

Ципфа, закон розсіяння С.К. Бредфорда тощо. Огляд основних бібліометричних законів представлений у праці [3].

Кожен запис ББД та відповідне джерело інформації аналізуються на основі певного бібліометричного закону. У випадку відхилення від певних бібліографічних правил бібліографічний запис потребує перевірки редактором ЕК.

Висновки. У роботі піднімається актуальна проблема якості даних у ББД ЕКБ. Запропоновано підходи до пошуку логічних помилок та протиріч у записах ББД. Авторські підходи ґрунтуються на протиріччі між значеннями атрибутів та пошуку помилок заснований на бібліометричних законах. У подальшому робота буде проводитись над практичної реалізацією запропонованих методів. Вищезазначені методи будуть реалізовані у підсистемі пошуку протиріч СВДЕК.

Література

1. Ярмолюк Р.С. Основні типи та джерела помилок у записах електронного каталогу / Р.С. Ярмолюк // Вісник Національного Університету «Львівська політехніка» Інформаційні системи та мережі, - 2010. - № 689. – С. 348-357.
2. Ярмолюк Р.С. Структурно-функціональна модель верифікації даних електронного каталогу / Р.С. Ярмолюк // Суч. пробл. діяльн. бібл. в ум. інф. сусп. м. 3-ої наук.-прак. конф., 29 вересня 2011р., Львів/ Національний університет "Львівська політехніка"; - Львів : Видавництво НУ "Львівська політехніка", 2011. – С. 217-224.
3. Бредихин С.В. Методы библиометрии и рынок электронной научной периодики / С.В. Бредихин, А.Ю. Кузнецов – Новосибирск, Москва, 2012. – С. 255.

УДК 681.518.5

Олександр Шкіль, Дмитро Коротких
Харківський національний університет радіоелектроніки

АНАЛІЗ НАДІЙНОСТІ ТЕСТІВ В КОМП'ЮТЕРНІЙ СИСТЕМІ ТЕСТУВАННЯ ЗНАНЬ

© Олександр Шкіль, Дмитро Коротких, 2013

Розглянуто питання аналізу надійності тестів з використанням коефіцієнту α -Кронбаха для розріджених матриць результатів тестування. Запропоновано метод доповнення розріджених матриць та методика розрахунку статистичних показників з використанням табличного процесора Microsoft Excel.

Ключові слова – тестові завдання, надійність тестів, дисперсія, кореляція.

The question of the tests security analysis with the use of the α -Cronbach coefficient for the disperse matrix of the testing results is considered. The method of the disperse matrix addition and method of the statistical indexes calculation with the use of the tabular processor Microsoft Excel are offered.

Keywords – test tasks, reliability of tests, dispersion, correlation.

Вступ. Одним з важливих показників якості тесту є його надійність. Під надійністю тесту розуміється стійкість або узгодженість результатів, одержаних при повторному його застосуванні до тих же випробуваних у різні моменти часу, при використанні різних наборів еквівалентних завдань або ж при зміні інших умов випробування. Коефіцієнт надійності - двовимірна описова статистика, кількісна міра варіативності (спільного розсіяння) двох змінних. В залежності від кількості повторних тестувань, шкал оцінювання та упорядкування результатів розрізняють такі коефіцієнти надійності: ретестової, Спирмена–Брауна, Гутмана, Рулона, Кьюдера–Ричардсона та α –Кронбаха. В якості чисельної міри надійності при аналізі результатів тестування в системі OpenTEST2 використовується коефіцієнт α –Кронбаха, який дозволяє оцінити надійність тесту за результатами одного тестування за наявності серед результатів частково правильних відповідей і не вимагає спеціального упорядкування матриці результатів тестування. Діапазон значень α –Кронбаха : для високонадійних тестів $\alpha > 0,9$, для тестів задовільної надійності $\alpha > 0,8$. Якщо $\alpha < 0,6$ надійність тесту викликає сумніви.

