

# РАСЧЕТ ПАРАМЕТРОВ МОДЕЛИ ПЕДАГОГИЧЕСКИХ ИЗМЕРЕНИЙ

**О.Г. Деменченко,**

начальник кафедры математики  
и информатики  
ФГОУ ВПО ВСИ МВД России,  
кандидат технических наук, доцент

*Рассмотрены теоретические основы оценки параметров модели измерения, предложены аналитические зависимости для модели Г. Раша. Описана технология поиска параметров модели Г. Раша средствами Microsoft Excel. Проведен анализ результатов подбора параметров модели и анализ качества моделей.*

*Theoretical bases of an estimation of parameters of model of measurement are considered, analytical dependences for G.Rasch model are offered. The technology of search of parameters of model of G.Rasch is described by means Microsoft Excel. The analysis of results of selection of parameters of model and the analysis of quality of models is carried out\*.*

Специфической чертой современного этапа развития теории и практики педагогических измерений следует считать широкое внедрение методов математического моделирования. Математические модели способствуют выполнению двух основных функций теории - объяснению и прогнозированию наблюдаемых результатов тестирования.

---

\* Demenchenok O.G. Calculation of parameters of model of pedagogical measurements.

Для практического применения моделей в педагогических измерениях необходимо, прежде всего, определить параметры модели. Затем неизбежно возникает вопрос об адекватности полученной модели – насколько хорошо модель справляется со своими функциями? Подбору параметров и оценке качества модели посвящена данная статья.

*Теоретические основы оценки  
параметров модели измерения*

Задача построения математической модели, связывающей уровни подготовленности испытуемых и трудности тестовых заданий с конкретными результатами выполнения теста, может быть сформулирована следующим образом. Из теоретических соображений выбран принципиальный вид зависимости вероятности правильного ответа. Эта зависимость – та или иная модель Item Response Theory (IRT).

Модель содержит ряд параметров:  $\theta$  – подготовленность испытуемых,  $\beta$  – трудность заданий,  $a$  – значение параметра крутизны графика функции задания или различающая способность задания для двух- и трехпараметрических моделей,  $c$  – мера возможного угадывания правильного ответа на задание, для трехпараметрической модели. Требуется так выбрать эти параметры, чтобы результаты конкретного тестирования и математическая модель наилучшим образом совпадали.

Решение этой задачи во многом зависит от того, что именно мы условимся считать «наилучшим», т.е. от принятого критерия оптимальности. Можно, например, считать критерием оптимальности минимум максимального отклонения между расчетными и экспериментальными данными; можно потребовать, чтобы в минимум обращалась сумма абсолютных величин отклонений и т.д. При каждом из этих требований мы получим свое решение задачи, свои значения параметров  $\theta$  и  $\beta$ .

Теоретическую базу всех методов и приемов, положенных в основу построения эмпирических математических моделей, составляет метод максимального (наибольшего) правдоподобия. В общем виде метод максимума правдоподобия можно сформулировать так: наилучшее описание явления то, которое дает наибольшую вероятность получить в результате измерений именно те значения, которые и были фактически получены [1].

В идеальном случае математическая модель полностью соответствует экспериментальным данным. Однако в действительности результаты измерения всегда содержат погрешности. В связи с этим имеет место множество гипотез, не противоречащих опытным данным. Метод максимального правдоподобия состоит в выборе такой гипотезы, при которой вероятность получить в процессе измерения фактически

наблюдаемые величины была бы максимальной. В формальной записи этот метод может быть представлен в виде максимума произведения вероятностей всех наблюдаемых независимых событий [2]:

$$p_1 \cdot p_2 \cdot \dots \cdot p_N = \prod_{i=1}^N p_i \rightarrow \max \quad (1)$$

Ещё один известный метод – метод наименьших квадратов – это метод теории ошибок для оценки неизвестных величин по результатам измерений, содержащим случайные ошибки. По этому методу требование наилучшего согласования расчетных  $x$  и экспериментальных  $x_0$  данных сводится к тому, чтобы сумма квадратов отклонений между ними обращалась в минимум:

$$\sum_{i=1}^N (x_i - x_{0i})^2 \rightarrow \min \quad (2)$$

Метод наименьших квадратов приводит к сравнительно простому математическому способу оценки параметров; кроме того, он допускает довольно веское теоретическое обоснование с вероятностной точки зрения.

