

10. Yen C.N., Carey W. Application of synthetic aperture processing to towed array data // JASA. – 1989. – Vol. 86. – P. 754–765.
11. Горбань И.И. Оптимальная пространственно-временная обработка сигналов в условиях движения антенны // Радиотехника. – 1990. – № 12. – С. 49–51.
12. Горбань И.И. Оптимальная пространственно-временная обработка шумового сигнала в условиях движения антенны // Радиотехника. – 1991. – № 11. – С. 24–27.
13. Горбань И.И. Оптимальная пространственно-временная обработка сигналов для подвижных антенн при воздействии локальных помех // Радиотехника. – 1993. – № 7. – С. 41–44.
14. Горбань И.И. Быстрые алгоритмы многоканальной пространственно-временной обработки сигналов с подавлением локальной помехи // Радиоэлектроника. – 1994. – № 4. – С. 9–13.
15. Калюжный А.Я. Лучевое представление поля эхолокационных сигналов в слоисто-неоднородной среде при сложном движении приемно-излучающих антенных систем // Акустический журн. – 1996. – Т. 40. № 4. – С. 599–605.
16. Калюжный А.Я. Оптимизация пространственно-временной обработки эхолокационных сигналов в условиях сложного движения приемно-излучающих антенных систем // Акустический журн. – 1996. – Т. 42. № 1. – С. 70–75.
17. Gorban I.I. Estimation of signal parameters when antenna is complicately moving in space // Belgium, Forum Acusticum. – 1996. – P. 222.
18. Gorban I.I. Space-time signal processing algorithms for moving antennas // JASA. – 1996. – Vol. 100. № 4. – Pt. 2. – P. 2638.
19. Gorban I.I. New approaches in optimization of space-time signal processing in hydroacoustics. Course notes of the tutorial. Oceans'98 // IEEE. – 1998. – 69 p.
20. Калюжный А.Я. Методы пространственно-временной обработки гидролокационных сигналов в сложных сигнально-помеховых условиях при наличии дестабилизирующих факторов: Дис. ... д-ра физ.-мат. наук: 01.04.06. – К., 1996. – 348 с.

Стаття надійшла до редакції 18.11.02.

050p-38 6623.51+4481.2846 623.51

УДК 658.336:007:681.3.06

НАДЕЖНОСТЬ ТЕСТОВОГО ЗАДАНИЯ, ДИАГНОСТИКА ПЕДАГОГИЧЕСКАЯ
ОЦЕНКА ЗНАНИЙ, МЕТОД ЭКСПЕРТНЫХ ОЦЕНОК

Н.М. Шибицька, канд. техн. наук, доц.
(Національний авіаційний університет)

МЕТОДИКА ДОСЛІДЖЕННЯ ВАЛІДНОСТІ ТЕСТОВИХ ЗАВДАНЬ

Проведено дослідження способів визначення надійності тестового контролю з урахуванням різних методик оцінювання. Розглянуто поняття валідності, яке визначає достовірність методу педагогічних вимірювань, що має важливе значення під час організації тестового контролю знань. Результати досліджень обґрунтовано практичним використанням методики.

Для поняття “тестування” існує досить багато визначень психологів. За визначенням К. Інгенкампа [1], тестування – це метод педагогічної діагностики, який повинен максимально відповідати основним критеріям, що визначають якість методу:

- об’єктивність, коли метод вимірів вважається об’єктивним, а результати вимірів максимально незалежні від методу;
- надійність, яка визначається тим, наскільки збігаються результати вимірів, що проведені в однакових умовах;
- валідність або достовірність методу;
- точність, що визначається як мінімально припустима помилка методу:

$$\varepsilon = \sigma / \sqrt{NM},$$

де σ – середньоквадратичне відхилення; N – обсяг вибірки; M – математичне сподівання.

Проведемо аналіз валідності тестового завдання з урахуванням різних методик оцінювання [2–4]:

- методика 1: система здійснює “жорстке” порівняння відповіді того, кого навчають, з еталоном (двобальна шкала);

- методика 2: система пропонує один із можливих варіантів відповідей, звертає до ефекту “узнавання” (двобальна шкала);
- методика 3: система пропонує можливі варіанти відповідей після порівняння відповіді з еталоном (рангова шкала);
- методика 4: система здійснює експертне оцінювання з урахуванням відносних ваг питань і відповідей (нечітка шкала [0,1]).

Основними параметрами, що визначають валідність методу, є [1; 5; 6]:

- складність: відповідність середніх результатів тестування середнім значенням репрезентативної вибірки;
- розподільна спроможність: можливість достовірно розділяти тих, що опитуються з різним рівнем знань;
- нормальний закон розподілу.

Коефіцієнт k_r розподільна спроможність тестового завдання (для нормального тесту $k_r = 1$) визначають за формулою:

$$k_r = \exp \left\{ \frac{\left| \bar{\mu} - \frac{\mu_{\max} + \mu_{\min}}{2} \right|}{2\sigma} \right\},$$

де $\bar{\mu}$ – середнє значення результатів тестування; μ_{\max} , μ_{\min} – максимальне і мінімальне значення результатів тестування; σ – стандартне відхилення розподілу.

