Міністерство освіти і науки України
Національна академія аграрних наук України
ННВК «Всеукраїнський науково - навчальний консорціум»
Вінницький національний аграрний університет
Верхівський сільськогосподарський коледж
Ладижинський коледж
Могилів-Подільський технолого-економічний коледж
Немирівський коледж будівництва та архітектури
Технологічно-промисловий коледж
Чернятинський коледж

ЗБІРНИК ТЕЗ
за матеріалами
ІІІ Всеукраїнської науково-практичної
інтернет-конференції

«ПРОФЕСІЙНА ПІДГОТОВКА ФАХІВЦЯ В КОНТЕКСТІ ПОТРЕБ СУЧАСНОГО РИНКУ ПРАЦІ»

27 лютого 2018 року

м. Вінниця

УДК 378.147

ОЦІНКА ЯКОСТІ ДИСТРАКТОРІВ У МАТЕМАТИЧНИХ ТЕСТАХ

Найко Д. А., доцент,
канд. фіз. -мат. наук, доцент ВНАУ

Розглядаються питання рівномірності розподілу якості дистракторів та розподілу якості тестових завдань у закритих математичних тестах.

У контексті ефективності закритих математичних тестів важливим моментом є якісний добір дистракторів. Числова оцінка правдоподібності дистракторів спирається на знаходження частки тих учасників тестування, які вибрали неправильні відповіді при тестуванні.

Припустимо, що 200 осіб отримали тестове завдання, з чотирма варіантами відповіді, з яких один – правильний, а решта три – дистрактори. Якщо це завдання правильно виконали 70% усіх учасників тестування, то його складність дорівнює $p = 0,7$. Решта 60 учасників вибрали за правильну відповідь один із дистракторів. Якщо дистрактори однаково привабливі, то кожен із дистракторів вибрали приблизно по 20 учасників. Іншими словами, вибір кожного із дистракторів підпорядкований майже рівномірному закону їхнього розподілу.

На практиці прийнято вважати, що дистрактор, який вибрали менше за 5% учасників від числа тих, які відповіли неправильно, треба вилучити. Це свідчить про дуже нерівномірний розподіл дистракторів за якістю, і отже залучення такого дистрактора до тесту є невдалим.

З іншого боку, рівномірність розподілу дистракторів лише допомагає контролювати їхню правдоподібність і не означає нічого більшого за це.

У наших міркуваннях ми не закладаємо ефекту вгадування. Адже, коли група осіб, у якій проводиться тестування, схильна до вгадування, то на практиці це спонукає до переоцінки складності тестових завдань. У разі, коли істинна складність тестового завдання дорівнює 0,5, а 50% тих учнів, які не знають правильної відповіді, вгадують її з імовірністю 0,5, то дослідна складність завдання дорівнює 0,75. Крім того, окремі вгадування не є «сліпими», бо правильну відповідь іноді можна вгадувати методом раціонального відбору за певними ознаками. Проте інколи це теж можна вважати проявом окремих знань. У цьому контексті виникає цікаве запитання: як за отриманою дослідною складністю тестового завдання знайти істинну складність завдання?

Ці міркування приводять до необхідності побудови якісних дистракторів. Неякісні дистрактори є однією з причин неправильної оцінки істинної складності тестового завдання.

Одним із методів аналізу якості дистракторів у тестових завданнях з однією правильною відповіддю є метод порогових груп [2]. Найважливіша тенденція цього аналізу проявляється у тому, що найсильніші групи екзаменованих осіб найчастіше вибирають правильну відповідь, а всі дистрактори, якщо вони якісні, мають приблизно однакову кількість «голосів».

Цілком валідні завдання тесту повинні мати хорошу роздільну здатність (дискримінативність). У практичній тестології можуть складатися різні, навіть парадоксальні ситуації. У групі з високим рівнем підготовленості може спостерігатися тенденція до неправильної відповіді на тестове завдання. У такому випадку ми маємо справу з негативною дискримінативністю. Для завдань однакової складності також можна спостерігати різну здатність диференціювати учасників тестування на слабкіших та сильніших.

