

УДК 303.714

А.А. Кравченко, А.С. Шантир, Д.С. Шантир

МІНІМІЗАЦІЯ ВИПАДКОВОЇ ПОХИБКИ ПРИ ОЦІНЦІ ЗНАТЬ ТЕСТАМИ ІЗ ЗАКРИТИМИ ЗАВДАННЯМИ

Вступ

Найбільш ефективним засобом вимірювання в процесі контролю знань є тести [1], використання яких забезпечує реалізацію всіх функцій контролю [2] та відповідає основним принципам контролю [3]. Тестування дає об'єктивну, порівнювану і навіть кількісну оцінку якості знань, що в свою чергу надає широкі можливості в керуванні навчальним процесом. Для побудови моделей тестування використовуються два основні підходи – класична емпірико-статистична теорія тестів та стохастична теорія тестів [4, 5].

Основу класичної емпірико-статистичної теорії тестів становлять концепція похибки результату вимірювання та статистичні методи аналізу. Головною метою застосування класичних тестів є встановлення відношення порядку між проявленими випробовуваними індивідами знаннями та визначення місця (рейтингу) кожного на заданій множині протестованих індивідів [6]. При встановленні відношень порядку виникають помилки внаслідок наявності випадкової та систематичної похибок результату вимірювання тестом. Вплив систематичної похибки на результати вимірювання тестом зводиться до додавання константи до істинного значення вимірюваної величини, що не має значення для інтервальних шкал. Випадкова похибка може бути зменшена за рахунок підвищення ретестової надійності тесту [5], але не може бути виключена.

Стохастична теорія тестів базується на латентно-структурному аналізі [7], згідно з яким поведінка (вирішення завдань тесту) індивіда, що тестується, є проявом деяких латентних характеристик (складність завдань у тесті, рівень знань індивіда) [8]. Головною метою латентно-структурного аналізу є вивчення поведінки індивідів для виведення латентних характеристик та класифікації індивідів за близькістю (рівністю) значень цих характеристик. Стохастична теорія тестів має велику кількість недоліків, детально описаних у літературі [5, 9]. Слід зазначити, що ряд недоліків пов'язаний з тим, що будь-які випадкові дані можуть бути приве-

дені у відповідність до деяких моделей тестування (наприклад, до моделі Раша).

Таким чином, незалежно від того, який підхід використовується при побудові моделей тестування знань, завжди виникають випадкові похибки вимірювання тестами, пов'язані з можливістю заповнювати тестові форми довільно. Оскільки випадкова похибка при вимірюванні тестами може бути значною, то розробка нових методів, що дають змогу мінімізувати цю похибку, є важливим напрямком дослідження.

Постановка задачі

Метою даної статті є розробка методу мінімізації випадкової похибки при вимірюванні тестами із закритими завданнями розширенням первинної вимірювальної інформації.

Метод тестування знань з оцінкою вибору

Тест є системою завдань специфічної форми з рівномірно зростаючою складністю [6]. Під специфічною формою мається на увазі формулювання завдань у вигляді набору висловлювань (варіантів вирішення), істинних або помилкових. При вимірюванні знань тестом результат вибору варіанта або варіантів вирішення завдання є первинною вимірювальною інформацією, причому правильному вирішенню відповідає оцінка знань 1, неправильному – 0. У даному випадку первинна вимірювальна інформація відображає лише результат вибору при ігноруванні при цьому способі, за яким виконується вибір (результат вибору є наслідком або вибору з використанням наявних у індивіда знань, або довільного вибору). Іншими словами, при вирішенні завдань індивід за умови наявності необхідних знань вирішить завдання правильно. За умови відсутності необхідних знань індивід має довільно вибрати варіант або варіанти вирішення завдання, причому імовірність правильного вирішення дорівнюватиме:

- $\frac{1}{k}$, якщо відомо, що один із варіантів вирішення правильний;
- $\frac{1}{2^k}$, якщо кількість правильних варіантів вирішення не відома, де k – кількість варіантів вирішення завдання.

Таким чином, факт наявності правильного вирішення завдання індивідом не несе інформації про спосіб вирішення цього завдання, що

унеможлиблює виключення випадкової похибки вимірювання.

Для усунення вказаного недоліку необхідно розширити первинну вимірювальну інформацію наданням індивіду можливості не лише вибрати правильну, на його думку, комбінацію варіантів розв'язання задачі, а й вказати ступінь впевненості в зробленому виборі – оцінити свій вибір.