Вихідними даними для розрахунку надійності тесту є матриця результатів тестування (МРТ), яка являє собою прямокутну матрицю, розмірністю $\{N \times (K + 1)\}$. Рядками матриці $\{N\}$ є результати тестування випробуваних, для зручності подальшого аналізу впорядковані у порядку убудування сумарної оцінки за тест, а стовпцями матриці $\{K\}$ є тестові завдання, які використовувалися при тестуванні даної групи (вибірки випробуваних). Останній стовпець представляє собою суму балів, набраних кожним випробуваним при проходженні тесту. В комірках матриці знаходяться бали, отримані кожним випробуваним за відповіді на тестові завдання $(B_{i,j})$. Для недихотомічної шкали оцінювання, яка враховує частково правильні відповіді, використовується загальна формула обчислення α –Кронбаха [1] :

$$\alpha = \frac{K}{K-1} \left(\frac{D_y - \sum_{i=1}^K D_i}{D_y} \right) = \frac{K}{K-1} \left(1 - \frac{\sum_{i=1}^K D_i}{D_y} \right),$$

де K -число стовпців МРТ, D_i – дисперсія по i -ому завданню (за відповідним стовпцем МРТ), D_y – дисперсія по тесту в цілому,

Коефіцієнт $\frac{K}{K-1}$ необхідний для врахування кінцевого числа елементів вибірки (K) и при $(K \rightarrow \infty)$, він прагне до 1.

Аналіз розріджених матриць результатів тестування. Особливістю проведення тестування в системі OpenTEST2 є те, що для кожного випробуваного формується свій сеанс тестування, тобто індивідуальна репрезентативна вибірка тестових завдань з тестової бази даних [2]. Отже, кожен випробовуваний відповідає не на всі завдання тесту, а тільки на M завдань зі свого сеансу тестування. Таким чином, одержана МРТ є розрідженою, тобто в деяких комірках немає результату, тому що даний випробуваний відповідного завдання не отримував. При цьому відмова від відповіді прирівнюється до невірної відповіді.

Визначимо коефіцієнт заповнення МРТ: $F = \frac{M * N}{K * N} = \frac{M}{K}$.

Для спрощення подальших розрахунків приймемо, що вага всіх завдань однакова і дорівнює 1, а кількість використань кожного завдання в тестуваннях однаково. Остання умова досяжна, якщо довжина сеансу тестування кратна числу тем тесту, а кількість завдань в кожній темі тесту однакова. В (K+1) стовпці МРТ наведені бали, отримані кожним з випробуваним в сеансі,

$B_i = \sum_{j=1}^M B_{i,j}$, $i = \overline{1, N}$. Оцінка для кожного випробуваного визначається як набраний бал за всі завдання сеансу тестування, віднесена до кількості завдань

в сеансі тестування (M): $Y_i = \frac{B_i}{M} * 100\%$. Таким чином, Y_i являє собою частку

(відсоток) правильних відповідей на завдання сеансу тестування або оцінку за 100-бальною шкалою оцінювання. При цьому $Y_i = B_i * \frac{100}{M} = B_i * H$, де H –

коефіцієнт перерахунку набраного балу у 100-бальну шкалу. Для реальних сеансів тестування ($M < 100$) коефіцієнт $H > 1$. Аналіз різних способів доповнення порожніх комірок МРТ будь-якими постійними значеннями відносно стовпців показав, що, незалежно від способу заповнення порожніх комірок, діапазон результуючих тестових балів звужується в $(1 / F)$ раз, що в свою чергу зменшує D_B і знижує значення коефіцієнта α –Кронбаха (навіть до негативних значень).

Альтернативним способом доповнення МРТ є доповнення по рядках. Для цього зробимо ряд припущень.

1. Якщо випробуваний в сеансі з M завдань отримав відповідну оцінку, то можна припустити, що і для решти (K-M) завдань тесту (стовпців МРТ) він міг би отримати такі ж оцінки.

2. Виходячи з того, що M завдань сеансу тестування обираються з тестової бази випадково, але в кожному сеансі є завдання різної трудності, то в цілому розподіл трудностей завдань в усіх сеансах (для всіх випробуваних) є приблизно однаковим.

