УДК 519.2:378.14

КЛАСИФІКАЦІЯ ШКАЛ ВИМІРЮВАННЯ ОЦІНКИ ЗНАНЬ

Матвіїшина Н.В., Бровченко Є.Е.

Вимір і оцінка знань традиційно залишаються важливими в процесі навчальної діяльності.

Оцінка – більш складне поняття, ніж вимір. Вона включає як кількісний і якісний опис, так і інтерпретацію оцінюваного результату. Вимір обмежується кількісним описом і не включає якісних описів або суджень про цінність знань, що вимірюються.

Предметом виміру в навчальному процесі є такі властивості тих, кого навчають, як засвоєння знань, розуміння навчального матеріалу і т.д.

Основною проблемою вимірів у загальному випадку є пошук числової (символьної) системи, що відповідає емпіричній, в якій відношення між числами зберігає відношення між самими об'єктами вимірів. У цьому зв'язку ставиться питання про застосування відповідної шкали.

Під шкалою розуміється упорядкована трійка ***<A,R,f>***, де

A – емпірична система;

R – числова система;

f – гомоморфне відображення ***A*** в ***R*** (в ідеальному випадку – ізоморфне відображення, що відповідає єдиній числовій системі, якщо вона адекватна емпіричній).

Роботи в області класифікації основних шкал вимірювання проводилися С.Стівенсом [3], В.П.Беспалько [1], Н.М.Розенбергом [4] і іншими.

За Стівенсом, найбільш уживаними є чотири типи вимірювальних шкал: номінальні (шкали найменувань), шкали порядку, шкали інтервалів і шкали відношень.

У цій статті пропонується класифікація перерахованих шкал у залежності від характеристик об'єктів шкалування, типів вимірюваних даних, припустимих статистичних операцій над даними.

Номінальна шкала (шкала найменувань). У шкалах найменувань числа використовуються просто як імена, символічні позначення об'єктів; самі числа можуть бути зовсім довільними. Таким чином, об'єкти шкалування класифіковано, класи позначені неспівпадаючими номерами.

Шкалу найменувань можна, наприклад, одержати, класифікуючи тих, кого навчають, згідно зі статтю, національністю, мовою навчання, а також класифікуючи експериментальний та контрольний методи викладання навчального предмета (позначивши їх літерами А і В).

Арифметичні процедури в межах номінальної шкали не мають змісту. Усі види числової обробки стосуються упорядкування об'єктів у кожному класі. Тому можна одержати, наприклад:

- частоту (числа спостережень),
- моду (значення, що має найбільшу частоту).

Методи, застосовувані в процесі цієї обробки, називаються статистикою якісних ознак.

Шкала порядку. Об'єкти шкалування розбиті на класи й упорядковані за ступенями зміни ознаки. Дані являють собою послідовні незв'язані ранги 1, 2, …, *n*. Ці ранги можна привласнити вимірам іншої шкали (наприклад, вихідні дані 157, 100, 48 ранжируються номерами 1, 2, 3). Інтервали між числами, що відносяться до конкретних об'єктів, не рівні, але їхній порядок можна планувати.

Шкалу порядку можна одержати, ранжуючи тих, кого навчають, згідно зі знаннями, вміннями, здібностями і т.д.

З даними порядкової шкали можливі такі операції:

- визначення медіани,
- визначення процентилей (Р-й процентиль являє собою точку, нижче якої лежить Р відсотків оцінок),

- визначення коефіцієнтів рангової кореляції:

- коефіцієнт кореляції Спірмена [2]

$$r_s = 1 - \frac{6\sum_{i=1}^{n}(X_i - Y_i)}{n(n^2-1)},$$

де $X_i - Y_i$ - різниця між рангом за X i–го об'єкта і його ж рангом за Y;

n – кількість об'єктів.

коефіцієнт кореляції Кендала [2]

$$\tau = \frac{P-Q}{n(n-1)/2},$$

де P – число «збігів» (порядок за X і Y однаковий);

Q – число «інверсій» (порядок за X і Y розрізнений).

Основне обмеження при операціях з порядковими шкалами – неприпустимість арифметичних дій з ранговими оцінками (обчислення суми, середнього арифметичного, середнього квадратичного відхилення), тому що результати цих операцій можуть нічого не говорити про кількість аналізованої властивості, якою володіють об'єкти. Наприклад, різниця між «рангами успішності» студентів A і B дорівнює трьом; різниця між рангами C і D дорівнює одиниці. Але нема рації в тому, що різниця в успішності між A і B у три рази вище, ніж між C і D. Результати арифметичних дій тут не можна інтепретувати так, щоб вони говорили що-небудь про кількість властивості, якою фактично володіють об'єкти.

Шкала інтервалів. Володіючи усіма властивостями порядкової шкали, шкала інтервалів дозволяє визначити розмір інтервалу (дистанцію) між двома точками шкали.

Використовуючи интервальну шкалу, варто не тільки точно визначити відношення еквівалентності, як у номінальній шкалі, відношення порядку, як у порядковій шкалі, але і визначити відношення між будь-якими двома інтервалами. Одиниця виміру і нульової точки вибираються довільно, і відношення між двома будь-якими інтервалами на ній не залежить від положення нуля шкали і розміру одиниці виміру. Рівняння шкали інтервалів відповідає лінійній функції $y=ax+b$. Константа b вказує на умовність нульової точки шкали.

