

Було проведено аналіз предметної галузі документування наукових досліджень в НАПН України, описано перелік вихідних документів, за яким разом із зазначеними вище переліком операцій з документами визначають інформаційну модель СЕД, на основі якої розробляється модель даних СЕД.

Формування такої моделі СЕД було здійснено за такою схемою:

1. Аналіз документів з документування наукових досліджень в НАПН України.
2. Формування переліку *спільних полів* цих документів, тобто полів, які використовуються не тільки в одному документі.
3. Ідентифікація полів.
4. Визначення документа-джерела кожного поля.
5. Визначення множини документів, де використовується кожне поле.
6. Визначення механізму первинного заповнення кожного поля.
7. Формування списків для полів, значення яких фіксовані і визначаються за відповідним списком.

Автором описано типи контенту мовою XML, та розроблено засіб синхронізації спільних полів у різних типах документів, процедури для роботи з таблицями в документах MS Word з використанням засобів OpenXML.

Таким чином для впровадження системи електронного документообігу в наукову установу рекомендується платформа SharePoint. Головними перевагами є безкоштовність, закритий код платформи, можливість удосконалення за допомогою мови програмування C# та безкоштовний програмний продукт SharePoint Designer, звичне для наукового працівника середовище роботи MS Office. Перспективним є розгортання «хмарної» версії ІС «Наукові дослідження» на платформі Microsoft Office 365, до складу якої включено оновлені версії SharePoint (2010 та 2013).

Список використаних джерел

1. Задорожна Н.Т. Документування НДР з використанням інформаційної системи «Наукові дослідження: Планування, контроль, моніторинг». Методичні рекомендації / Н.Т. Задорожна, А.В. Кільченко, Х.В. Серета, С.М. Тукало та ін. // Ін-т інформаційних технологій і засобів навчання НАПН України. – Київ, 2014. – 75с. – Бібліогр.: 28 назв. – Укр. – Деп. в ДНТБ України

2. Положення про порядок планування і контролю виконання наукових досліджень у Національній академії педагогічних наук України [Затверджено: Постановою Президії НАПН України від 23 червня 2011 року, протокол № 1-7/9-198 із змінами, внесеними Постановою Президії НАПН України від 20 грудня 2012 року, протокол № 1-7/14-403] – 38 с. [Електронний ресурс]. – Режим доступу : <http://www.twirpx.com/file/740206/>.

3. Задорожна Н.Т. Методологія інформатизації наукової та управлінської діяльності установ НАПН України на основі веб-технологій / Н. Т. Задорожна, Т. В. Кузнецова, А. В. Кільченко, Х. В. Серета, С. М. Тукало, О. О. Каплун, Л. А. Лупаренко. – К.: Атіка, 2014. – 160 с.

4. DOCFLOW Україна: все об електронном документообороте, бизнес-конференция [Електронний ресурс] – Режим доступу: <http://www.docflow.ua/conference/>. – Назва з екрану.

5. Задорожна Н.Т. Інформаційна система менеджменту наукових досліджень в НАПН України / Н. Т. Задорожна, В. А. Петрушко, С. М. Тукало // Інформаційні технології в освіті: Збірник наукових праць. Випуск 15, Херсон : Видавництво ХДУ, 2013. – С. 129-137.

Диховичний О.О., Дудко А.Ф.
НТУУ «КПІ»

Комплексна методика аналізу якості тестів з вищої математики

Сьогодення характеризується вираженою інформатизацією суспільства. Це зумовлює необхідність активного використання інформаційно-комунікаційних технологій в освіті для оцінювання знань студентів. Так в НТУУ «КПІ» кафедрою математичного аналізу та теорії ймовірностей організовано комп'ютерне тестування з використанням створеного викладачами кафедри комплексу дистанційних курсів «Вища математика» [1].

Розширення кола студентів, охоплених тестуванням, вимагає розширення бази тестових завдань. Це, в свою чергу, призводить до необхідності аналізу якості тестів. Основу аналізу традиційно складають статистичні методи, а саме методи КТТ або ІРТ.