Функцію щільності можливості розподілу знаходять за виразом

$$f(\mu, \mu_{sr}, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} * \exp\left(-\frac{(\mu - \mu_{sr})^2}{2\sigma^2}\right),$$

де μ – тестування, для якого будується розподіл; μ_{sr} – середньоарифметичний розподіл.

Асиметрія A_s визначається скісністю функції розподілу і буває лівоскошеної (рис. 1) і правоскошеної (рис. 2).

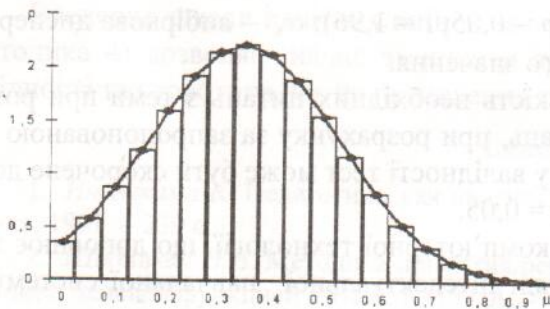


Рис. 1. Функція щільності розподілу за методикою 1

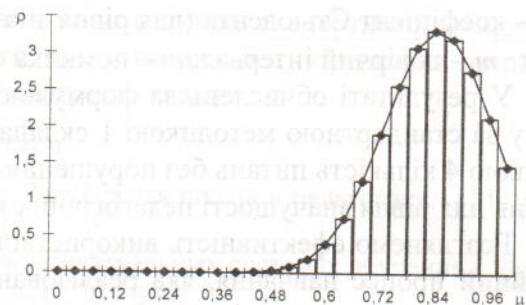


Рис. 2. Функція щільності розподілу за методикою 2

Вибірка розподілена за нормальним законом, і тест вважається валідним, якщо $|A_s| \leq 0,2$ (рис. 3).

Розподільну спроможність визначають за ексцесом E_x . Для функцій розподілу плосковерхих і гостроверхих ексцесів E_x – негативний. Ексцес вважають незначним, якщо $E_x \leq 0,5$.

На рис. 4 наведено порівняльний аналіз функцій щільності розподілу за різними методиками.

Коефіцієнт складності k_c тестового завдання можна визначити як міру відхилення середніх значень результатів від середини нормального закону розподілу:

$$k_c = \frac{\left| \bar{\mu} - \frac{\mu_{\max} + \mu_{\min}}{2} \right|}{2\sigma} A_s,$$

де A_s – асиметрія.

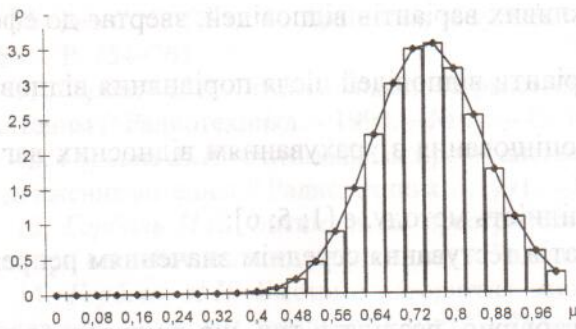


Рис. 3. Функція щільності розподілу за методикою 3

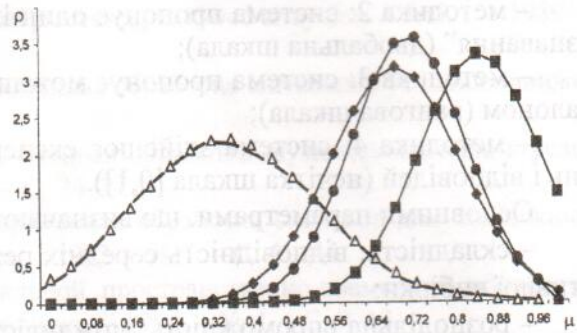


Рис. 4. Порівняльний аналіз функцій щільності розподілу:

—△— методика 1; —■— методика 2; —●— методика 3; —◆— методика 4

Коефіцієнт складності k_c тестового завдання набуває значення в діапазоні $[0, 1]$, для нормально розподіленого тесту $k_c = 1$.

Відповідно до теорії інформації один знак російської або української мови (буква, цифра) містить символічну інформацію 1,19 біт/знак. На підставі цих даних розраховують інформаційну ємність сторінки або цілого підручника.

У роботі [5] наведено визначення достатньої кількості тестових питань для виявлення знань тих, яких навчають. Традиційно тестові опитування оцінюють за двобальною системою 0 або 1, що містять по 1 біту інформації. Допускаючи, що одне питання з варіантами відповідей містить 0,5 КВ, для виявлення знань, які відповідають вивченому за підручником, необхідно задати 2500 питань, із припустимою в педагогіці помилкою вимірів 5%.

При такому підході значно зростають витрати часу, що виділяється до контролю знань, у порівнянні з долею часу, що виділяється на процес закріплення знань на практиці.