Розміри інтервалів допускають виконання всіх арифметичних операцій:

- визначення середнього арифметичного шкали:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i ,$$

де $x_1,...,x_n$ - значення деякої характеристики;

n – кількість об'єктів, що володіють даною характеристикою.

- визначення дисперсії:

$$D = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

- визначення середнього квадратичного відхилення

$$\sigma = \sqrt{D}\,.$$

Шкала відношень. На додаток до всіх характеристик інтервальної шкали, шкала відношень володіє й абсолютною нульовою точкою.

Шкали відношень відповідають перетворенню $y=ax$. Формування таких шкал можливо, якщо знайдені способи зрівнювання відношень розмірів, що відповідають точкам на шкалі.

Арифметичні дії у випадку шкали відношень застосовні не тільки до інтервалів між розмірами, але і до самих цих розмірів. Прикладами перемінних, що задовольняють вимогам шкали відношень, є час вирішення завдання, кількість інформації (за Шенноном).

Характерні риси шкал подані в таблиці 1.

При оцінці знань тих, кого навчають, традиційно використовується чотирибальна шкала. Цю шкалу можна назвати змістовною шкалою порядку, якщо співвіднести її з визначеним рівнем засвоєння.

Розглянемо рівні засвоєння, запропоновані В.П.Беспалько [1].

І рівень: розпізнавання, розрізнення, класифікація об'єктів.

ІІ рівень: репродуктивний, тобто рівень засвоєння, коли той, кого навчають, здатний відтворювати, обговорювати, аналізувати, застосовувати інформацію для вирішення типових завдань.

ІІІ рівень: продуктивний, тобто відтворення і застосування засвоєної інформації для вирішення виникаючих (нетипових) завдань.

IV рівень: застосування засвоєної інформації для вирішення практичних завдань, що знаходяться за межами того класу явищ, на якому йшло формування знань.

Таблиця 1

Вид шкали	Характеристики об'єктів шкалування	Тип вимірюваних даних	Статистичні операції над даними	Приклади
Номінальна	Класифіковані, класи позначені неспівпадаючими номерами	Дискретні; немає масштабу і нульової точки	Припустимі: визначення частоти, моди; неприпустимі: обчислення медіани, середнього, арифметичні операції	Класифікація за статтю, національністю, мовою навчання; номер у списку
Порядкова	Розбиті на класи, упорядковані за ступенем зміни ознаки	Дискретні; немає масштабу і нульової точки	Припустимі: визначення частот, процентилей, коефіцієнтів рангової кореляції; неприпустимі: середнє арифметичне, середнє квадратичне відхилення	Ранжування тих, кого навчають за знаннями, вміннями, здібностями, бал за *к*-бальною шкалою
Інтервальна	Вимірюється інтервал між об'єктами щодо деякої властивості	Дискретні і безперервні; нульова точка довільна і не вказує на відсутність властивості	Неприпустимі: визначення середнього геометричного, коефіцієнта варіації; всі інші - припустимі	Календарний час, швидкість
Відношень	Об'єктам привласнюються числа, що мають властивості об'єктів інтервальної шкали	Дискретні і безперервні; нульова точка свідчить про відсутність оцінюваної властивості	Припустимі практично всі статистичні операції	Час вирішення завдання, кількість інформації

Наведемо приклад використання порядкової шкали при аналізі результатів навчальної діяльності.

Є такі вихідні дані: екзаменаційні оцінки студентів (20 чоловік) економічного факультету за дисципліною «Інформатика та комп'ютерна техніка» і оцінки, отримані цими студентами за комплексну контрольну роботу (ККР). Визначимо: результат виміру X – екзаменаційні оцінки; результат виміру Y – оцінка за ККР. Дані були проранжировані: 1-й ранг привласнюється найкращій відповіді, 2-й ранг – відповіді, близькій до кращої, …, 20-й ранг – найгіршій відповіді. Результати наведені в таблиці 2.

Зробивши необхідні обчислення, і підставивши відповідні значення у формулу коефіцієнта кореляції Спірмена, одержимо, що

$$r_s = 0.88$$

Значення r_s свідчить про досить сильний прямий зв'язок між отриманими оцінками на іспиті та оцінками за ККР.

Викладені вище поняття про способи виміру і можливі вимірювальні шкали складають основу для об'єктивного контролю знань тих, яких навчають.

Таблиця 2

Студент №	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
X	1	4	17	5	6	2	7	16	3	8	9	15	14	19	10	11	20	12	13	18
Y	3	6	14	7	4	2	5	18	1	8	16	17	9	19	13	11	20	10	12	15

ЛІТЕРАТУРА

1. Беспалько В.П. Основы теории педагогических систем. – Воронеж, 1977.
2. Гласс Дж., Стэнли Дж. Статистические методы в педагогике и психологии. – М.: Прогресс, 1976.
3. Стивенс С. Математика, измерение и психофизика //Экспериментальная психология. – М., 1960.
4. Розенберг Н.М. Проблемы измерений в дидактике. – К.: Вища школа, 1979.