Методика аналізу якості тестів з вищої математики розроблена на підставі досвіду застосування методів класичної теорії тестів та ІРТ до результатів тестування студентів НТУУ «КПІ». Загальна схема використання методики зображена на рис. 1 і включає в себе наступні етапи:

- формування таблиці результатів;
- аналіз окремих тестових завдань і тесту в цілому;
- приймання остаточних рішень.

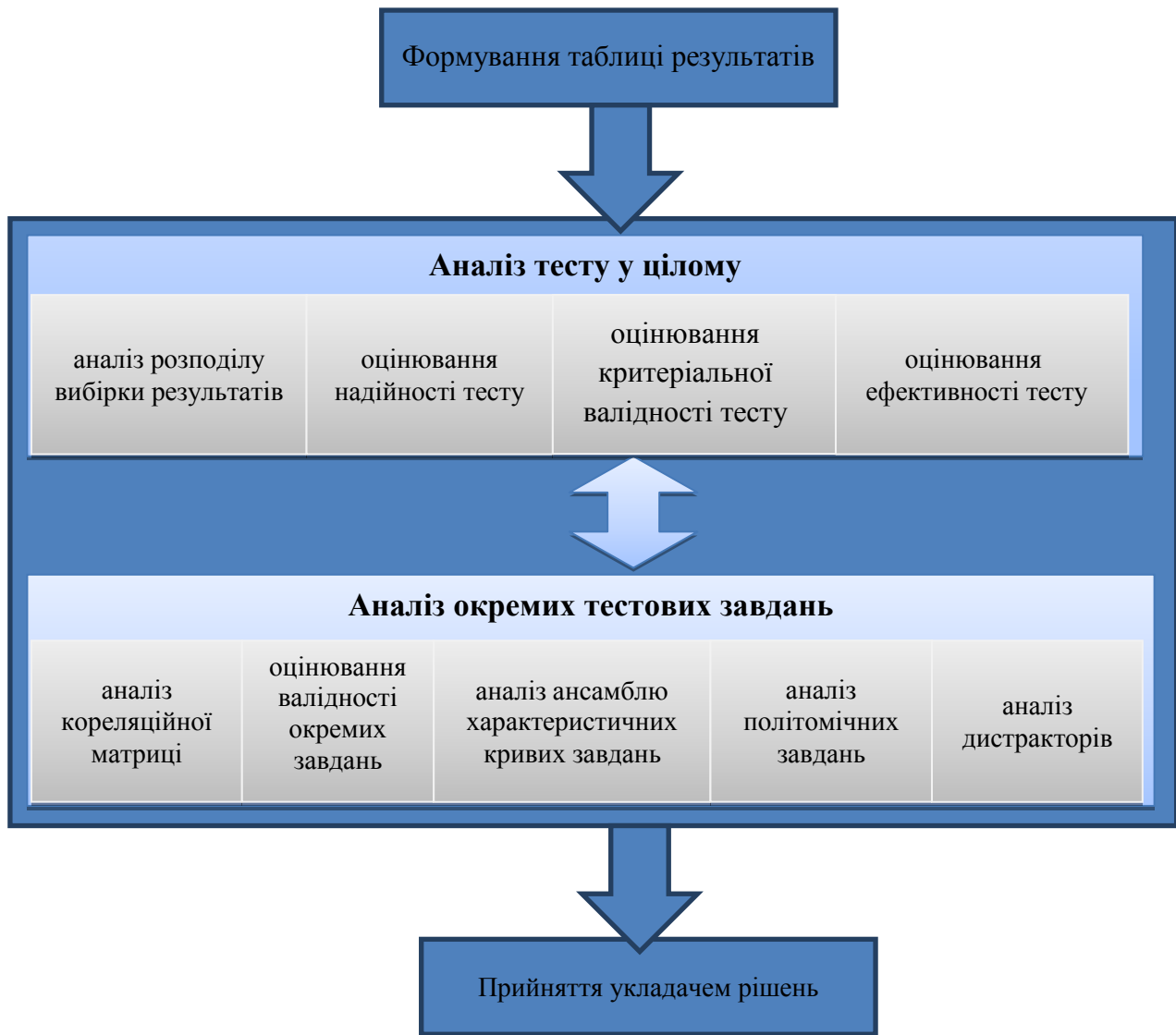


Рис. 1

На **першому етапі** здійснюється формування таблиці результатів тестування (рис. 2). З таблиці виокремлюються всі рядки, в яких містяться лише нулі або лише найвищий бал за завдання. Обчислюються індивідуальні бали студентів X_i , $i = \overline{1, N}$, де N - кількість студентів, і загальні бали за завдання R_j , $j = \overline{1, K}$, де K - кількість завдань тесту. Далі всюди у формулах використовуються ці позначення. Здійснюється впорядкування матриці результатів тестування спочатку в порядку спадання значень X_i потім – значень R_j .

ПІБ іспитника i	Номер завдання j															X_i
	№1	№2	№3	№4	№5	№6	№7	№8	№9	№10	№11	№12	№13	№14	№15	
Артем Юрійович Лашко	0,27	0,4	0,53	0,53	0	0,53	0	0,53	0,53	0	0,53	0,53	0,53	0,53	0,53	5,97
Рустам Олегович Срохін	0,27	0	0,53	0	0	0	0,53	0	0,53	0	0	0	0,53	0,53	0	2,92
Олександр Євгенійович Хорунжий	0,27	0,53	0,53	0,53	0	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	7,16
Микита Миколайович Кузімович	0,4	0,53	0,53	0	0	0	0	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	5,7
Дар'я Вікторівна Кравець	0,4	0,53	0,53	0	0	0	0	0,53	0,53	0	0,53	0,53	0,53	0	0	4,11
Андрій Геннадійович Захарчук	0,27	0,53	0,53	0	0	0,53	0	0	0	0,53	0,53	0,53	0,53	0,53	0,53	5,04
