УДК 303.2, 37.091.26

АНАЛІЗ ТЕСТІВ РІЗНИХ ФОРМ З МАТЕМАТИЧНОГО АНАЛІЗУ КЛАСИЧНИМИ МЕТОДАМИ ТА ЗАСОБАМИ IRT

Гуртовий Ю.В, Кондрашова К.М.

Створено тестові завдання трьох форм з математичного аналізу. Проаналізовано їх за класичною теорією та засобами IRT у порівняльній формі. Підкреслено необхідність даних методик для контролю якості освіти.
Ключові слова: тестування, математичний аналіз, засоби IRT.

Созданы тестовые задания трех форм по математическому анализу. Они проанализированы по классической теории и средствами IRT, в сравнительной форме. Подчеркнута необходимость данных методик для контроля качества образования.
Ключевые слова: тестирование, математический анализ, средства IRT.

Three forms of test items in mathematical analysis are created. The test items are analyzed according to the classical theory and IRT in the comparative form. The necessity of these methods for quality control of education is underlined.
Key words: testing, mathematical analysis, IRT tools.

Вступ. Актуальною на сьогодні є проблема об'єктивізації та стандартизації контролю в освіті [1, с. 12]. Традиційні форми і методи контролю серйозно критикуються і починає складатися інша система діагностики рівня сформованості знань та вмінь учнів – тестування [2, с. 34].

Науковий рівень тестування (розробка тестів, проведення тестування, обробка і надання результатів в Україні) поки що не відповідають міжнародним стандартам у галузі тестування [4, с. 18]. А це стає справді нагальною потребою, в світлі рішень на загальнодержавному рівні відносно перспектив незалежного тестування в подальшому. Також поширюється розуміння того, що для розробки науково-обґрунтованих педагогічних тестів блискучого знання змісту предмета недостатньо. Необхідна методика, яка включає в себе сукупність методів аналізу змісту і форм завдань, а також науковий підхід для процесу збору і обробки інформації, потрібні методи розрахунку параметрів завдань, методи експертизи якості тестів і тестових завдань. Хоча ці проблеми носять виключно науковий характер, але вони не надумані, а продиктовані самою практикою [5, с. 42].

Відомо, що тільки правильно складений тест дає можливість повністю відповідати сучасним цілям навчання й освіти. Багато викладачів мають певні труднощі в розробці таких тестів та їх правильному методичному застосуванні [2, с. 16]. Таким чином, можна констатувати, що в цьому плані необхідні певні рекомендації. Фундаментальні дослідження тестування, як методу педагогічної діагностики, висвітлені в працях С.Аванесова, В.Беспалька, К.Інгекампа, П.Клайна, А.Майорова, Л.Долінера та ін.

Тому ознайомлення студентів педагогічних вищих навчальних закладів з тестовими технологіями, безперечно, сприяють підвищенню ефективності навчального процесу, формуванню професійної компетентності майбутніх фахівців освітньої галузі.

1.Три системи тестових завдань. Для математико-статистичної обробки емпіричних даних та обробки даних у межах сучасної теорії створення тестів, було проведено дослідження оцінювання знань у студентів першого курсу фізико-математичного факультету на оцінку залишкових знань з курсу математичного аналізу, що вивчаються на першому курсі. Було розроблено тести трьох форм по 6 рівноцінних варіантів кожної форми:

а) тест із вибором однієї правильної відповіді (15 завдань у кожному варіанті);

б) тест на знаходження співвідношення між елементами двох рядів даних (10 завдань у кожному варіанті);

в) тест із множинним вибором (10 завдань у кожному варіанті).

Наведемо по одному завданню із кожного тесту:

А) тестове завдання із вибором однієї правильної відповіді:

Знайти точки розриву функції

$$y=\begin{cases} x, x\le 0 \\ x^2-1, 0<x\le 1 \\ \dfrac{1}{1-x}, x>1 \end{cases}$$ і встановити їх тип:

а) $x=1$ – другого роду;

б) $x=0$ – другого роду, $x=1$ – другого роду;

в) $x=0$ – другого роду, $x=1$ – першого роду;

г) $x=0$ – першого роду;

д) $x=0$ – першого роду, $x=1$ – другого роду.