#### *Критерии подбора параметров*

##### *модели Раша по методу максимума правдоподобия*

Вернемся к задаче подбора параметров модели. Предположим, из теоретических соображений выбрана модель Г. Раша. По этой модели вероятность результата  $x$  при решении  $i$ -ым тестируемым  $j$ -го задания определяется выражением:

$$P_{ij}(x) = \frac{e^{x_{ij}(\theta_i - \beta_j)}}{1 + e^{\theta_i - \beta_j}} \quad (3)$$

где  $x_{ij}$  – результат выполнения задания,  $x = 1$  при правильном ответе,  $x = 0$  при неправильном ответе;  $\theta_i$  – уровень подготовленности  $i$ -го тестируемого;  $\beta_j$  – уровень трудности  $j$ -го задания.

По методу максимального правдоподобия наилучшим будет признан тот набор значений  $\theta$  и  $\beta$ , при котором произведение расчетных вероятностей фактически полученных результатов максимально:

$$\prod_{i=1}^n \prod_{j=1}^m P_{ij}(x) = \prod_{i=1}^n \prod_{j=1}^m \frac{e^{x_{ij}(\theta_i - \beta_j)}}{1 + e^{\theta_i - \beta_j}} \rightarrow \max \quad (4)$$

где  $n$  – количество испытуемых;  $m$  – число тестовых заданий.

Особенность модели Г. Раша заключается в том, что она оперирует только разностью параметров  $\theta$  и  $\beta$ . Поэтому можно получить бесконечное количество «наилучших» решений, отличающихся на произвольную постоянную величину:

$$\theta - \beta = (\theta + C) - (\beta + C) = \theta - \beta.$$

Для определенности целесообразно ввести ограничение, например, в

виде равенства нулю среднего значения уровня подготовленности всех испытуемых:

$$\bar{\theta} = 0. \quad (5)$$

В этом случае наилучшее решение будет единственным. Для реализации в вычислительных алгоритмах желательно преобразовать ограничение (5) в неравенство:

$$|\bar{\theta}| \leq \varepsilon, \quad (6)$$

где  $\varepsilon$  – некоторая допустимо малая величина, например равная 0,001.

Выражений (4) и (6) достаточно для непосредственного численного решения задачи подбора параметров модели.

По методу наибольшего правдоподобия можно получить ещё один вариант решения, основанный на том, что в точке экстремума (максимума или минимума) производная функции равна нулю.

Зная, что  $e^a e^b = e^{a+b}$ , преобразуем числитель условия (4):

$$\frac{e^{\sum_{i=1}^n \sum_{j=1}^m x_{ij} (\theta_i - \beta_j)}}{\prod_{i=1}^n \prod_{j=1}^m (1 + e^{\theta_i - \beta_j})} \rightarrow \max \quad (7)$$

Введём обозначения:  $R_i$  – сумма результатов выполнения тестовых заданий  $i$ -м тестируемым ( $R_i$  также называют исходным баллом):

$$R_i = \sum_{j=1}^m x_{ij} \quad (8)$$

$s_j$  – сумма результатов выполнения  $j$ -го задания всеми тестируемыми

$$s_j = \sum_{i=1}^n x_{ij} \quad (9)$$

Используя (8) и (9), заменим

$$\sum_{i=1}^n \sum_{j=1}^m x_{ij} \theta_i = \sum_{i=1}^n R_i \theta_i \quad (10)$$

$$\sum_{i=1}^n \sum_{j=1}^m x_{ij} \beta_j = \sum_{j=1}^m s_j \beta_j \quad (11)$$

Таким образом, уравнение (7) принимает вид

$$\frac{e^{\sum_{i=1}^n R_i \theta_i - \sum_{j=1}^m s_j \beta_j}}{\prod_{i=1}^n \prod_{j=1}^m (1 + e^{\theta_i - \beta_j})} \rightarrow \max \quad (12)$$

Так как логарифм – монотонно возрастающая функция, то максимум функции правдоподобия совпадает с максимумом ее логарифма:

$$\ln \left( \frac{e^{\sum_{i=1}^n R_i \theta_i - \sum_{j=1}^m s_j \beta_j}}{\prod_{i=1}^n \prod_{j=1}^m (1 + e^{\theta_i - \beta_j})} \right) \rightarrow \max \quad (13)$$