Марія Олександрівна Ігнатовська	0,27	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0,53	0	0,53	7,16
Антон Олександрович Чічков	0,4	0,53	0,53	0,53	0	0,53	0,53	0,53	0,53	0,53	0,53	0	0,53	0	0,53	6,23
...																...
R_j	13,65	16,58	16,96	11,13	5,06	11,13	9,54	15,9	19,61	15,37	20,14	14,31	18,55	13,25	14,31	

Рис. 2

На **другому етапі** проводиться аналіз окремих тестових завдань і тесту в цілому.

Аналіз тесту в цілому полягає у наступному:

- аналіз розподілу вибірки результатів;
- оцінювання надійності тесту;
- оцінювання критеріальної валідності тесту;
- оцінювання ефективності тесту.

Аналіз розподілу вибірки результатів. Здійснюються графічна інтерпретація емпіричних даних у вигляді гістограми, аналіз розподілу вибірки індивідуальних балів та обчислення вибірових характеристик.

Оцінювання надійності тесту. Обчислюється коефіцієнт надійності тесту r , знайдений на основі методу розщеплення, та коефіцієнт надійності Кронбаха α .

Для оцінювання коефіцієнта надійності r впорядковані результати тестування ділять на дві частини: в одну включають дані про студентів за парними номерами, а в іншу – за непарними номерами завдань. Формула для обчислення коефіцієнта надійності r має наступний вигляд:

$$r = \frac{2\tilde{r}}{1 + \tilde{r}},$$

де \tilde{r} – коефіцієнт надійності для половини завдань тесту:

$$\tilde{r} = \frac{N \sum_{i=1}^N Z_i Y_i - \left(\sum_{i=1}^N Z_i \right) \left(\sum_{i=1}^N Y_i \right)}{\sqrt{N \sum_{i=1}^N (Z_i)^2 - \left(\sum_{i=1}^N Z_i \right)^2} \sqrt{N \sum_{i=1}^N (Y_i)^2 - \left(\sum_{i=1}^N Y_i \right)^2}},$$

де Z_i – бал i -го студента за парними номерами завдань тесту, Y_i – бал i -го студента за непарними номерами завдань тесту.