Б) Тестове завдання на знаходження співвідношення між елементами двох рядів даних:

Знайти і встановити тип точок розриву для кожної функції:

а) $y = ctgx$ 1) $x = 0$ – першого роду;

б) $y = tgx$ 2) $x = -2$ – першого роду усувного типу;

в) $y = \begin{cases} 0, x = 0 \\ \dfrac{\sin x}{x}, x \neq 0 \end{cases}$

3) $x = \pm\pi + 2\pi\kappa, \kappa \in Z$ – другого роду;

г) $y = \begin{cases} \dfrac{|x|}{x}, x \neq 0 \\ 0, x = 0 \end{cases}$

4) $x = 0$ – першого роду усувного типу;

д) $y = \dfrac{x^3 + 8}{x + 2} - 4$

5) $x = \pm\dfrac{\pi}{2} + 2\pi\kappa, \kappa \in Z$ – другого роду.

В) Тестове завдання із множинним вибором:

Знайти точки розриву функції $y = tgx$, і встановити тип точок розриву:

а) $x = \pm 2\pi\kappa, \kappa \in Z$ – другого роду;

б) $x = \dfrac{\pi}{2} + \pi\kappa, \kappa \in Z$ – другого роду;

в) $x = \pm\dfrac{\pi}{2} + 2\pi\kappa, \kappa \in Z$ – першого роду;

г) $x = \pm\dfrac{\pi}{2} + \pi\kappa, \kappa \in Z$ – першого роду;

д) $x = \pm\dfrac{\pi}{2} + 2\pi\kappa, \kappa \in Z$ – другого роду.

У тестуванні брали учать 50 студентів трьох напрямів підготовки “Статистика”, “Інформатика”, “Фізика”. Кожен студент отримав по одному варіанту тестів трьох названих форм та інструкцію до них. Кожна правильна відповідь оцінювалася в один бал, неправильна – нуль балів. Обробка результатів для кожного тесту проводилася окремо.

2. Математико-статистична обробка емпіричних даних.

а) Обробка класичними методами. Підраховано індивідуальні бали досліджуваних та кількість правильних відповідей випробовуваних на кожне завдання тесту. Емпіричні дані представимо графічно за допомогою гістограми частот (рис. 1).

Мода в кожному з трьох випадків відповідно дорівнює: а) 9; б) 5; в) 6. Оскільки значення моди для кожного тесту практично співпали з середнім значенням балів, то можна вважати, що три форми тестів забезпечують нормальний розподіл індивідуальних балів репрезентативної вибірки студентів, а це говорить про те, що нормативно-орієнтовані тести правильно сконструйовані.

Обчислення асиметрії дало змогу визначити ступінь відхилення розподілу спостережуваних частот вибірки від симетричного розподілу: а) $\overline{A}=-0{,}026$; б) $\overline{A}=0{,}097$; в) $\overline{A}=-0{,}106$. Асиметрія розподілу балів від’ємна для а) та в), це говорить про те, що більшість студентів отримали оцінки нижче середнього балу; додатна для б) – більшість студентів отримали оцінки вище середнього балу.

Щоб отримати уявлення про те, чи є полігон частот і гістограма гостро вершинними чи плоскими, було знайдено ексцес: а) $\overline{E}=-0{,}9$; б) $\overline{E}=-0{,}71$; в) $\overline{E}=-0{,}82$. Ексцес від’ємний, тому крива – плоско вершинна.

На наступному кроці було обчислено показник зв’язку між результатами студентів з окремих завдань тесту за допомогою коефіцієнта кореляції. Результати підрахунку звели у відповідну матрицю. Аналіз значень коефіцієнта кореляції дозволяє виділити: а) 4, 6, 11, 15; б) 3, 5, 6, 9; в) 1, 6, 8, 9, 10 завдання, які від’ємно корелюють з іншими завданнями тесту. Таким чином, дані завдання, для підвищення гомогенності змісту, краще змінити або видалити з тесту.

