варіанти відповідей на нього довільно перемішані. Кожне таке згенероване тестування може використовуватись на довільній кількості груп. Для

тестування у паперовому вигляді система генерує PDF-файл готовий до друку, що містить задану кількість варіантів (рис. 1). Одні роздруковані варіанти можуть використовуватись на різних групах, але в кожній групі один варіант можна використати лише один раз. Список правильних відповідей до цих варіантів не надається, тому навіть тестувальник під час проведення тестування не знає точно, які відповіді правильні, і не впливатиме на результати.

Окрім питань, файл містить бланк для відповідей. Після проведення тестування в паперовому вигляді тестувальник має перенести результати із заповнених бланків у систему, використовуючи форму для введення результатів (рис. 2).

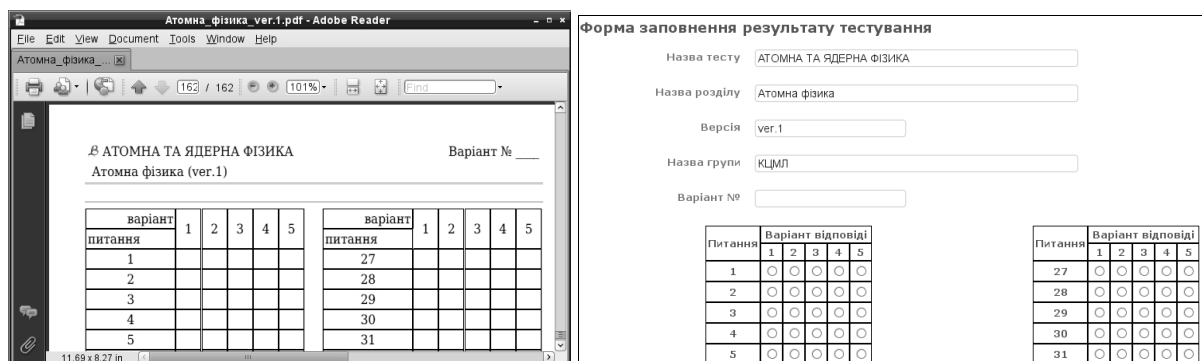


Рис. 2. Бланк відповідей та форма введення відповідей в інформаційну систему

Проведення тестування на комп'ютері спрощує процес, але потребує наявності техніки.

Після отримання перших результатів тестування можна починати процес статистичного аналізу за певними групами респондентів. З надходженням результатів наступних тестувань розрахунки уточнюватимуться, а їх достовірність збільшуватиметься.

### Побудова профілів складності питань та рівнів знань респондентів

Хоча кількість різних моделей IRT сягає десятків, для побудови профілів використовується переважно функція Раша. Це результат добре розробленої теорії. Проте застосування функції Раша пов'язане із труднощами оцінювання параметрів, які входять до рівняння нелінійно. Вибір профілюючої функції значною мірою довільний і ґрунтується на інтуїції й досвіді оператора, а це утруднює автоматизацію цієї процедури. Однак існують функції, які мають гарну адаптивну здатність у сполученні із простою розрахунків і лінійною залежністю від параметрів – це сплайни [10].

В інформаційну систему інтегровано дві авторські сплайн-моделі, що застосовуються для розрахунку та побудови характеристичних кривих рівнів складності питань та знань респондентів. Ці моделі мають такі переваги: дають змогу виконувати розрахунки, враховуючи усі емпіричні дані, модель профілів подано компактно, вся інформація про моделі міститься у значеннях вузлових точок сплайна, під час розрахунку виконується ітераційне уточнення рівнів, процес розрахунку та побудови профілів повністю автоматизовано, для оптимізації використовується трифакторний критерій.

У зв'язку із значною кількістю розрахунків та процедурою ітераційного уточнення час, що витрачається на побудову профілів, перевищує допустимий для генерації web-сторінки. Розрахунки профілів, як і генерація варіантів для проведення пробних тестувань, ставиться до черги на виконання. Швидкість просування черги залежить від потужності сервера та його навантаження, але просувається в межах 10 хвилин. Після виконання завдання черги користувачів буде сповіщено електронною поштою, і вони зможуть продовжити працювати з інформаційною системою.

Завдяки автоматизації процесу розрахунків та побудові профілів в інформаційній системі користувачі можуть використовувати її для вдосконалення тестів без глибоких знань із статистики та математики. Результат розрахунку профілів питань окремого розділу та респондентів певних груп подається зведено у вигляді карток з основними показниками (рис. 3). Кожна картка містить ідентифікатор питання чи респондента, рівень складності або знань у логітах та межі

диференціальної здатності. Є можливість сортувати за ідентифікатором, рівнями складності і знань та за величиною проміжку диференціальної здатності.