Виходячи з цих припущень, запропоновано доповнювати МРТ по рядках середніми значеннями набраних балів по кожному рядку (середнім балом, набраними, кожним випробуваним за сеанс, тобто $\delta = \frac{B_i}{M}$). Підсумкова оцінка

доповненого рядку буде :

$$Y_i = \frac{B_i + \frac{B_i}{M} \cdot (K - M)}{K} \cdot 100\% = \frac{B_i + \frac{B_i \cdot K}{M} - B_i}{K} \cdot 100\% = \frac{\frac{B_i \cdot K}{M}}{K} \cdot 100\% = \frac{B_i}{M} \cdot 100\% .$$

При такому способі заповнення порожніх комірок підсумкова оцінка за тест для i -го випробуваного Y_i не зміниться, отже, і не зміниться і діапазон підсумкових оцінок в цілому за тест (розмах вибірки) $R = Y_{\max} - Y_{\min}$, і заниження α -Кронбаха не буде.

Якщо виходити з того, що завдання в тесті якісні, що підтверджує висока диференціююча здатність кожного завдання і достатня кореляція кожного завдання (стовпця МРТ) з результатами по тесту в цілому (стовпця V), то таке доповнення стовпців істотно не змінює а ні кореляцію, а ні дисперсію стовпців МРТ. Таким чином, коефіцієнт α -Кронбаха, розрахований за МРТ, доповненою середніми значеннями результатів по рядках, в цілому відображає характер розподілу і варіацію тестових оцінок i , як наслідок, відображає реальну надійність тесту. При цьому слід враховувати, що наведена методика дає достатньо достовірні результати при $F > 0,5$. При $F < 0,5$ достовірність розрахованого за цією методикою α -Кронбаха значно знижується.

Розрахунок статистичних показників з використанням MS Excel. Особливістю будь-якої клієнт-серверної технології, в тому числі і комп'ютерного тестування під керуванням єдиного тестового серверу, є необхідність наявності постійного зв'язку з тестовим сервером. При проведенні тестування наявність такого зв'язку є необхідною умовою, оскільки реалізація алгоритму тестування та оцінювання здійснюється на стороні серверу. Розрахунок статистичних показників характеризується великою кількістю арифметичних обчислень при мінімальних зверненнях до бази даних. При розрахунку різних статистичних показників наявність постійного зв'язку з сервером не є обов'язковою умовою. Для зменшення обчислювального навантаження на тестовий сервер достатньо на клієнтський комп'ютер один раз передати результати тестування (наприклад, МРТ) та всі розрахунки вести безпосередньо на клієнтському комп'ютері. При цьому формат передачі даних

має бути стандартизованим, щоб користувач не витрачав час на переформатування даних та не створював власних конверторів. Стандартним форматом передачі табличних даних між сервером і клієнтом є CSV-формат при використанні в якості клієнтської програми табличного процесора Microsoft Excel MS Office. Статистична обробка МРТ з використанням табличного процесора MS Excel має ряд переваг, а саме, можливість обробки даних без зв'язку з сервером, реалізація складних процедур обробки статистичних даних з використанням макросів Excel на мові MS Visual Basic, можливість створення індивідуальних процедур обробки, підключення зовнішніх програм обробки статистики, видалення завідомо недостовірних результатів тестування тощо.

На підставі вищенаведених розрахунків розроблена власна процедура доповнення та аналізу МРТ. На рис.1 показано екранну форму відображення МРТ в MS Excel з розрахунками трудностей, кількості використань, кореляції за Пірсоном та надійності тестів для доповненої матриці. При натисканні зеленої кнопки порожні клітинки (комірки) доповнюються середніми оцінками по кожному рядку (перший стовпець), та знову розраховуються трудність (Diff), дисперсія даного стовпця (D_new) і кореляція доповненого стовпця з результатами по тесту (Rxy_new). Крім того в останньому рядку наводиться розраховане значення уточненого коефіцієнта надійності α -Кронбаха. (a_modify).