На останньому кроці, за допомогою підрахунку значень коефіцієнта бісеріальної кореляції, оцінено валідність окремих завдань тесту. Аналіз значень коефіцієнта бісеріальної кореляції вказує на невдалі тестові завдання. Завдання можна вважати валідним, якщо значення $(r_{bis})_j \approx 0{,}5$, але оскільки вибірка у нас невелика, то будемо вважати завдання валідним, якщо значення $(r_{bis})_j$ перевищує 0,3. Завдання: а) 6, 11, 15; б) 5, 9; в) 6, 9, 10 не задовольняють дану вимогу. Отриманий висновок говорить про низьку валідність цих завдань. Ці завдання потрібно визнати невдалими і для вдосконалення тесту їх необхідно видалити.

Потрібно виключити завдання з низькою валідністю і ті, які негативно корелюють з іншими завданнями. Отже для кожної з форм тесту ми вилучаємо лише а) 6, 11, 15; б) 5, 9; в) 6, 9, 10.

Підсумувавши отримані результати за класичною теорією, ми дійшли висновку, що розроблені тести правильно сконструйовані, але потребують удосконалення кількох завдань.

б) Обробка засобами IRT. Аналіз результатів складається з кількох етапів. На першому етапі ми за однопараметричною моделлю Раша, яка представлена формулами:

$$P_j(Q)=\frac{e^{1{,}7(Q-\beta_j)}}{1+e^{1{,}7(Q-\beta_j)}}, \quad P_i(Q)=\frac{e^{1{,}7(Q_i-\beta)}}{1+e^{1{,}7(Q_i-\beta)}}$$

обчислили та проілюстрували інваріантність оцінок рівня підготовки випробовуваних від складності завдань тесту. Проаналізували взаємне розміщення побудованих характеристичних кривих завдань тесту. Для побудови графіка функції $P_j(Q)$ необхідно вибрати декілька значень незалежної змінної Q, а потім обчислити значення функції $P_i(Q)$. Після нанесення значень функції на координатній площині і з’єднавши отримані точки, графік функції $P_j(Q)$ має вигляд (рис. 2).