Розглянемо можливі способи вирішення [10] тестового завдання індивідом на прикладі завдання з чотирма варіантами вирішення ($k = 4$), один з яких правильний. За наявності у індивіда знань для вирішення цього завдання в достатньому обсязі вибір правильного варіанта буде зроблено з імовірністю $P = \frac{1}{k-3} = 1$, оскільки буде

відкинуто три достовірно невірні. За наявності знань для вирішення завдання в недостатньому обсязі індивід може відкинути один або два достовірно невірних варіантів і виконати вибір із можливих варіантів з імовірністю $P = \frac{1}{k-1} \approx 0,33$ або $P = \frac{1}{k-2} = 0,5$, відповідно. За відсутності знань, необхідних для вирішення завдання, індивід може виконати довільний вибір з імовірністю $P = \frac{1}{k} = 0,25$. Тестова

форма для даного випадку наведена на рис. 1. При заповненні тестової форми індивід має вибрати правильний, на його думку, варіант вирішення і поставити відмітку в потрібному рядку та стовпчику, що відповідає ступеню його впевненості у своєму виборі. На рис. 1 ступінь впевненості у вибраному варіанті чисельно відповідає імовірності правильного вибору P у відсотках.

При збільшенні кількості варіантів вирішення завдання слід також збільшувати кількість ступенів впевненості в зробленому виборі (рис. 2).

При невідомій кількості правильних варіантів вирішення завдання кількість можливих комбінацій варіантів вирішення, які становлять правильне вирішення завдання, зростає. Тому доцільно надати можливість індивіду вказати ступінь впевненості в зробленому виборі в діапазоні $P \in \left[\frac{1}{2^k}; \frac{1}{2^k - (2^k - 1)} \right]$ (рис. 3).

Варіанти вирішення завдання	Ступінь впевненості в зробленому виборі (оцінка вибору)			
	Не знаю: впевненість > 25 %	Можливо: впевненість > 33 %	Найімовірніше: впевненість > 50 %	Точно: впевненість 100 %
1				
2				
3				
4				

Рис. 1. Тестова форма: закрите завдання з чотирма варіантами вирішення, один з яких правильний

Варіанти вирішення завдання	Ступінь впевненості в зробленому виборі (оцінка вибору)				
	$P = \frac{1}{k}$	$P = \frac{1}{k-1}$	$P = \frac{1}{k-2}$...	$P = \frac{1}{k-(k-1)}$
1					
2					
3					
⋮					
k					

Рис. 2. Тестова форма: закрите завдання з k варіантами вирішення, один з яких правильний

Варіанти вирішення завдання	Правильний варіант	Ступінь впевненості в зробленому виборі (оцінка вибору)
		$P \in \left[\frac{1}{2^k}; \frac{1}{2^k - (2^k - 1)} \right]$
1		
2		
3		
⋮		
k		

Рис. 3. Тестова форма: закрите завдання з k варіантами вирішення, якщо кількість правильних варіантів вирішення не відома

Таким чином, в результаті вимірювання тестом з оцінкою вибору отримуємо первинну вимірювальну інформацію, яка відображає результат і спосіб вирішення завдання. Оцінка знань при цьому розраховується за формулою

$$x = a f(P),$$

де a – коефіцієнт, що відображає правильність вирішення завдання (коефіцієнт набуває значення +1, якщо завдання вирішено правильно, і -1, якщо завдання вирішено неправильно); $f(P)$ – функція оцінки вибору, значення якої залежить від ступеня впевненості індивіда в зробленому виборі (імовірності, з якою зроблено вибір). Як функцію оцінки вибору доцільно використати функцію $f(P) = P$, проте можна використовувати й інші функції залежно від мети тестування.

При вирішенні тестових завдань індивід має множину стратегій їх вирішення, які залежать від наявних у нього знань. З точки зору індивіда, який застосовує стратегію довільного вирішення завдання (саме ця стратегія зумовлює наявність випадкової похибки вимірювання тестами), найоптимальнішою є оцінка вибору $P = \frac{1}{k}$ (або $P = \frac{1}{2^k}$), оскільки за умови неправильного вирішення завдання оцінка знань не є мінімальною. Таким чином, індивід, що проходить тестування, зацікавлений, навіть у випадку застосування стратегії довільного вирішення завдань, адекватно оцінювати свій вибір. У разі правильного довільного вирішення завдання оцінка знань не буде максимальною, що зменшить випадкову похибку вимірювання тестом.

