

СТАНДАРТИЗАЦИЯ С ИСХОДНЫМИ ДАННЫМИ ПРЕДСТАВЛЕННЫМИ ОТНОШЕНИЕМ ЛИНЕЙНОГО ПОРЯДКА

Гомонов Д.А.,

научный руководитель канд. техн. наук Даничев А.А.

Сибирский федеральный университет

Для обработки первичных данных полученных в шкалах равных отношений разработаны и широко применяются методы стандартизации. В данной работе рассматривается подход к стандартизации в случае строгого ранжирования.

Существуют различные способы получения информации от респондента, эта информация может быть как наименее информативная (дихотомическая шкала), так и очень детальной Метод Анализа Иерархий (Саати, шкала равных отношений). Для психологических тестов наиболее важно получение «быстрой» информации, возможность не задумываться над ответом. Поэтому наиболее распространены дихотомические шкалы (да/нет, нравится/не нравится) и бальные оценки: отлично (3), хорошо (2), плохо (1), очень плохо (0).

При получении экспертных данных в порядковых шкалах объекты либо непосредственно ранжируются экспертом, либо производятся парные сравнения объектов друг с другом. Такая форма суждений наиболее естественна для человека, однако требует дополнительных разъяснений при проведении теста.

Шкалы измерений принято классифицировать по типам измеряемых данных, которые определяют допустимые для данной шкалы математические преобразования, а также типы отношений, отображаемых соответствующей шкалой.

Типы шкал:

- 1) Номинальные
- 2) Порядковые
- 3) Интервальные
- 4) Равных отношений

Для равных отношений отработаны механизмы первичной обработки данных: удаление выбросов, проверка однородности выборки, определение вида распределения и т.д. Результат тестирования испытуемого x_i , вычисленный с помощью диагностической модели $y_i = u(x_i)$, обычно называют первичной тестовой оценкой или, часто, «сырым» баллом. Для лучшего понимания этого результата в ряду других результатов производится его дальнейшее искусственное преобразование, основанное на анализе эмпирического распределения тестовых оценок в репрезентативной выборке испытуемых. Процедура такого преобразования носит название стандартизации. Известно три основных вида стандартизации первичных тестовых оценок: 1) приведение к нормальному виду; 2) приведение к стандартной форме; 3) квантильная стандартизация.

Для порядковых шкал способы обработки и анализа данных разрознены и известны лишь узким специалистам. Поэтому психологи могут столкнуться с проблемой при обработке результатов тестирования.

В данной работе рассматривается применение порядковых шкал в методике косвенного измерения системы самооценок (КИСС). Данная методика демонстрирует, естественность получения информации (расставить изображения в порядке предпочтения), а так же высокую эффективность и глубину диагностики самооценки по

сравнению с другими методами. Однако с математической обработкой результатов возникают серьёзные проблемы (так в этой методике вместо меры близости на порядковых шкалах используется коэффициент ранговой корреляции между парами ранжирований).

Стандартизации в случае строгого ранжирования:

Автор предлагает следующий подход к стандартизации в случае строгого ранжирования:

- 1) использование меры близости на отношениях порядка;
- 2) изучение распределения первичных тестовых оценок;
- 3) приведение, распределение тестовых оценок к нормальному виду [2].

Представления отношений

Информация об отношениях может быть представлена *матрицами отношений*. Строки и столбцы матрицы $\|t_{ij}\|$ отношения P соответствуют элементам множества A . Для представления отношений линейного порядка используется матрица $\|t_{ij}\|$ с элементами

$$t_{ij} = \begin{cases} 1, & \text{если } (a_i, a_j) \in P, (a_j, a_i) \notin P \\ & \text{(эксперт предпочитает объект } a_i \text{ объекту } a_j), \\ 0, & \text{если } (a_i, a_j) \in P, (a_j, a_i) \in P \\ & \text{(эксперт считает объекты } a_i \text{ и } a_j \text{ равноценными),} \\ -1, & \text{если } (a_i, a_j) \notin P, (a_j, a_i) \in P \\ & \text{(эксперт предпочитает объект } a_j \text{ объекту } a_i), \end{cases} \quad (1)$$

Меры близости на отношениях

Меры близости позволяют определить, насколько близки или далеки точки зрения экспертов и являются одним из основных инструментов, используемых при анализе и обработке экспертной информации [4].

$$d(P_1, P_2) = \sum_{i < j} |t_{ij}^1 - t_{ij}^2| \quad (2)$$

Формула (2) определяет единственную меру близости на ранжированиях [1].

Часто, для вычисления расстояний между ранжированиями, используют другие метрики, что не всегда корректно и не отображает специфику порядковых шкал.

