

МАТЕМАТИЧЕСКОЕ ОБОСНОВАНИЕ КОМПЬЮТЕРНОЙ ПРОГРАММЫ САЗЫКИНА RANG

Был проведен математический анализ правомерности использования для целей комплексной сравнительной характеристики различных субъектов предложенного В.Л. Сазыкиным алгоритма, реализованного в авторской компьютерной программе Rang. Факторный анализ и ступенчатый регрессионный метод показали правильность предложенного Сазыкиным алгоритма, его достаточную достоверность при относительной простоте, что делает его интуитивно понятным для практических специалистов.

Необходимость комплексной сравнительной оценки деятельности различных субъектов возникает очень часто. В некоторых сферах деятельности, например таких, как медицина, важна именно комплексная сравнительная оценка, поскольку отдельные показатели (критерии деятельности) субъективны и неточны. Для целей сравнительной комплексной оценки одним из нас в 1994 году была предложена компьютерная программа Rang [1, 2], в основу которой был заложен оригинальный алгоритм. Учитывая то, что программа первоначально предназначалась для практических врачей, требовался простой, наглядный и вместе с тем достаточно достоверный алгоритм сравнительной оценки. Суть предложенного алгоритма программы Rang заключалась в следующем. Программа расставляет сравниваемые субъекты по местам сначала по каждому показателю, а в итоге суммирует места, полученные по всем показателям (критериям) и расставляет сравниваемые субъекты по полученным цифрам суммы мест от лучшего к худшему. Для каждого показателя предусмотрен «вес» (по умолчанию – 1) от 1 до 9.

Целью данной работы было проверить с помощью математических методов объективность, достоверность, а значит правомерность использования программы В.Л. Сазыкина Rang. В качестве примера применения программы Rang мы провели сравнительный анализ эпидемиологической обстановки по туберкулезу (распределение по степени благополучия) в территориях, входящих в Приволжский федеральный округ. Затем с помощью математических методов [3-7] проверялась степень достоверности программы.

В таблице 1 приведены параметры 12 эпидемиологических показателей для каждой из 14 территорий Приволжского федерального округа, которые послужили основой для анализа.

Результаты обработки параметров таблицы 1 с помощью программы Rang приведены в таблице 2. Полученная сумма мест для каждой сравниваемой территории, которая в программе

Rang служит основой для распределения сравниваемых субъектов «от лучшего к худшему», была учтена нами в качестве самостоятельного параметра в математическом анализе.

Для математического обоснования разработанной методики оценок была построена матрица исследования со следующими параметрами-столбиками:

1. Заболеваемость туберкулезом общая (по ф. N8)
2. Заболеваемость туберкулезом (по ф. N33)
3. Заболеваемость детей туберкулезом
4. Смертность от акт. туберкулеза (по ГКС)
5. Распространенность ФКТ
6. Болезненность туберкулезом общая
7. Смертность от акт. туберкулеза (по ф. N33)
8. Абациллировано в/в, в %
9. Бациллярность, на 100 тыс. населения
10. Доля ФКТ среди в/в
11. Заболеваемость подростков
12. Распространенность деструктивного туберкулеза
13. Балл по программе Сазыкина Rang – сумма мест (далее «балл по программе Rang»)

В качестве строчек-наблюдений рассматривались 14 регионов Приволжского федерального округа.

Проведенный на такой матрице исследования факторный анализ [3-5] выявил следующие групповые обусловленности: число факторов – 5.

Таблица. Сумма квадратов нагрузок по факторам

| Номер фактора | Сумма квадратов нагрузок |
|---------------|--------------------------|
| 1 | 4,854 |
| 2 | 2,154 |
| 3 | 1,632 |
| 4 | 1,282 |
| 5 | 2,234 |

Таблица. Объединение по фактору 1

| Номер | Название параметра | Нагрузка |
|-------|---|----------|
| 1 | Заболеваемость туберкулезом общая (по ф. 8) | 0,8992 |
| 2 | Заболеваемость туберкулезом (по ф. 33) | 0,9043 |
| 6 | Болезненность туберкулезом общая | 0,8656 |
| 9 | Бациллярность (на 100 тыс. населения) | 0,6244 |
| 12 | Распространенность деструктивного туберкулеза | 0,8346 |
| 13 | Балл по программе Rang | 0,7420 |

Таблица 1. Параметры эпидобстановки по туберкулезу на территории Приволжского федерального округа в 2003 г.