Для вимірювання роздільної здатності тестових завдань використовуються різні показники. Найуживаніші з них базуються на понятті кореляції або на методі порогових груп.

У випадку використання дихотомічної шкали, найпростішим показником роздільної здатності є *індекс дискримінативності*, який визначається таким чином. За результатами загального тестування всіх його учасників поділяють на 3−4 рівні за їх чисельністю групи. Тоді індексом дискримінативності певного тестового завдання називається число

$$d = p_{слн} - p_{слб},$$

де $p_{слн}$ − частка учасників тестування найсильнішої групи, які правильно відповіли на дане завдання тесту; $p_{слб}$ − частка учасників тестування найслабкішої групи, які правильно відповіли на дане завдання тесту. Оскільки найбільше значення кожної з цих часток може дорівнювати 1 (коли всі представники групи відповіли правильно), а найменше може дорівнювати 0 (коли всі представники групи відповіли неправильно), то величина $d = p_{слн} - p_{слб}$ може набувати значень від −1 до 1.

Зважаючи на простоту цього показника, він є дуже прийнятним для використання на рівні студентської групи чи лекційного потоку кількох груп. Вважається практично задовільним, якщо $d \geq 0,4$.

Визначення роздільної здатності завдань тесту з використанням поняття коефіцієнта кореляції може здійснюватися через визначення: φ -коефіцієнта кореляції; точково-бісеріального коефіцієнта кореляції; бісеріального коефіцієнта кореляції; тетрахоричного коефіцієнта кореляції.

φ -коефіцієнта кореляції є окремою формою коефіцієнта кореляції Пірсона для випадку дихотомічної шкали вимірювання. Він відображає зв'язок складностей завдань тесту.

Коефіцієнтом кореляції випадкових величин X та Y називається число

$$\rho(X,Y)=\frac{M(X\cdot Y)-M(X)\cdot M(Y)}{\sqrt{D(X)\cdot D(Y)}},$$

де M − математичне сподівання, D − дисперсія.

Якщо результати тестування подаються у дихотомічній шкалі, то коефіцієнт кореляції між i -м та j -м завданнями тесту називають φ - *коефіцієнтом кореляції* і обчислюють за формулою

$$\varphi_{ij}=\frac{p_{ij}-p_i\cdot p_j}{\sqrt{p_i q_i\cdot p_j q_j}},$$

де p_{ij} − частка учасників, які впоралися з i -м та j -м завданнями тесту; p_i − частка учасників, які впоралися з i -м завданням ($q_i=1-p_i$); p_j − частка учасників, які впоралися з j -м завданням ($q_j=1-p_j$).

Точково-бісеріальний коефіцієнт кореляції також використовують у випадку дихотомічної шкали. Він відображає зв'язок між оцінкою за виконання певного завдання та критеріальною оцінкою і також є окремою формою коефіцієнта кореляції Пірсона.

В основу використання бісеріального коефіцієнта кореляції взято припущення про те, що якість завдань розподілена за нормальним законом.

Якщо потрібно вдатися до дихотомізації нормального закону розподілу випадкової величини, то використовується так званий тетрахоричний коефіцієнт кореляції.

Ми не обговорюємо питання про те, в якій ситуації найдоцільнішим є використання того чи іншого показника роздільної здатності. Зауважимо лише, що у випадку, коли складність завдань близька до середньої, то вибирати можна будь-який з цих показників. Очевидно, що найпростішим є використання індекса дискримінативності d . Проте у випадку, коли потрібно перевірити на значущість отриману числову оцінку, треба вибирати один із коефіцієнтів кореляції.

Література

1. Найко Д.А. Особливості побудови тестових завдань з математичних дисциплін / Д.А. Найко // Сучасні інформаційні технології та інноваційні методи навчання у підготовці фахівців: методологія, теорія, досвід, проблеми. //**Зб. наук. пр. – Випуск 48 / редкол. – Київ-Вінниця**: ФОП Тарнашинський О. В., 2017. – С. 154 – 160.

2. Вимірювання в освіті: Підручник / За редакцією О. В. Авраменко. − Кіровоград: «КОД», 2011. – 360 с.