Изучение распределения первичных тестовых оценок

Если ответы строго ранжированы то произвольный ответ соответствует строгому упорядочиванию или перестановке. Для того, чтобы знать закон распределения расстояний для базового ранжирования необходимо знать количество ранжирований с заданным расстоянием. В таблице 1 представлен пример для трёх объектов. Всего возможно шесть строгих ранжирований. Максимально возможное расстояние между ранжированиями (a,b,c) и (c,b,a) равно шести.

Таблица 1. Количество перестановок для трёх объектов

| Расстояние | 0 | 2 | 4 | 6 |
|------------|---|-----|-----|---|
| Ранг 1 | a | b a | c b | c |
| Ранг 2 | b | a c | a c | b |
| Ранг 3 | c | c b | b a | a |

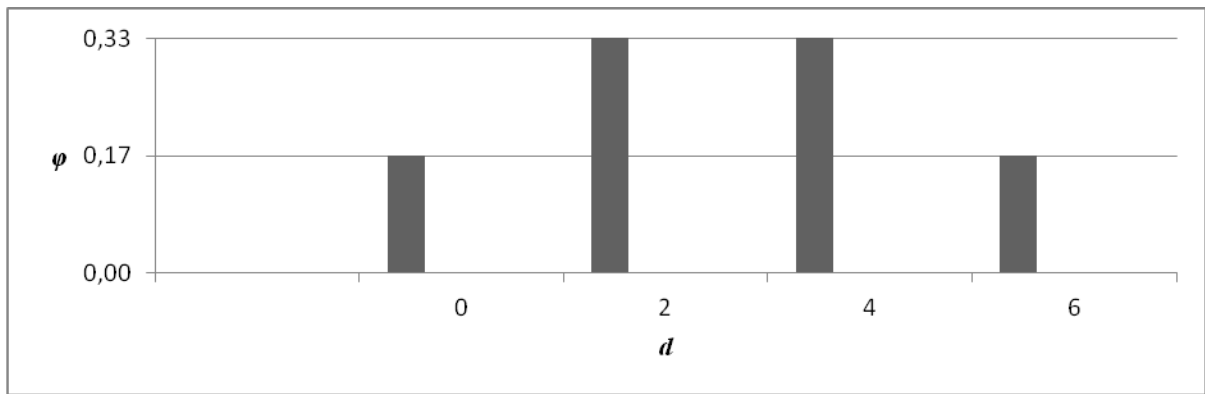


Рисунок 1. Плотность распределение возможных расстояний до случайного ранжирования

Расстояние между строгими ранжированиями соответствует числу инверсий в соответствующих перестановках. В свою очередь количество инверсий в перестановках соответствует числу Mahonian [3]:

$$M(0,0) = 1,$$

$$M(0,c) = 0 \quad \text{for } c \neq 0$$

$$M(r,c) = \sum_{k=0}^r M(r-1, c-k) \quad \text{for } r > 0$$

(3)

Вычисляется $M(r,c)$ по рекуррентной формуле.

Первые числа Mahonian, начиная с $M(0,0)$, представлены в Таблице 2.

Максимально возможное число перестановок равно $N=m(m-1)$, где m – количество объектов. Таким образом вероятность $P(d_i=d)$ того что расстояние d_i до случайного ранжирования равно d определяется как $M(m,d)/N$.

Таблица 2. Первые числа Mahonian

| Количество объектов r | Возможное количество инверсий c | | | | | | | | | | | | | | | |
|-----------------------|---------------------------------|---|----|----|----|----|----|-----|-----|----|----|----|----|----|----|----|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| 1 | 1 | | | | | | | | | | | | | | | |
| 2 | 1 | 1 | | | | | | | | | | | | | | |
| 3 | 1 | 2 | 2 | 1 | | | | | | | | | | | | |
| 4 | 1 | 3 | 5 | 6 | 5 | 3 | 1 | | | | | | | | | |
| 5 | 1 | 4 | 9 | 15 | 20 | 22 | 20 | 15 | 9 | 4 | 1 | | | | | |
| 6 | 1 | 5 | 14 | 29 | 49 | 71 | 90 | 101 | 101 | 90 | 71 | 49 | 29 | 14 | 5 | 1 |

Рассмотренный подход позволил изменить методику подсчёта ключей КИСС и повысить точность и эффективность его работы. Что было апробировано на учащихся 7-9 классов.

Список литературы

1. Литвак, Б. Г. Экспертная информация: методы получения и анализа. М.: Радио и связь, 1982. — 255 с.

2. Дюк, В. А. «Компьютерная психодиагностика». URL: <http://psyfactor.org/lib/dyuk5.htm> (дата обращения: 28.01.2012)

3. The OEIS Foundation Inc. URL: <http://oeis.org/A008302> (дата обращения: 27.01.2012)

4. Даничев, А.А. Методы и алгоритмы обработки данных в порядковых шкалах для систем поддержки принятия решений: дис. ... канд. техн. наук: 05.13.01 – Красноярск: КГТУ, 2005. 127 с.