Питання	Порівнювати				
	#5	#8	#11	#12	#13
Ідентифікатор питання	0,12	0,58	1,90	1,39	1,99
Складність в логіці	-0,32	-0,80	1,52	0,45	1,59
Верхня межа диф.здатності	0,66	0,99	0,41	0,36	0,36
Нижня межа диф.здатності	-0,20	0,24	-0,36	-0,28	-0,41
	0,12	0,58	1,90	1,39	1,99
	-0,32	-0,80	1,52	0,45	1,59
	0,66	0,99	0,41	0,36	0,36
	-0,20	0,24	-0,36	-0,28	-0,41
	1,32	1,52	1,73	1,15	1,27
	0,53	0,99	0,41	0,36	0,36
	-0,20	0,24	-0,36	-0,28	-0,41
	1,82	1,15	1,27	0,41	1,27
	1,21	0,41	-0,45	0,58	-0,16
	1,21	1,59	1,39	0,28	0,66
	0,49	0,58	0,41	-0,41	0,00
	-0,28	-0,36	-0,20	-0,71	-0,58
	1,21	1,10	2,09	1,82	1,52
	0,53	0,36	1,32	1,45	0,94
	-0,28	-0,24	0,28	1,05	-0,28
	1,27	1,21	1,32	1,32	1,90
	0,58	-0,16	0,71	0,62	1,52
	1,52	2,09	1,39	1,39	1,39
	0,85	0,58	0,75	0,80	0,71
	-0,08	-0,08	0,20	0,08	-0,08
	1,73	1,45	1,82	1,82	2,09
	1,21	0,80	1,39	1,15	1,73
	-0,62	-0,04	0,41	0,24	1,27
	1,39	1,05	2,20	-0,04	1,27
	0,49	-0,36	1,99	-0,45	0,62
	-0,32	-0,36	1,99	-0,45	0,62
	0,41	1,73	1,19	0,49	1,20
	-0,12	0,58	-0,08	-0,62	0,16
	-0,62	1,27	0,16	-0,36	-0,36

Респонденти	Порівнювати				
	#1	#2	#3	#4	#5
Ідентифікатор респондента	-0,71	-0,99	-0,80	-1,15	-1,05
Рівень знань в логіці	0,16	0,45	0,08	0,24	-0,20
Верхня межа диф.здатності	-1,21	-1,05	-0,85	-1,27	-1,52
Нижня межа диф.здатності	-0,53	0,49	0,04	0,41	-0,58
	-0,80	-1,05	-1,21	-1,21	-1,66
	0,04	0,53	-0,49	0,12	-0,66
	-1,52	-1,66	-1,52	-2,09	-0,90
	-0,90	-0,16	-0,99	-0,16	-0,08
	-1,15	-0,80	-2,09	-1,99	-0,90
	-0,45	0,45	-1,59	-0,41	-0,08
	-2,09	-1,59	-0,85	-0,20	-0,85
	-1,45	-0,58	-0,62	0,32	-1,05
	-1,73	-1,90	-1,90	-1,90	-1,39
	-0,80	0,12	-0,90	0,20	-0,71
	-1,45	-1,32	-1,32	-1,27	-1,45
	-0,49	0,32	-0,66	0,16	-0,62
	-1,90	-1,66	-1,66	-1,66	-1,27
	-1,27	-0,41	-1,05	-0,20	-0,32
	-2,09	-1,73	-1,73	-1,99	-0,99
	-1,52	-0,66	-0,80	0,16	-0,66
	-1,73	-1,27	-1,27	-1,66	-1,66
	-1,21	-0,41	-0,58	0,20	-0,99
	-1,39	-2,31	-1,82	-0,85	-1,59
	-0,75	0,00	-1,82	-1,05	-0,58

Рис. 3. Картки профілів питань та респондентів

Є можливість переглянути кожен профіль у графічному вигляді, що допомагає краще зрозуміти розраховані характеристики (рис. 4).