Для результатів, розподілених за нормальним законом, довжину тесту визначають за формулою
$$N = t^2 \sigma_x^2 / \Delta^2, \quad (1)$$

де t – коефіцієнт Стюдента (для рівня значущості $p = 0,05$; $t = 1,96$); σ_x^2 – вибіркова дисперсія; $\Delta = t \cdot m$ – довірчий інтервал; m – помилка середнього значення.

У результаті обчислень за формулою (1) кількість необхідних питань з теми при розрахунку за стандартною методикою 1 складає 90 питань, при розрахунку за запропонованою методикою 4 кількість питань без порушення принципу валідності тест може бути скорочене до 20 питань для рівня значущості педагогічних вимірів $p = 0,05$.

Розглянемо ефективність використання нової комп'ютерної технології, що доповнює традиційний процес навчання, яка реалізована у вигляді інтелектуальної навчальної системи з функціями декларування і тестування знань із дисципліни.

Для вирішення задачі оцінювання скористаємося не абсолютними середніми значеннями величин до і після експерименту, а частотними характеристиками величин для визначення χ^2 -критерію:

$$\chi^2 = \sum_{K=1}^m (V_K - P_K)^2 / P_K,$$

де m – загальне число груп, на які розділилися результати спостережень; V_K – частоти результатів спостережень, зроблених після експерименту, %; P_K – частоти результатів спостережень до експерименту, %.

У процесі проведення експерименту зібрано дані про успішність операторів ПК загальною кількістю біля 2000 осіб. Для статистичної обробки виділено чотири класи даних залежно від успішності тих, яких навчають.

До впровадження комп'ютерної методики в процес навчання успішність розподілилася так:

– 55% “задовільно”;

- 30% “добре”;
- 5% “відмінно”;
- 5% “незадовільно”.

Експертна система навчання декларує знання в зручній формі і дозволяє на етапі тестового контролю оцінити рівень навчання оператора, завдяки чому педагог звільняється від рутинної роботи перевірки елементарних знань із теми дисципліни.

Після впровадження комп'ютерної методики навчання і тестування студентів із розділів і тем дисципліни було отримано такі дані, що характеризують успішність:

- 38% “задовільно”;
- 50% “добре”;
- 10% “відмінно”;
- 2% “незадовільно”.

Проаналізуємо дані, отримані в ході експерименту, з метою дослідження ефективності запропонованої методики 4 комп'ютерного навчання в порівнянні з методикою 1 традиційного навчання (рис. 5).

Частоти результатів спостережень розподілилися в такій послідовності:

- до експерименту: 55, 30, 5, 5 %;
- після експерименту: 38, 50, 10, 2 %.

кількість груп, на які розділилися результати спостережень, дорівнює чотирьом.

Таким чином,

$$\chi^2 = (38-55)^2/55 + (50-30)^2/30 + (10-5)^2/5 + (2-5)^2/5 = 41,78.$$

Порівняємо знайдене значення критерію χ^2 з табличним. Для трьох степенів вільності і з вірогідністю припустимої помилки 0,1% табличне значення критерію $\chi^2 = 16,27$.

Отже, запропонована методика навчання експериментально підтвердилася з припустимою помилкою не більш ніж 1%.

Введення шкали нечітких оцінок за запропонованою методикою експертного оцінювання (методика 4) дозволяє значно зменшити кількість тестових питань без порушення принципу валідності та скоротити час на проведення тестового контролю знань.

Список літератури

1. *Ингенкамп К.* Педагогическая диагностика // Сер. Зарубежная школа и педагогика. – М.: Педагогика, 1991. – 239 с.
2. *Шибицкая Н.Н.* Методика принятия решений в интеллектуальных системах обучения с нечетко заданной целевой функцией // Вісн. КМУЦА. – 1998. – №1. – С. 307–314.
3. *Кудиненко А.В., Шибицкая Н.Н.* Проблемные вопросы управления в эргатических системах обучения // Кибернетика и вычислительная техника. – 1998. – Вып. 116. – С. 29–41.
4. *Шибицкая Н.Н.* Метод идентификации объектов в эргатической системе управления процессом обучения // Кибернетика и вычислительная техника. – 1998. – Вып. 121. – С. 52–58.
5. *Булах І.Е.* Комп'ютерна діагностика навчальної успішності. – К.: ЦМК МОЗ України; УДМУ, 1995. – 221 с.
6. *Свиридов А.П.* Основы статистической теории обучения и контроля знаний. – М.: Высш. шк., 1989. – 262 с.

Стаття надійшла до редакції 02.09.02.

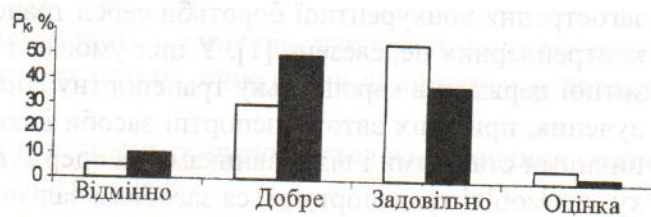


Рис. 5. Ефективність методик:
□ – методика 1; ■ – методика 4