Експериментальна перевірка методу тестування знань з оцінкою вибору

Експериментальна перевірка ефективності розробленого методу тестування була проведена в такому порядку:

- збір первинної вимірювальної інформації;
- оцінка знань протестованих індивідів, виходячи з первинної вимірювальної інформації, отриманої без використання і з використанням оцінки вибору;
- аналіз отриманих результатів.

Загальна кількість протестованих з використанням розробленого методу становила 136 студентів чотирьох-п'яти років навчання. Тестування проводилося з різних предметів у різних академічних групах, причому кількість тестових завдань залежала від предмета, по якому проводилося тестування.

Мінімальна кількість протестованих по одному предмету студентів дорівнювала 25. За формулою Брукса і Каррузера [11], для заданої вибірки оптимальним числом інтервалів групування m є:

$$m = 5 \lg n = 5 \lg 25 \approx 7,$$

де n – об'єм вибірки. Для подальшого аналізу зручно вибрати шкалу оцінювання знань із максимальною кількістю балів, яка дорівнює оптимальному числу інтервалів групування m .

Оцінка знань виконувалась двічі. Вперше – виходячи з первинної вимірювальної інформації, отриманої з використанням оцінки вибору. Вдруге – виходячи з первинної вимірювальної

інформації, отриманої без використання оцінки вибору, причому ця інформація отримана відкиданням оцінки вибору з первинної вимірювальної інформації, отриманої з використанням оцінки вибору, тобто одні і ті ж знання, проявлені студентами при тестуванні, оцінені двічі – виходячи з різної первинної вимірювальної інформації.

Візьмемо гіпотезу нормальності статистичного розподілу ідеальних (зроблених без похибок) оцінок знань по предметах. Якщо гіпотеза справджується, результати оцінювання знань по предмету тестуванням повинні мати наближено нормальний розподіл, і наближення буде кращим, якщо випадкова похибка вимірювання тестом буде меншою. Виконаємо оцінку знань студентів, виходячи з різної первинної вимірювальної інформації, та побудуємо розподіли оцінок знань по предметах. На рис. 4–6 наведено розподіли оцінок знань по трьох різних предметах.

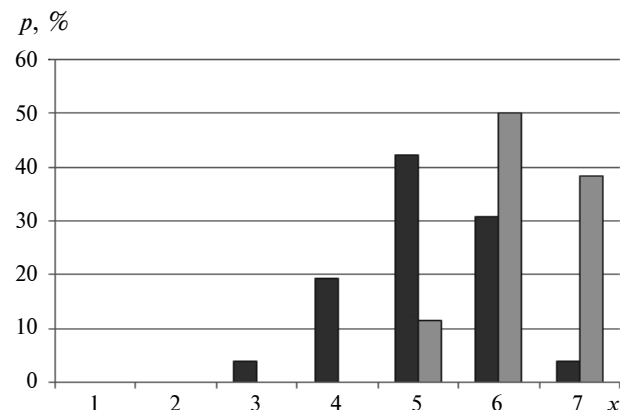


Рис. 4. Розподіл оцінок знань студентів по предмету 1; тут і на рис. 5 і 6: ■ – тестування з оцінкою вибору; ■ – класичне тестування