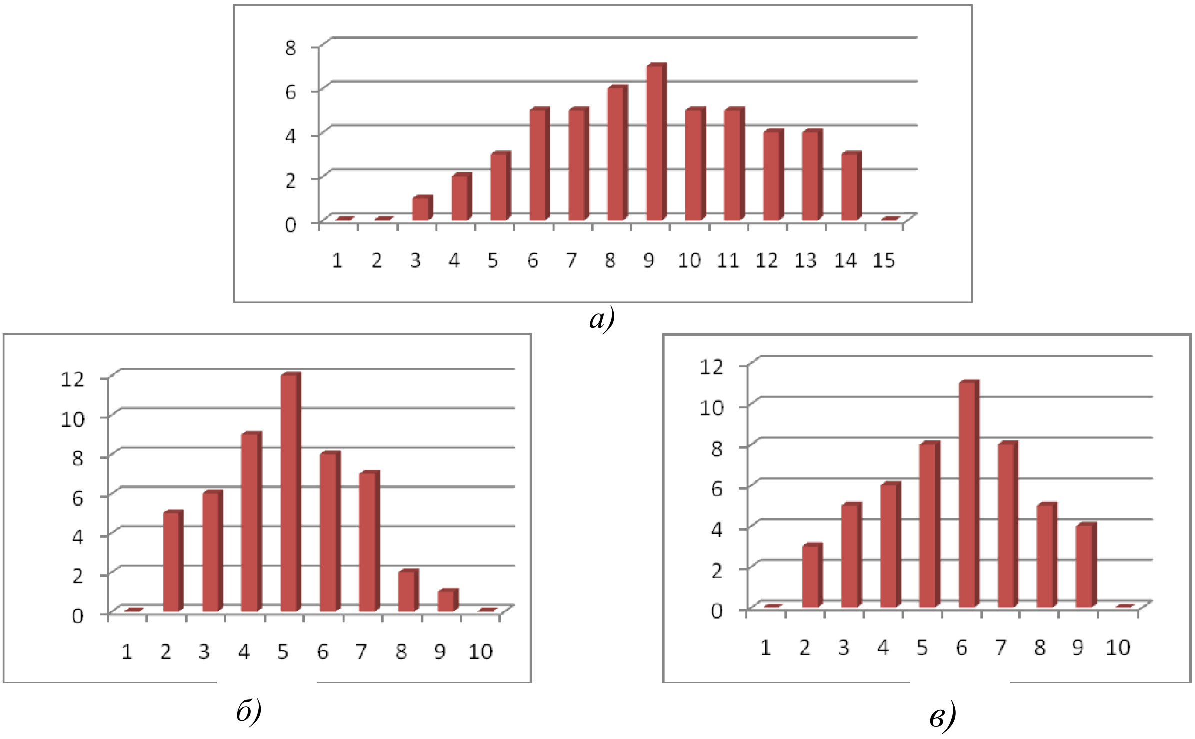


Рис. 1. ***Частота прояву кожного балу: а) тест із вибором однієї правильної відповіді; б) тест на знаходження співвідношення між елементами двох рядів даних; в) тест із множинним вибором***

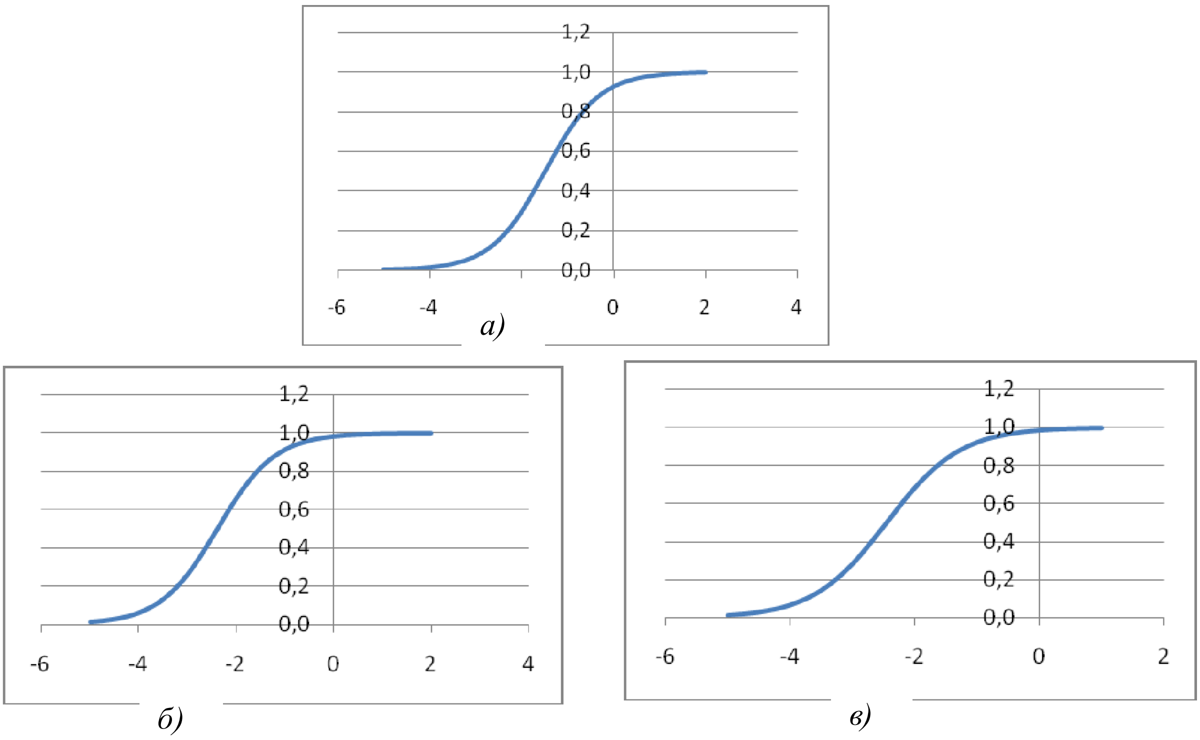
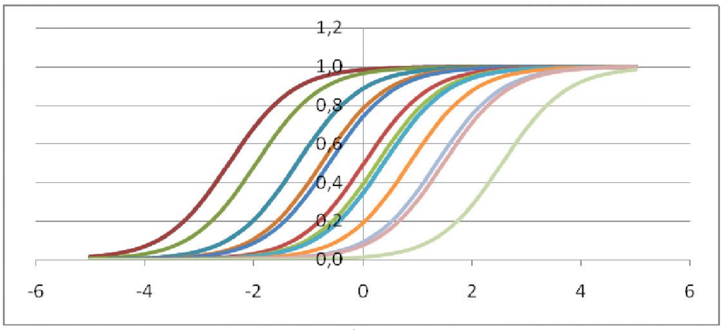
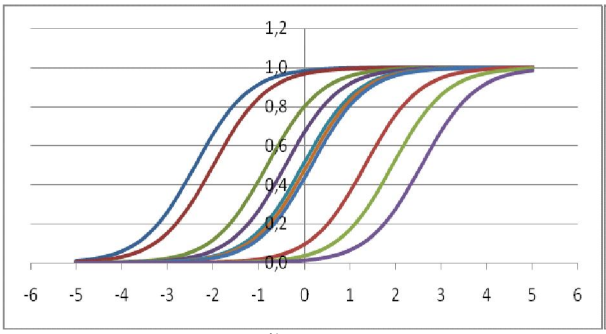