Для нахождения максимума функции правдоподобия найдем и приравняем нулю частные производные логарифмической функции правдоподобия по каждому из аргументов:

$$\frac{\partial \left( \sum_{i=1}^n R_i \theta_i - \sum_{j=1}^m s_j \beta_j - \sum_{i=1}^n \sum_{j=1}^m \ln(1 + e^{\theta_i - \beta_j}) \right)}{\partial \theta_i} = R_i - \sum_{j=1}^m \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}} = 0$$

или  $R_i - \sum_{j=1}^m \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}} = 0$  (14)

$$\frac{\partial \left( \sum_{i=1}^n R_i \theta_i - \sum_{j=1}^m s_j \beta_j - \sum_{i=1}^n \sum_{j=1}^m \ln(1 + e^{\theta_i - \beta_j}) \right)}{\partial \beta_j} = -s_j + \sum_{i=1}^n \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}} = 0,$$

или  $s_j - \sum_{i=1}^n \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}} = 0.$  (15)

Нетрудно заметить, что выражение (14) представляет собой разность между суммой результатов выполнения заданий  $i$ -м испытуемым и суммой вероятностей правильных ответов на все задания для того же испытуемого. Аналогично, уравнение (15) представляет собой разность между суммой результатов выполнения задания всеми тестируемыми и суммой вероятностей правильных ответов на то же задание для всех испытуемых.

В итоге мы получаем второй вариант решения по методу наибольшего правдоподобия. Он состоит из  $n$  уравнений (14) (по одному уравнению на каждого тестируемого для  $i=1, 2 \dots n$ ) и  $m$  уравнений (15) (по одному на каждое задание), которые нужно численно решить с учетом дополнительного ограничения, например (6), для нахождения параметров модели Г. Раша. Наличие аналитических зависимостей для частных производных упрощает и существенно ускоряет поиск решения. Поэтому многие алгоритмы подбора параметров модели Г. Раша основаны на уравнениях (14) и (15).

#### *Критерий подбора параметров модели по методу наименьших квадратов*

Сформулируем условие оптимальности подбора параметров модели, используя метод наименьших квадратов. По этому методу приближённые значения параметров находятся из условия минимизации суммы квадратов отклонений между моделью и наблюдаемыми результатами.

Отклонение (различие) между моделью и наблюдаемыми результатами может быть найдено как разность результата выполнения задания  $x_{ij}$  и вероятности правильного ответа  $P_{ij}$ :

для правильного ответа  $x_{ij} - P_{ij} = 1 - P_{ij}$ ,

для неправильного ответа  $x_{ij} - P_{ij} = 0 - P_{ij} = -P_{ij}$ .

Разность  $x_{ij} - P_{ij}$  показывает, насколько расчетная вероятность правильного ответа отличается от идеального значения: для правильного ответа  $P_{ij}$  должна стремиться к единице, а для неправильного ответа – к нулю.

Тогда критерий минимума квадратов отклонений для обеих рассматриваемых моделей формулируется в виде условия:

$$\sum_{i=1}^n \sum_{j=1}^m \left( x_{ij} - \frac{1}{1 + e^{-a_j(\theta_i - \beta_j)}} \right)^2 \rightarrow \min \quad (16)$$

В ходе пробных расчетов выявлено, что в ходе поиска решения иногда отмечается существенный рост значений  $\theta$ . Поэтому условие (6) дополнено вторым ограничением, устанавливающим пределы изменения  $\theta$  от  $-5$  до  $+5$ :

$$|\theta|_{\max} \leq 5 \quad (17)$$

*Подбор параметров модели в электронной таблице Microsoft Excel.*

Все расчеты могут быть выполнены в Microsoft Excel, доступном и широко известном средстве автоматизации вычислений. Excel способен находить экстремум функции одной или нескольких переменных. Кратко опишем технологию поиска параметров модели Г. Раша по методу максимального правдоподобия для представленных в табл.1 результатов тестирования (принцип решения остальных задач аналогичен). Сначала вводим данные табл.1 в клетки В2:О35 (рис.1).

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>G</b>	<b>H</b>	<b>I</b>	<b>J</b>	<b>K</b>	<b>L</b>	<b>M</b>	<b>N</b>	<b>O</b>
<b>1</b>		<b>4</b>	<b>5</b>	<b>7</b>	<b>6</b>	<b>9</b>	<b>8</b>	<b>10</b>	<b>11</b>	<b>13</b>	<b>12</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>
<b>2</b>	<b>25</b>	0	1	0	1	0	0	0	0	0	0	0	0	0	0
<b>3</b>	<b>4</b>	1	0	1	0	1	0	0	0	0	0	0	0	0	0
<b>4</b>	<b>33</b>	1	0	1	0	0	0	1	0	0	0	0	0	0	0
<b>5</b>	<b>1</b>	1	1	1	1	0	0	0	0	0	0	0	0	0	0

Рис.1. Ввод исходных данных

Находим сумму баллов каждого тестируемого: в клетку Р2 вводим =СУММ(В2:О2) и копируем эту формулу в клетки Р3:Р35. Для суммы баллов по каждому заданию в В36 вводим =СУММ(В2:В35) и копируем в С36:О36.