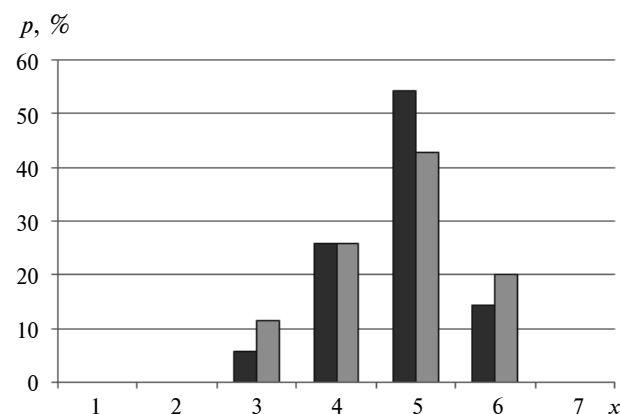


Рис. 5. Розподіл оцінок знань студентів по предмету 2

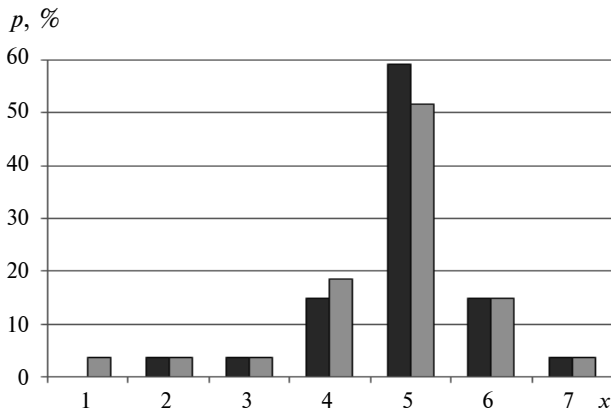


Рис. 6. Розподіл оцінок знань студентів по предмету 3

Перевіряємо відповідність отриманих статистичних розподілів нормальному розподілу, використовуючи критерій χ^2 [12]. Результати наведено в таблиці.

Таблиця. Відповідність отриманих статистичних розподілів нормальному розподілу

Тестування	Характеристика розподілу	Предмет 1	Предмет 2	Предмет 3
З оцінкою вибору	Число ступенів свободи	6		
	χ^2	0,337	1,213	10,09
	Імовірність спрavedження гіпотези	0,996	0,976	0,136
Класичне	Число ступенів свободи	6		
	χ^2	0,508	2,263	24,77
	Імовірність спрavedження гіпотези	0,994	0,9	0,001

З таблиці видно, що для всіх трьох предметів статистичні розподіли оцінок знань, отримані з використанням вимірювання тестом з оцінкою вибору, відповідають нормальному розподілу з більшою імовірністю, ніж статистичні розподіли оцінок знань, отриманих із використанням вимірювання класичним тестом. Таким чином, випадкова похибка вимірювання тестом з оцінкою вибору мінімізована порівняно з вимірюванням класичним тестом.

Статистичні розподіли оцінок по різних предметах наближені до нормального розподілу неоднаково, що свідчить про різну якість тестів. Тестування з оцінкою вибору дає можливість максимально наблизити отриману на основі вимірювання тестами оцінку до ідеальної. Таким чином, тестування з оцінкою вибору може слугувати для перевірки показників якості тестів (надійність, валідність), оскільки за умови низької ймовірності спрavedження гіпотези

зи нормальності розподілу та мінімальної випадкової похибки вимірювання зміст, цілісність та структура тесту викривляють первинну вимірювальну інформацію про наявні знання – такі тести не можуть використовуватися в навчальному процесі.

Структура інформаційної системи тестування знань з оцінкою вибору

Метод тестування з оцінкою вибору може бути застосований для вилучення випадкових похибок вимірювання тестами при побудові будь-яких моделей тестування, при цьому складність реалізації системи тестування із застосуванням інформаційної техніки істотно не зростає. На рис. 7 наведено структуру інформаційної системи тестування з оцінкою вибору. База тестових завдань $\{Q_i\}$ містить тестові завдання та інформацію про їх складність. Підсистема вибору і видачі i -го тестового завдання виконує вибір тестового завдання з бази за заданим алгоритмом (або планово при реалізації моделі класичного тестування, або виходячи з оцінки вирішення попереднього завдання та складності наявних у базі завдань при реалізації адаптивного тестування). Підсистема оцінки вирішення i -го тестового завдання виконує оцінку вирішення тестового завдання із врахуванням інформації про складність цього завдання. Підсистема оцінки знань за результатами тесту виконує сумування оцінок при вирішенні всіх запропонованих завдань, перево-