Коефіцієнт надійності Кронбаха α знаходять за формулою

$$\alpha = \frac{L}{L-1} \left(1 - \frac{\sum_{j=1}^L S_j^2}{S_X^2} \right),$$

де S_j^2 – вибіркова дисперсія розподілу балів за j -е завдання; S_X^2 – вибіркова дисперсія розподілу індивідуальних балів.

В якості нижньої межі припустимих значень коефіцієнта надійності обрано 0,7 [2, 336].

Оцінювання критеріальної валідності тесту. Як критерій зазвичай беруться оцінки експертів, виставлені при традиційній перевірці знань учнів без використання тестів. Для оцінювання валідності тесту використовується формула

$$\rho = \frac{\sum_{i=1}^N (X_i - \bar{X})(X_{m_i} - \bar{X}_m)}{N \sqrt{S_X^2 \cdot S_{m_x}^2}},$$

де $X_i - \bar{X}$ – відхилення тестового бала, виставленого i -му студентові, від середнього бала тесту; $X_{m_i} - \bar{X}_m$ – відхилення бала, виставленого i -му студентові у експертів, від середнього арифметичного експертних оцінок; S_X^2 – дисперсія розподілу балів за тестом; $S_{m_x}^2$ – вибіркова дисперсія розподілу балів, виставлених експертами.

При значенні $\rho \geq 0,7$ вважається, що валідність тесту висока.

Оцінювання ефективності тесту. Будується інформаційна функція тесту, оцінюється його ефективність. Визначається інтервал значень параметра підготовленості студента, для яких даний тест є найінформативнішим.

Інформаційна функція тесту в цілому обчислюється за формулою [3, 58]

$$I(\theta) = \sum_{j=1}^K I_j(\theta),$$

де $I_j(\theta)$ – інформаційна функція j -ого завдання; θ – параметр підготовленості студента.

Графік інформаційної функції правильно складеного тесту повинен бути дзвоноподібним, але не занадто загостреної форми. Якщо у інформаційної функції виявляється кілька локальних екстремумів, то тест потребує вдосконалення. При цьому, якщо кількість завдань у тесті невелика, то потрібно додавати завдання проміжної складності, щоб ліквідувати «провали» між сусідніми екстремумами. Якщо кількість завдань у тесті досить велика, то його потрібно поділити на два тести.

Найкращим виявляється тест на тому інтервалі значень параметра θ , на якому інформаційна функція набуває найбільших значень.

Аналіз тестових завдань передбачає:

- аналіз кореляційної матриці завдань;

- оцінку валідності окремих завдань;
- аналіз ансамблю характеристичних кривих завдань;
- аналіз політомічних завдань;
- аналіз дистракторів завдань множинного вибору.

Аналіз кореляційної матриці завдань. Обчислюється кореляційна матриця завдань. Для того, щоб тестові завдання не дублювали одне одного, значення кореляції між ними має бути не занадто високим ($r_{ij} \leq 0,3$). За від'ємної кореляції даного завдання з більшістю інших завдань необхідно переглянути завдання, усунути наявні в ньому помилки, або навіть вилучити з тесту.

Оцінювання валідності окремих завдань. Для оцінювання валідності j -го завдання обчислюються значення індексу розрізняльних характеристик дихотомічних завдань D_j , коефіцієнта точкової бісеріальної кореляції між дихотомічним завданням та індивідуальними балами, виставленими студентам і коефіцієнт кореляції Пірсона між політомічним завданням та індивідуальними балами ρ_j , виставленими студентам.

Значення індексу розрізняльних характеристик дихотомічних завдань знаходять за формулою:

$$D_j = p_{uj} - p_{lj},$$

де p_{uj} – частка правильних відповідей на j -те завдання для 27% студентів з найвищими індивідуальними балами; p_{lj} – частка правильних відповідей на j -те завдання для 27% студентів з найнижчими індивідуальними балами.