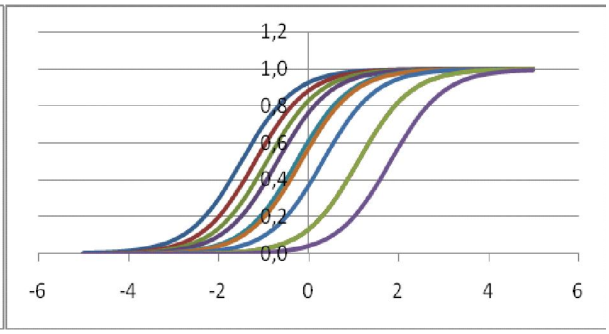
Рис. 2. ***Характеристичні криві 1-го завдання тестів: а) тест із вибором однієї правильної відповіді; б) тест на знаходження співвідношення між елементами двох рядів даних; в) тест із множинним вибором***



а)

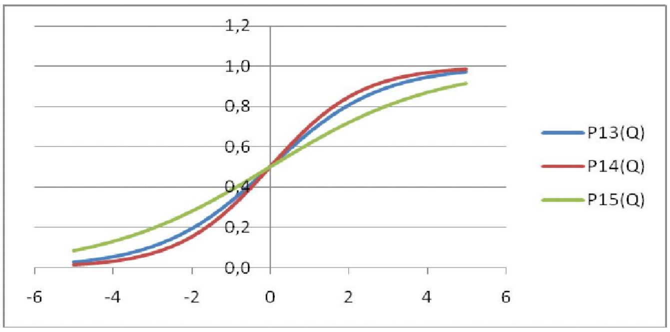


б)

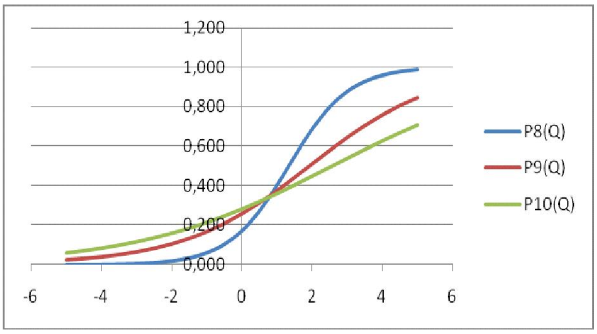


в)

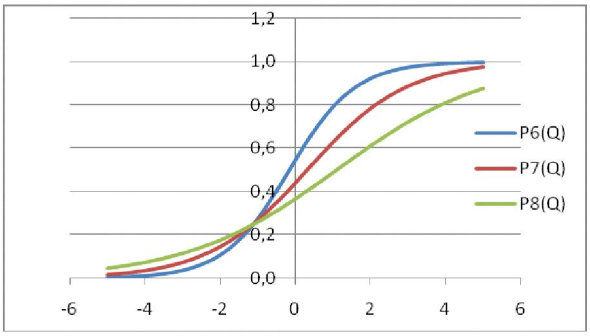
Рис. 3. ***Характеристичні криві 15(10) завдань: а) тесту із вибором однієї правильної відповіді; б) тесту на знаходження співвідношення між елементами двох рядів даних; в) тесту із множинним вибором***