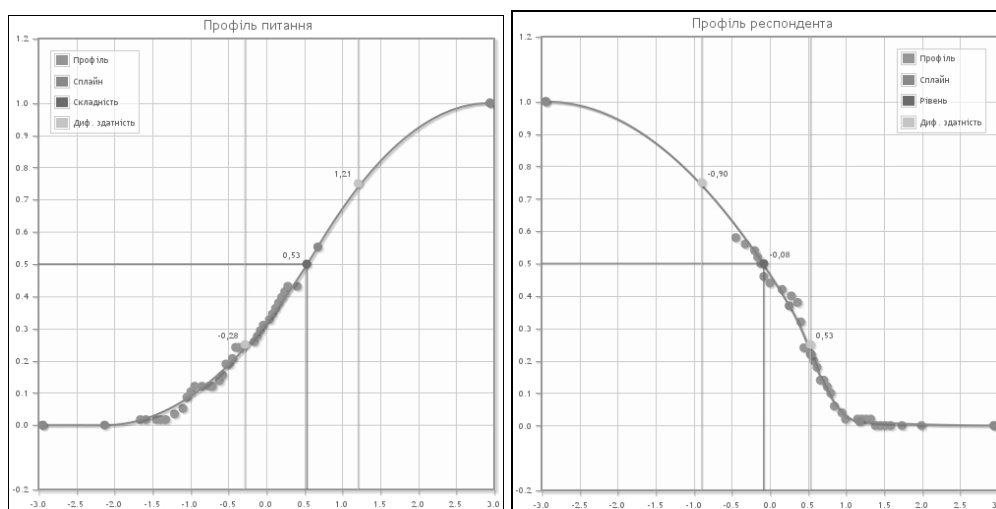


Рис. 4. Графічний вигляд довільних профілів питання та респондента

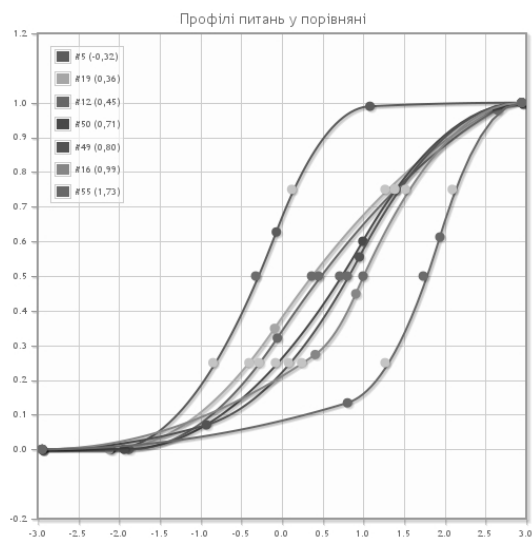


Рис. 5. Порівняння профілів питань

Отримані розрахунки мають використовуватись рецензентами та самим автором для роботи над вдосконаленням тесту. Оцінювання профілів тестових питань вимагає значного числа респондентів з різним рівнем знань. Тому практично доцільною є ітераційна процедура оцінювання характеристик тесту. Спочатку виявляють найочевидніші недоліки й установлюють загальні характеристики на деякій контрольній групі респондентів.

Для практичного застосування важливішою є можливість порівняння окремих профілів питань чи респондентів (рис. 5).

У разі досягнення необхідного рівня вірогідності оцінок профілів тестових питань тест оптимізують (наприклад, вилучають малоінформативні питання). Формують паспорт тесту, що містить профілі питань. Наявність цих характеристик дає можливість об'єктивніше оцінювати рівень знань респондентів. З цього моменту тест можна вважати завершеним інформаційним продуктом з нормованими характеристиками. Тест переводиться на стадію “застосування”, а зміни в нього не вносяться, за винятком умов, що потребують перегляду тесту.

### Висновки

Такий підхід до формування, випробування й застосування тестів здатний гарантувати необхідний (або заданий) рівень вірогідності тестування, перетворивши тест на вимірювальний інструмент із нормованими характеристиками щодо похибок. Функціональні можливості для цього процесу планується вдосконалювати на web-платформі, що дасть змогу оперативно обмінюватись інформацією між учасниками. Результати поточних розробок інформаційної системи “Logit” подано за адресою <http://logit.kdpu.edu.ua/>.

1. *Комплекс нормативних документів для розробки складових систем стандартів вищої освіти. Додаток 1 до наказу Міністерства освіти України № 285 від 31 липня 1998 р.* – К., 1998. – 115 с.
2. *Доровской В.А. Дихотомическое тестирование и оценивание эффективности обучения диспетчеров полиэргатических систем / В.А. Доровской.* – Кривой Рог, 2001. – 412 с.
3. *Ким В.С. Тестирование учебных достижений.* – Уссурийск: Издательство УГПИ, 2007. – 214 с.
4. *Чельщикова М.Б. Теория и практика конструирования педагогических тестов: учебное пособие / М.Б. Чельщикова.* – М.: Логос, 2002. – 432 с.
5. *Аванесов В.С. Композиция тестовых заданий: Учебная книга. 3 изд., доп.* – М.: Центр тестирования, 2002. – 240 с.
6. *Бахрушин В.С., Горбань О.М. Тестові технології в освіті: Проблеми якості тестів // Наукові записки Академії наук вищої школи України.* – 2011. – Т. 6. – с. 24 – 34.
7. *Rasch G. Probabilistic Model for Some Intelligence and Attainment Tests.* – Chicago: Univ. of Chicago Press, 1981. - 199 p.
8. *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.*
9. *Baker F. B. The basics of item response theory. 2nd ed. : ERIC, 2001.–. 176 p.*
10. *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.