На основі практичного досвіду обрано наступні інтерпретації значень D_j : $D_j \geq 0,4$ – завдання вдало дібрано; $0,3 \leq D_j \leq 0,39$ – можливо необхідна деяка корекція завдання; $0,2 \leq D_j \leq 0,29$ – завдання, яке потребує перегляду; $D_j \leq 0,19$ – завдання, яке необхідно вилучити з тесту.

Коефіцієнт кореляції Пірсона обчислюється за формулою

$$\rho_j = \frac{\sum_{i=1}^N (X_i - \bar{X})(X_{ij} - \bar{X}_j)}{\sqrt{\left(\sum_{i=1}^N (X_i - \bar{X})^2\right) \left(\sum_{i=1}^N (X_{ij} - \bar{X}_j)^2\right)}}$$

де $X_i - \bar{X}$ – відхилення тестового бала, виставленого i -му студентові, від середнього бала за тестом; $X_{ij} - \bar{X}_j$ – відхилення тестового бала, виставленого i -му студентові за j -те завдання від середнього бала за j -е завдання.

Формула для обчислення точкової бісеріальної кореляції має вигляд

$$\rho_j = \frac{\bar{X}_{1j} - \bar{X}_{0j}}{S_X} \sqrt{\frac{N_{1j} \cdot N_{0j}}{N(N-1)}}$$

де \bar{X}_{1j} – середнє значення індивідуальних балів студентів, які правильно виконали j -те завдання тесту; \bar{X}_{0j} – середнє значення індивідуальних балів студентів, які неправильно виконали j -те завдання тесту; S_X – стандартне відхилення вибірки індивідуальних балів; N_{1j} – кількість студентів, які правильно виконали j -те завдання тесту; N_{0j} – кількість студентів, які неправильно виконали j -те завдання тесту; N – загальна кількість студентів, $N = N_{1j} + N_{0j}$.

Обчислюється бісеріальний коефіцієнт кореляції [4, 424] між дихотомічним завданням та індивідуальними балами ρ_{bisj} , виставленими студентам, за формулою

$$\rho_{bisj} = \frac{\bar{X}_{1j} - \bar{X}}{S_X} \sqrt{\frac{p_j}{Y_j}}$$

де \bar{X}_{1j} – середнє значення індивідуальних балів, виставлених студентам, які правильно виконали j -те завдання тесту; \bar{X} – середнє значення індивідуальних балів, виставлених студентам; S_X – стандартне відхилення вибірки індивідуальних балів; p_j – частка студентів, які відповіли на j -те завдання правильно; Y_j – ордината стандартної нормальної кривої в точці з абсцисою, яка дорівнює z -оцінці, що відповідає величині p для j -го завдання (визначається за відповідною таблицею).

Для правильно складених завдань мають виконуватись наступні умови [4, 433]:

$$\rho_j \geq \frac{2}{\sqrt{N-1}}; \rho_{bisj} \geq \frac{\sqrt{pq/(N-1)}}{Y}$$

Аналіз ансамблю характеристичних кривих завдань. Поняття характеристичної кривої є одним з центральних понять IRT, в основу якої покладено ідею Г. Раша [5, 105] впровадження наступних латентних параметрів:

- підготовленості студента $\theta_i, i=\overline{1, N}$, де N – кількість студентів;
- складності завдання тесту $\beta_j, j=\overline{1, K}$, де K – кількість завдань в тесті.