Рассчитываем начальные значения параметров модели:

- для уровней подготовленности испытуемых в Q2 вводим  $=\text{LN}(\text{P2}/(\text{14}-\text{P2}))$  и копируем в клетки Q3:Q35 (14 – число тестовых заданий);
- для уровней трудности заданий в B37 вводим  $=\text{LN}((\text{34}-\text{B36})/\text{B36})$  и копируем в C37:O37 (34 – количество испытуемых).

Определяем вероятности результатов выполнения тестовых заданий, для чего в B42 вводим  $=\text{ЕСЛИ}(\text{B2}>0;1/(\text{1}+\text{EXP}((-1)*(\$R2-\text{B}\$38)));1-1/(\text{1}+\text{EXP}((-1)*(\$R2-\text{B}\$38))))$  и копируем в клетки B42:O75. Если значение в клетке B2 больше нуля (задание выполнено правильно), то рассчитывается вероятность правильного ответа по формуле  $1/(\text{1}+\text{EXP}((-1)*(\$R2-\text{B}\$38)))$ ; в противном случае – вероятность неправильного ответа  $1-1/(\text{1}+\text{EXP}((-1)*(\$R2-\text{B}\$38)))$ . Для расчета параметров двухпараметрической модели вместо сомножителя (-1) следует записать адрес клетки со значением дифференцирующей способности задания.

Зададим параметры модели, равные начальным:

- выделим начальные значения уровней подготовленности испытуемых Q2:Q35 и вставим значения в R2:R35, для чего нажмем кнопку *Копировать*, щелкнем в клетке R2 и выберем в меню *Правка – Специальная вставка – Значения*;
- для уровней трудности выделяем B37:O37 и вставляем значения в B38:O38.

В клетку R36 введем формулу модуля среднего значения уровней подготовленности  $=\text{ABS}(\text{CPЗНАЧ}(\text{R2}:\text{R35}))$ .

Найдем произведение всех вероятностей: в R38 введем  $=\text{ПРОИЗВЕД}(\text{B42}:\text{O75})$ . В клетке R38 появляется 1,66431E-58, т.е. начальное значение вероятности всех ответов очень мало –  $1,66 \cdot 10^{-58}$ . Так как механизм поиска решения Excel требует, чтобы значения критерия и, следовательно, его изменения при подборе параметров не были бы слишком малы, то удобнее искать максимум логарифмической функции правдоподобия. Для этого в клетку R39 введем  $=\text{Ln}(\text{R38})$ .

Проверим, установлен ли механизм поиска решения: выберем в меню *Сервис – Надстройки*, поставим отметку в пункте *Поиск решения* и нажмем *ОК*. Далее нужно выполнить команду *Сервис – Поиск решения* и заполнить диалоговое окно (рис.2):

- 1) кликнуть левой клавишей мыши в поле *Установить целевую ячейку* (т.е. клетку, значение которой необходимо максимизировать или минимизировать), переместить указатель мыши и кликнуть на ячейке R39 с формулой критерия оптимальности;
- 2) указать направление поиска, отметив пункт *Равной максимальному значению*;

- 3) в поле *Изменяя ячейки* ввести адреса ячеек, значения которых будут варьироваться в процессе поиска решения. В нашем случае это R2:R35; V38:O38 (удобно вводить, выделяя клетки с помощью мыши, второй диапазон при нажатой клавише *Ctrl*);
- 4) с помощью кнопки *Добавить* ввести ограничение: значение  $V36 \leq 0,001$ ;
- 5) нажав кнопку *Параметры*, указать параметры поиска в соответствии с рис.3. Это повысит точность подбора параметров.

После щелчка на кнопке *OK* получим решение поставленной задачи. В клетке R38 находится максимальное значение вероятности всех данных, равное  $5,09 \cdot 10^{-48}$ . Значения уровней подготовленности, найденные по методу наибольшего правдоподобия, записаны в клетках R2:R35, уровней трудности заданий – в ячейках V38:O38.