а)

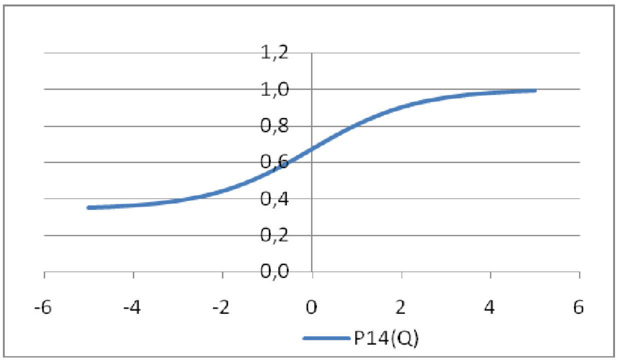


б)

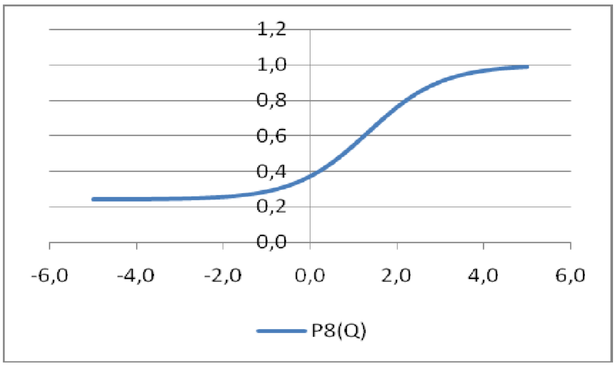


в)

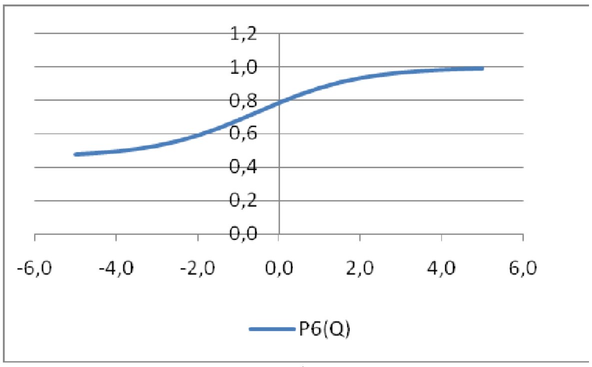
Рис. 4. ***Характеристичні криві трьох завдань: а) тесту із вибором однієї правильної відповіді; б) тесту на знаходження співвідношення між елементами двох рядів даних; в) тесту із множинним вибором***



а)



б)



в)

Рис. 5. ***Характеристичні криві завдань: а) тесту із вибором однієї правильної відповіді; б) тесту на знаходження співвідношення між елементами двох рядів даних; в) тесту із множинним вибором***

Отже, випробовуваний з рівнем підготовки, який дорівнює складності 1-го завдання тесту, відповість на нього правильно з імовірністю 0,5. Для випробовуваних з рівнем знань набагато більшими, ймовірність правильної відповіді прямує до одиниці. Якщо рівень підготовки набагато менший – імовірність правильної відповіді прямує до нуля. Характеристичні криві: а) 15; б) 10; в) 10 завдань наведені на рис. 3.

Зміщення кривої вправо говорить про збільшення складності даного завдання, вліво – зменшення складності завдання.

На наступному кроці побудовано двохпараметричну модель А.Бірнбаума, щоб з'ясувати умовну ймовірність правильного виконання *j*-го завдання випробовуваними з різними значеннями Q за формулою $P_j(Q)=\frac{e^{1,7a_j(Q-\beta_j)}}{1+e^{1,7a_j(Q-\beta_j)}}$, $a_j=\frac{(r_{bis})j}{\sqrt{1-[(r_{bis})_j]^2}}$.