Імовірність правильної відповіді i -го студента на j -ге дихотомічне завдання тесту, тобто завдання, виконання якого оцінюється альтернативно: виконане правильно або виконане неправильно, визначається наступним чином:

$$P_{ij} = \frac{1}{1 + \exp(-(\theta_i - \beta_j))}, \quad i = \overline{1, N}, j = \overline{1, K}.$$

Залежність імовірності від неперервного параметра θ при фіксованому значенні β_j називають характеристичною кривою j -го завдання тесту (рис. 3):

$$P_j(\theta) = \frac{1}{1 + \exp(-(\theta - \beta_j))}, \quad j = \overline{1, K}.$$

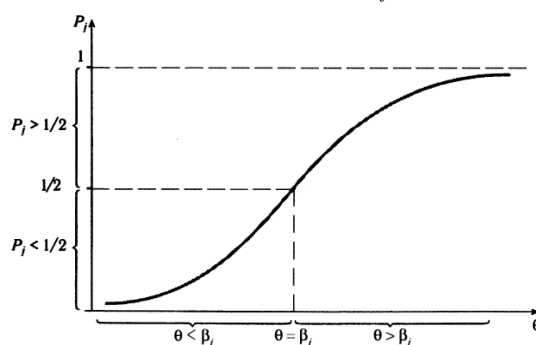


Рис. 3

Перевіряється рівномірність покриття характеристичними кривими завдань інтервалу у межах від -5 до 5 логітів. Якщо рівномірність покриття порушено характеристичними кривими, розміщеними зліва (справа) на графіку, то відповідні завдання мають мінімальне (максимальне) значення складності. Їх необхідно вилучити з тесту, якщо і інші показники цих тестових завдань вказують на неправильність їх складання. Якщо ж прогалина спостерігається в середині між характеристичними кривими, то для її заповнення до тесту необхідно додати завдання проміжної складності.

Аналіз політомічних завдань. Проводиться сумісний аналіз характеристичних кривих рівнів та інформаційних кривих політомічних завдань. Детально такий аналіз описано в роботі [3].

Аналіз дистракторів завдань множинного вибору. Дистрактори, які вибирають менше 5% студентів, повинні бути вилучені або замінені. Аналіз дистракторів завдань множинного вибору проводиться із застосуванням характеристик кривих IRT-моделей множинного вибору.

На **третьому етапі** аналізу на основі сумісного аналізу взаємодоповнюючих характеристик IRT та КТТ робляться висновки щодо якості як окремих тестових завдань, так і тесту в цілому, приймаються остаточні рішення.

На основі аналізу розподілу результатів вибірки, знайдених значень коефіцієнтів надійності та критеріальної валідності тесту, графіка інформаційної функції тесту оцінюється якість тесту в цілому. Якщо суттєво порушено нормальність розподілу вибірки індивідуальних балів, коефіцієнти надійності та критеріальної валідності менші від 0,7, на графіку інформаційної функції є кілька екстремумів, то тест вважається складеним неправильно. В цьому випадку необхідно докорінно змінити зміст всього тесту.

Якщо ж на неправильність складання тесту вказує лише одна або дві з перелічених вище характеристик, тест допускає вдосконалення:

- якщо значенні коефіцієнтів надійності або валідності тесту нижче 0,7, кроки вдосконалення тесту визначаються на основі аналізу якості окремих завдань;
- за наявності кількох екстремумів на графіку інформаційної функції вдосконалення здійснюється шляхом додавання завдань, проміжної складності, або поділом тесту на два тести.

На основі аналізу кореляційної матриці завдань, ансамблю характеристичних кривих завдань, інформаційної функції та характеристик кривих підрівнів політомічних завдань, дистракторів

завдань множинного вибору та значень коефіцієнтів валідності завдань виявляють неправильно складені завдання, визначають подальші кроки.

Якщо неправильність складання занадто легкого (занадто складного) завдання, виявленого на етапі аналізу ансамблю характеристичних кривих, слідує із всіх вище перелічених його характеристик, то таке завдання вважають складеним неправильно. У подальшому його вилучають з тесту.

Якщо неправильність складання завдання множинного вибору виявлена на підставі аналізу кореляційної матриці та значень коефіцієнтів валідності завдань, то проводиться детальний аналіз дистракторів цього завдання, на основі якого можна віднайти невдалий дистрактор або помилку в тексті завдання. Якщо ж невдалий дистрактор або помилку не знайдено, завдання вилучається з тесту.