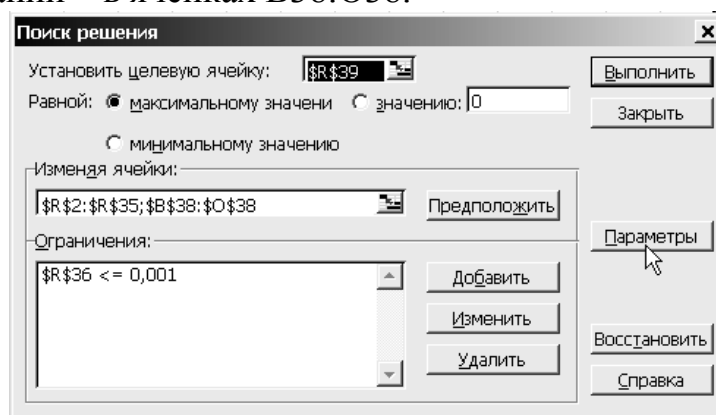


Рис.2. Окно поиска решения

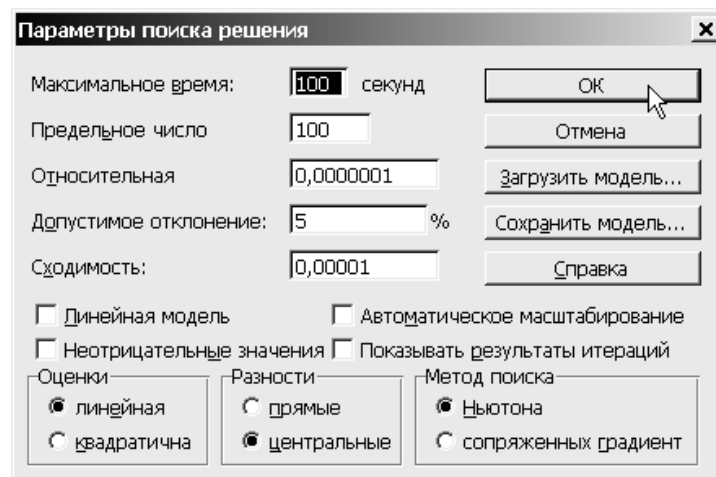


Рис.3. Рекомендуемые параметры поиска решения



### *Анализ результатов подбора параметров модели*

Насколько отличаются параметры модели измерения, найденные по разным методикам? Для ответа на этот вопрос сравним результаты решения одной и той же задачи, полученные разными способами. Для примера взята задача подбора параметров модели Г. Раша из известной работы Б.Д.Райта и М.Х. Стоуна [3].

Анализ результатов подбора параметров модели свидетельствует:

- 1) значения параметров существенно зависят от выбранной модели и метода расчета;
- 2) характер изменения параметров  $\theta$  и  $\beta$  для всех рассмотренных моделей и методов подобен. Это подтверждает корреляционный анализ – коэффициент корреляции для уровней подготовленности составляет от 0,92 до 0,99, для уровней трудности от 0,94 до 0,99, что свидетельствует об очень сильной статистической взаимосвязи;
- 3) значения дифференцирующей способности заданий, найденные по методу максимального правдоподобия и программой Winsteps хорошо согласованы (коэффициент корреляции 0,90). Однако эти значения слабо согласованы с результатами по методу наименьших квадратов (коэффициент корреляции 0,17);
- 4) равным начальным уровням подготовленности испытуемых (т.е. одинаковым суммам результатов тестирования  $R_i$ ) соответствуют равные итоговые значения по всем методикам расчета для модели Г. Раша, за исключением метода наименьших квадратов;
- 5) аналогично, равным начальным уровням трудности тестовых заданий (т.е. одинаковым суммам выполнения задания всеми тестируемыми  $s_j$ ) соответствуют равные итоговые значения по всем методикам расчета для модели Г. Раша, за исключением метода наименьших квадратов;
- 6) для двухпараметрической модели равным начальным уровням  $\theta$  и  $\beta$  могут соответствовать разные итоговые значения, как по методу максимального правдоподобия, так и по методу наименьших квадратов.

#### **Литература**

1. Львовский Б.Н. Статистические методы построения эмпирических формул. - М.: Высшая школа, 1988.
2. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. - М.: ЮНИТИ, 1998.
3. Wright B.D., Stone M.H. Best Test Design. - Chicago: MESA PRESS. 1979.