| № | Показатели | Р-ка Башкортостан | Кировская обл. | Р-ка Марий Эл | Р-ка Мордовия | Нижегородская обл. | Оренбургская обл. | Пензенская обл. | Пермская обл. | Самарская обл. | Саратовская обл. | Р-ка Татарстан | Р-ка Удмуртия | Ульяновская обл. | Р-ка Чувашия |
|----|---|-------------------|----------------|---------------|---------------|--------------------|-------------------|-----------------|---------------|----------------|------------------|----------------|---------------|------------------|--------------|
| 1 | Заболеваемость туберкулезом общая (по ф. N8) | 59,3 | 62,6 | 59,3 | 80,4 | 83,3 | 87,3 | 70,4 | 94,1 | 75 | 81,6 | 65,5 | 80,3 | 67,5 | 71,2 |
| 2 | Заболеваемость туберкулезом (по ф. N33) | 47,1 | 52,5 | 53,8 | 70,4 | 66 | 72,2 | 59,5 | 75,1 | 57 | 68,2 | 57,8 | 68 | 53,3 | 58,3 |
| 3 | Заболеваемость детей туберкулезом | 6 | 13 | 5,4 | 12,9 | 22,6 | 12,6 | 13,2 | 6,8 | 17,1 | 14,8 | 15 | 8,9 | 10,7 | 7,6 |
| 4 | Смертность от акт. туберкулеза (по ГКС) | 12,2 | 9,7 | 10,2 | 14 | 22,3 | 15,9 | 15,8 | 21,2 | 19,6 | 17,2 | 12,2 | 19,4 | 15,1 | 12 |
| 5 | Распространенность ФКТ | 19,2 | 7,1 | 8,2 | 31,3 | 31,6 | 11,6 | 19,1 | 30,4 | 22,3 | 20,4 | 12,6 | 14 | 16,1 | 14 |
| 6 | Блезненность туберкулезом общая | 177,7 | 226,9 | 100,2 | 249,2 | 261,2 | 255,1 | 216,3 | 291,4 | 215,4 | 314,7 | 212 | 264,9 | 230,4 | 255,6 |
| 7 | Смертность от акт. туберкулеза (по ф. N33) | 10,6 | 12,3 | 9,8 | 10,4 | 18,3 | 9,6 | 13,8 | 20 | 13 | 13,1 | 11,6 | 20,1 | 13,7 | 13 |
| 8 | Абациллировано в/в, в % | 87,3 | 75,1 | 82,5 | 75,4 | 62,6 | 89 | 78,8 | 75,9 | 75,2 | 77,8 | 73,4 | 64 | 66,1 | 63 |
| 9 | Бациллярность, на 100 тыс. населения | 57,5 | 85,7 | 74,9 | 62 | 82 | 84,6 | 69 | 114,2 | 79,2 | 87,8 | 69,7 | 87,2 | 66 | 82,1 |
| 10 | Доля ФКТ среди в/в | 1,7 | 0,7 | 0,8 | 8,5 | 6,5 | 0,2 | 2,5 | 0,4 | 1,6 | 0,9 | 0,2 | 1 | 3 | 1,4 |
| 11 | Заболеваемость подростков | 16,7 | 18,4 | 20,1 | 39 | 29,1 | 33,4 | 19,2 | 33,5 | 31,7 | 26,4 | 19,4 | 14,4 | 19,5 | 14,4 |
| 12 | Распространенность деструктивного туберкулеза | 49,5 | 48,4 | 37,8 | 72,7 | 90,7 | 82,5 | 63,2 | 90,7 | 65,5 | 73,2 | 55,9 | 80,5 | 70,2 | 80,5 |

Согласно суммам квадратов нагрузок по факторам наибольший вклад в общую дисперсию дает первый фактор. В этот же фактор входит и балл из программы Rang – так называемая сумма мест (таблица – объединение по фактору 1). То есть он определяет в большей степени природу поведения общей дисперсии (в данном случае по эпидобстановке по туберкулезу).

Таким образом, результаты качественной групповой обусловленности, полученные факторным анализом, выявили правильность алгоритма, предложенного В.Л. Сазыкиным для комплексной сравнительной оценки объектов и реализованного в компьютерной программе Rang.

На базе ступенчатого регрессионного метода [6, 7] были получены регрессионные модели для основных показателей эпидобстановки по туберкулезу, в которых параметром-аргументом был балл, полученный по программе Rang (сумма мест).

Модель 1. Заболеваемость туберкулезом общая (по ф. N8)

Значимый по вкладу параметр 13 – балл по программе Rang

$$y = +(46.7987)*(x_{13})^{*0} + (.2949)*(x_{13})^{*1} + (.0003)*(x_{13})^{*2} + (.0000)*(x_{13})^{*3}$$

где * – умножение, ** – возведение в степень.

| Характеристики модели | Значения |
|----------------------------|----------|
| Коэффициент детерминации | 0,89 |
| Средняя абсолютная ошибка | 3,39 |
| Средняя ошибка в процентах | 4,57 |

Модель 2. Заболеваемость туберкулезом (по ф. N33)