	А	В	С	Д	Е	Ф	Г	Н	І	Ј	К	Л	М
1													
2													
3		1	2	3	4	5	6	7	8	9	10	Сумма_X по теме	X*2
4	100	1	1	1	1	1	1	1	1	1	1	10	100
5	80	0,8	0,8	0,8	0,8	1	1	1	1	0	0,8	8	64
6	80	0,8	0,8	0,8	0,8	0,8	0	1	1	1	1	8	64
7	60	0	1	0,6	1	0	0,6	0,6	0,6	0,6	1	6	36
8	60	0	0,6	1	1	0,6	0,6	1	0	0,6	0,6	6	36
9	60	0,6	0	0	0,6	0,6	1	0,6	0,6	1	1	6	36
10	60	0	1	0,6	0,6	0,6	1	0	0,6	0,6	1	6	36
11	40	0,4	0	1	0,4	1	0,4	0,4	0	0	0,4	4	16
12	20	0	0,2	0	1	0	0,2	0,2	0,2	0	0,2	2	4
13	20	0,2	0,2	0,2	0	0,2	0	1	0	0,2	0	2	4
14													
15	Сумма	3,8	5,6	6	7,2	5,8	5,8	6,8	5	5	7	58	396
16	N	10	10	10	10	10	10	10	10	10	10	10	
17	Diff	0,38	0,58	0,6	0,72	0,58	0,58	0,68	0,5	0,5	0,7	5,8	
18	SSx	1,396	1,584	1,44	0,976	1,396	1,556	1,296	1,62	1,62	1,3	14,184	59,6
19	D_old	0,14	0,158	0,144	0,098	0,14	0,158	0,13	0,162	0,162	0,13	1,418	5,96
20	D_new												
21	Rxy_old	0,897	0,882	0,804	0,503	0,653	0,577	0,428	0,835	0,831	0,841		
22	Rxy_new												
23	n	10											
24	a_modify	0,847											

Рис. 1. Екранна форма обробки розрідженої МРТ в MS Excel

Для даного варіанту тесту та доповнення порожніх комірок $\alpha = 0,847$, що є цілком задовільним значенням. Запропонований підхід дозволяє проводити аналіз якості тестових завдань та надійності тестів в автономному режимі без завантаження тестового сервера, що істотно підвищує оперативність апробації та коригування тестів.

Література

1. Анастаси А. Психологическое тестирование: пер. с англ. / А. Анастаси, С. Урбина. – 7-е международ. издание. – СПб.: Питер, 2003 – 688 с.
2. Шкиль А., Каук В., Напрасник С., Цимбалюк Е., Хоменко Р. Новые функциональные возможности компьютерной системы тестирования знаний OPENTEST2 // Педагогические измерения. – 2009. – № 2. – С.86–103.

Євген Міюшкович, Юлія Міюшкович
Національний університет «Львівська політехніка»

МЕТОДОЛОГІЯ РЕДАГУВАННЯ ПАРАМЕТРІВ ТЕСТОВИХ ЗАВДАНЬ В ФОРМАТІ MOODLE XML

© Євген Міюшкович, Юлія Міюшкович, 2012

Розглянуто методологію редагування тестових завдань за допомогою параметрів тегів Moodle XML.

Ключові слова: Moodle XML, формат, тест, редагування, параметр, тег, атрибут.

The work demonstrates methodology of tests editing with the Moodle XML tag options.

Keywords: Moodle XML, format, test, edit, option, tag, attribute.

Вступ. Стандартні засоби системи Moodle надають можливість редагувати тільки одне тестове питання в один момент часу. Редагування великої кількості тестових питань, розміщених в банку питань курсу, створює необхідність розробки методології редагування тестових завдань з мінімальними затратами часу.

Система Moodle передбачає експорт тестових завдань в чотирьох форматах: GIFT, IMS QTI 2.0, Moodle XML, XHTML. Розглянемо особливості побудови Moodle XML формату.

Аналіз структури формату Moodle XML. Moodle XML – специфічний формат імпорту/експорту тестових питань, побудований за принципами розширюваної мови розмітки XML. Даний формат був розроблений Говардом Міллером (Howard Miller) [1] і використовується в модулі «Тест» системи Moodle.