На рис. 4 наведені характеристичні криві трьох завдань різної крутизни.

Для порівняльної характеристики якості завдань під час диференціації знань випробовуваних розглянемо помітно різні по крутизні криві: а) 14-го та 15-го завдань тесту. Крива P14(Q) дуже крута, що відповідає великому значенню a_{14}, а крива P15(Q) дуже полога: a_{15} 0. Для випробовуваних з різним рівнем підготовки ймовірність правильного виконання 14-го завдання суттєво відрізняється. Для 15-го завдання ймовірність досить однакові; б) 8-го та 10-го завдань тесту. Крива P8(Q) дуже крута, що відповідає великому значенню a_8, а крива P10(Q) дуже полога: $a_{10} \longrightarrow 0$. Для випробовуваних з різним рівнем підготовки ймовірність правильного виконання 8-го завдання суттєво відрізняється. Для 10-го завдання ймовірність досить однакові; в) 6-го та 8-го завдань тесту. Крива P6(Q) дуже крута, що відповідає великому значенню a_6, а крива P8(Q) дуже полога: $a_8 \longrightarrow 0$. Для випробовуваних з різним рівнем підготовки ймовірність правильного виконання 6-го завдання суттєво відрізняється. Для 8-го завдання ймовірність досить однакові.

Можемо зробити висновок, що завдання, де $a \longrightarrow 0$ є безкорисними під час диференціації випробовуваних груп, оскільки вони не несуть інформації про індивідуальні різності студентів.

Трьохпараметричну модель А.Бірнбаума характеризує вплив на достовірність попередніх результатів аналізу ефекту відгадування правильної відповіді. Було використано формулу:

$$P_j\{x_{ij}=1/\beta_j\}=c_j+(1-c_j)\frac{e^{1,7a_j(Q-\beta_j)}}{1+e^{1,7a_j(Q-\beta_j)}} ,$$

де параметр c_j характеризує ймовірність правильної відповіді випробовуваних на *j*-те завдання тесту. Для порівняльної характеристики, побудували характеристичну криву для а) 14-го; б) 8-го; в) 6-го завдань (рис. 5).

Порівнюючи з кривою двохпараметричної моделі для а) 14-го; б) 8-го помітним є те, що дані криві є більш пологими. Таким чином, ефект угадування понижує диференціальну можливість завдань тестів. Отже, введення третього параметра не лише понижує точність оцінок параметрів Q і β, але й погіршує сходження ітераційних методів, використаних для підвищення точності оцінок латентних параметрів Q і β.

Висновки. Інтерпретація результатів за класичною та сучасною теорією тестів підкреслює необхідні корективи завдань кожної з форм тестів для подальшого використання їх під час проведення підсумкового контролю знань. Отже, експериментальна перевірка ефективності розробленої методики тестування дає підстави наголосити на важливості математико-статистичної обробки емпіричних даних перед запровадженням у навчальний процес.

Література

1. Фіцула М. М. Педагогіка : навч. посіб. / М. М. Фіцула. – 2-ге вид., випр., доп. – К. : Академвидав, 2007. – 560 с.

2. Малихін А. Тести у навчальному процесі сучасної школи / А. Малихін // Рідна школа. – 2001. – № 8.

3. Фігурська Л. В. До проблеми забезпечення ефективності тестових технологій як засобу контролю професійної підготовки майбутніх спеціалістів / Л. В. Фігурська // Професійні компетенції та компетентності вчителя : матеріали регіонального науково-практичного семінару. – Тернопіль : Вид-во ТНПУ ім. В. Гнатюка, 2006. – С. 40–44.

4. Майоров А. Н. Теория и практика создания тестов для системы образования / А. Н. Майоров. – М. : Народное образование, 2000 – 352 с.

5. Практикум педагогічної майстерності : навч. посіб. / кол. автор. : Л. М. Сергеєва, А. О. Молчанова, О. В. Пащенко та ін. ; за ред. В. В. Олійника. – К. : ТОВ “Етіс Плюс”, 2008. – 184 с.