Значимый по вкладу параметр 13 – балл по программе Rang

$$y = +(39.0804)*(x_{13})^{*0} + (.2477)*(x_{13})^{*1}$$

где * – умножение, ** – возведение в степень.

| Характеристики модели | Значения |
|----------------------------|----------|
| Коэффициент детерминации | 0,85 |
| Средняя абсолютная ошибка | 3,79 |
| Средняя ошибка в процентах | 6,17 |

Таблица 2. Распределение территории Приволжского федерального округа по благополучию эпидобстановки по туберкулезу

Наилучшая возможная сумма мест: 12
 Наихудшая возможная сумма мест: 168

| Итоговое место | Сравниваемые субъекты | Сумма мест |
|----------------|-----------------------|------------|
| 1 | Р-ка Марий Эл | 36,5 |
| 2 | Р-ка Башкортостан | 43,0 |
| 3 | Кировская область | 58,0 |
| 4 | Р-ка Татарстан | 65,0 |
| 5-6 | Пензенская обл. | 85,0 |
| 5-6 | Ульяновская обл. | 85,0 |
| 7 | Чувашская р-ка | 86,0 |
| 8 | Оренбургская обл. | 91,5 |
| 9 | Самарская обл. | 101,5 |
| 10 | Р-ка Мордовия | 105,0 |
| 11 | Удмуртская р-ка | 109,5 |
| 12 | Саратовская обл. | 118,0 |
| 13 | Пермская обл. | 131,5 |
| 14 | Нижегородская обл. | 144,5 |

Модель 4. Смертность от акт. туберкулеза (по ГКС)

Значимый по вкладу параметр 13 – балл по программе Rang

$$y = +(5.2805)*(x_{13})^{**0} + (.1137)*(x_{13})^{**1}$$

где * – умножение, ** – возведение в степень.

| Характеристики модели | Значения |
|----------------------------|----------|
| Коэффициент детерминации | 0,89 |
| Средняя абсолютная ошибка | 1,49 |
| Средняя ошибка в процентах | 9,6 |

Модель 6. Болезненность туберкулезом общая
 Значимый по вкладу параметр 13 – балл по программе Rang

$$y = +(112.9348)*(x_{13})^{**0} + (1.3444)*(x_{13})^{**1}$$

где * – умножение, ** – возведение в степень.

| Характеристики модели | Значения |
|----------------------------|----------|
| Коэффициент детерминации | 0,82 |
| Средняя абсолютная ошибка | 22,23 |
| Средняя ошибка в процентах | 9,51 |

Модель 12. Распространенность деструктивного туберкулеза
 Значимый по вкладу параметр 13 – балл по программе Rang

$$y = +(27.0500)*(x_{13})^{**0} + (.4635)*(x_{13})^{**1}$$

где * – умножение, ** – возведение в степень.

| Характеристики модели | Значения |
|----------------------------|----------|
| Коэффициент детерминации | 0,90 |
| Средняя абсолютная ошибка | 5,57 |
| Средняя ошибка в процентах | 8,11 |

Во всех представленных моделях для показателей эпидобстановки по туберкулезу мы имеем высокий коэффициент детерминации (≥ 0.8) и небольшие средние абсолютные и относительные ошибки. Это говорит о правильности предложенного Сазыкиным В.Л. алгоритма комплексной сравнительной оценки субъектов, реализованного в компьютерной программе Rang с точки зрения количественных характеристик регрессионных моделей.

ВЫВОДЫ

Результаты качественной групповой обусловленности, полученные факторным анализом, выявили правильность алгоритма, предложенного В.Л. Сазыкиным для комплексной сравнительной оценки объектов и реализованного в компьютерной программе Rang. С точки зрения количественных характеристик регрессионных моделей можно сделать аналогичный вывод о правильности и достоверности предложенного В.Л. Сазыкиным алгоритма комплексной сравнительной оценки различных субъектов, которые можно сравнивать по одним и тем же показателям.

Список использованной литературы:

1. Сазыкин В.Л., Рогачев С.И. Компьютерная программа ранжирования сравниваемых объектов («Rang»). Рационализаторское предложение. Удостоверение №1213 от 23.12.1994
2. Сазыкин В.Л. Сравнительная оценка противотуберкулезной работы с помощью компьютерной программы RANG. IV(XIV) Съезд научно-медицинской ассоциации фтизиатров, тезисы докладов, 1999 г., Йошкар-Ола.
3. Харман Г. Современный факторный анализ. М., Статистика, 1972.
4. Иберла К. Факторный анализ. М., Статистика, 1980.
5. Kaiser H.F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, 23, 187-200, 1958.
6. Драйпер Н., Смит Г.. Прикладной регрессионный анализ. М., Статистика, 1973.
7. Brandon D.B. Developing Mathematical Models for Computer Control, USA Journal, V.S, №7, 1959.