Якщо неправильність складання політомічного завдання виявлена одночасно на основі сумісного аналізу коефіцієнтів кореляції цих завдань з іншими завданнями, аналізу інформаційних функцій та характеристичних кривих підрівнів завдань, значень коефіцієнтів валідності завдань, то таке завдання вилучають з тесту. Якщо ж на неправильність політомічного завдання вказують одна-дві з його характеристик, то завдання потребує доопрацювання, яке виконується на основі аналізу інформаційної функції та характеристичних кривих підрівнів завдання [3].

Таким чином:

- Повноцінний досконалий аналіз якості тесту можливий лише за поєднання взаємодоповнюючих методів КТТ та ІРТ.
- Аналіз якості тесту у цілому неможливий без аналізу якості окремих тестових завдань.
- Сумісний аналіз тесту з використанням методів КТТ та ІРТ дозволяє не лише виявити невдалість тесту, але і вказати на причини такої «невдалості».
- На основі такого аналізу у подальшому можливе здійснення покращення якості як окремих тестових завдань, так і тесту у цілому.
- Розроблена методика дозволяє проводити швидкий аналіз якості завдань та тестів, але не може замінити ретельний аналіз тесту укладачем.
- Накопичений досвід використання створеної методики підтверджує подальшу її перспективність.

Список використаних джерел

1. Про розвиток та досвід експлуатації комплексу дистанційної освіти «Вища математика» / [І.В. Алексеева, В.О. Гайдей, О.О. Диховичний та ін.] // Дидактика математики: проблеми і дослідження : міжнар. зб. наук. робіт. – Вип. 31. – Донецьк: Вид-во ДонНТУ, 2009. – С. 49-56.
2. Чельшкова М. Б. Теория и практика конструирования педагогических тестов / М. Б. Чельшкова. – М.: Логос, 2002. – 431 с.
3. Диховичний О.О. Застосування інформаційної функції для аналізу та підвищення ефективності тестів з вищої математики [Електронний ресурс] / О. О. Диховичний, А. Ф. Дудко // Інформаційні технології і засоби навчання. – Том 41 №3 (2014). – 2014. – С. 55-69 – Режим доступу: http://journal.iitta.gov.ua/index.php/itlt/article/view/1040/790#.U7xiCZR_tbE
4. Крокер Л. Введение в классическую и современную теорию тестов: учебник / Л. Крокер, Дж. Алгина; пер. с англ. Н. Н. Найденовой, В. Н. Симкина, М. Б. Чельшковой; под общ. ред. В. И. Звонникова, М. Б. Чельшковой. – М.: Логос, 2010. – 668 с.
5. Диховичний О. О., Дудко А. Ф. Автоматизована система аналізу результатів комп'ютерного тестування з вищої математики / О. О. Диховичний, А. Ф. Дудко // Наукові праці ДонНТУ. Серія: «Педагогіка, психологія і соціологія». – №2 (14). – Донецьк, 2013. – С. 103-110.

Микитенко П.В.

Національний педагогічний університет імені М.П. Драгоманова

Сучасні комп'ютерно-орієнтовані технології освітніх вимірювань: аналіз, порівняння та вибір

Аналізуючи сучасні пропозиції у сфері розроблення програмного забезпечення та досвід використання комп'ютерно-орієнтованих технологій в вищих навчальних закладах, можна констатувати, що комп'ютерних технологій для підтримки педагогічного контролю та діагностики якості знань студентів існує достатньо, щоб обрати ефективний варіант для реалізації цілей освітніх вимірювань. Однак попри таку їх кількість, існують проблеми з їх адаптацією у вищих навчальних закладах України та високими витратами для отримання повного пакету послуг, оскільки програмне забезпечення з випробувальним періодом (демоверсія), не має тих потрібних функціональних характеристик.

Деякі науковці-педагоги наголошують, що питання інформатизації навчального процесу пов'язане з використанням в ньому комп'ютерних технологій. Комп'ютерні технології зазвичай