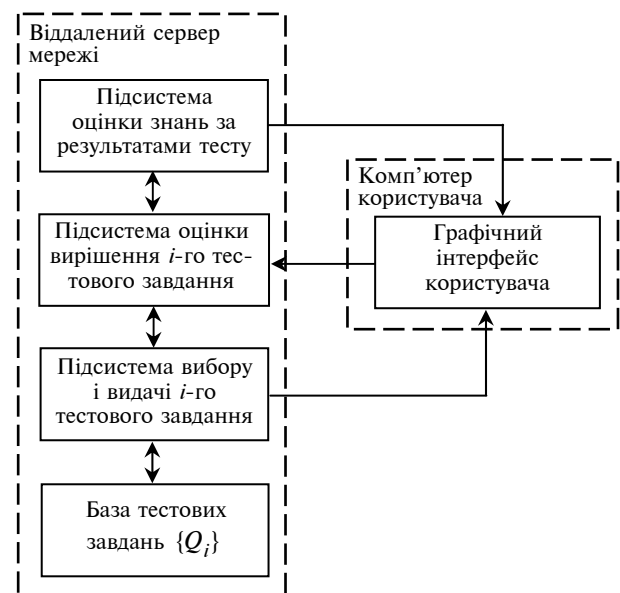


Рис. 7. Структура інформаційної системи тестування з оцінкою вибору

дить отриману оцінку в задану шкалу оцінювання знань та видає результат користувачу (особі, що тестується) через графічний інтерфейс користувача. Перелічені підсистеми реалізуються програмно та розміщуються на віддаленому сервері мережі. На комп'ютері користувача розміщено лише графічний інтерфейс користувача, який забезпечує проведення тестування та обмін супутньою інформацією з підсистемами.

Висновки

Експериментальна перевірка розробленого методу мінімізації випадкової похибки вимірю-

вання тестами із закритими завданнями показала його ефективність порівняно із звичайним тестуванням.

Запропонована структура інформаційної системи тестування з оцінкою виборів дає можливість реалізовувати для побудови моделей тестування як класичну емпірико-статистичну теорію тестів, так і стохастичну теорію тестів.

Подальші дослідження можуть бути спрямовані на застосування розробленого методу для підвищення ефективності адаптивних тестів та перевірки показників якості тестів, що використовуються в навчальному процесі.

А.А. Кравченко, А.С. Шантырь, Д.С. Шантырь
МИНИМИЗАЦИЯ СЛУЧАЙНОЙ ПОГРЕШНОСТИ ПРИ ОЦЕНКЕ ЗНАНИЙ ТЕСТАМИ С ЗАКРЫТЫМИ ЗАДАНИЯМИ

Предложен метод минимизации случайной погрешности при измерении тестами с закрытыми заданиями путем расширения первичной измерительной информации. Проведена экспериментальная проверка предложенного метода, которая подтвердила его эффективность. Приведена структура информационной системы для реализации разработанного метода тестирования.

A.A. Kravchenko, A.S. Shantyr, D.S. Shantyr
MINIMIZATION OF RANDOM ERROR OF KNOWLEDGE ESTIMATION USING CLOSED TESTS

By employing the expansion of primary measuring data, we propose the method of random error minimization when measuring by optional choice tests. In addition, we experimentally verify this method as well as confirm its efficiency. We develop the system of information structure for implementing the proposed testing method.

1. *Чельшкова М.Б.* Теория и практика конструирования педагогических тестов: Учеб. пос. – М.: Логос, 2002. – 432 с.
2. *Талызина Н.Ф.* Теоретические основы контроля в учебном процессе. – М.: Знание, 1983. – 96 с.
3. *Аванесов В.С.* Основы научной организации педагогического контроля в высшей школе: Учеб. пос. – М.: МИСиС, 1987. – 168 с.
4. <http://www.znannya.org>. – Портал знань.
5. *Дружинин В.Н.* Экспериментальная психология: Учеб. для вузов. – 2-е изд., доп. – СПб.: Питер, 2003. – 320 с.
6. *Аванесов В.С.* Теория и методика педагогических измерений (материалы публикаций). – М.: ЦТ и МКО УГТУ-УПИ, 2005. – 98 с.
7. *Лазарсфельд П.* Латентно-структурный анализ и теория тестов // Мат. методы в социальных науках. – М.: Прогресс, 1973. – С. 42–53.
8. *Lord F.M.* Application of Item Response Theory to Practical Testing Problems. Hillsdale N-J. Lawrence Erlbaum Ass., 1980. – 266 p.
9. *Клайн П.* Справочное руководство по конструированию тестов: Введение в психометрическое проектирование / Пер. с англ.; Под ред. Л.Ф. Бурлачука. – К.: ПАН Лтд., 1994. – 288 с.
10. *Акофф Р., Эмери Ф.* О целеустремленных системах / Пер. с англ.; Под ред. И.А. Ушакова. – М.: Сов. радио, 1974. – 272 с.
11. *Шторм Регина.* Теория вероятностей. Математическая статистика. Статистический контроль качества. – М.: Мир, 1970. – 368 с.
12. *Володарський Є.Т., Кошева Л.О.* Статистична обробка даних: Навч. посіб. – К.: НАУ, 2008. – 308 с.