Методология

АНАЛИЗ МОДЕЛЕЙ ДЛЯ АДАПТИВНОГО ТЕСТИРОВАНИЯ

Олег Деменчёнок
Восточно-Сибирский институт МВД России
AskSystem@yandex.ru

Проведён анализ эффективности моделей педагогического измерения в условиях адаптивного тестирования. Показано, что наиболее эффективна двухпараметрическая модель при условии достаточного количества заданий с высокой дифференцирующей способностью, равномерно распределённых по всему диапазону измерений. Если указанное условие не выполняется, то предпочтительными являются Partial Credit Model и модель Раша.

Ключевые слова: адаптивное тестирование, математическая модель измерения, информационная функция.

Введение

Одним из наиболее перспективных направлений совершенствования педагогических измерений может стать адаптивное тестирование.

Традиционное тестирование основано на тестах с фиксированной последовательностью заданий. Так как испытуемые имеют раз-

ный уровень подготовленности, то и тестовые задания должны варьироваться по сложности. Очевидно, что для конкретного студента часть заданий может оказаться слишком лёгкой, а некоторые задания будут для него слишком сложными. Ответы на такие задания мало информативны: способность «среднего» студента правильно ответить на простейшие задания, наряду с неспособностью справиться с олимпиадными задачами, не дают надёжной основы для измерения уровня подготовленности этого студента.

Гораздо более информативны ответы на задания, соответствующие уровню подготовленности испытуемого. Адаптивный тест приспосабливается к возможностям испытуемого: при правильном ответе следующее задание будет чуть более трудным, при неправильном ответе — более лёгким. Таким образом, поддерживается примерное равенство уровней подготовленности испытуемого и трудности заданий $\theta \approx \beta$, а средняя вероятность правильного ответа будет близка к 0,5[1].

При адаптивном тестировании испытуемый с высоким уровнем подготовки получит набор трудных заданий, а слабый студент — лёгкие задания. Такой подход к тестированию можно проиллюстрировать соревнованиями по прыжкам в высоту, на которых планка постепенно устанавливается на ту высоту, которую спортсмен потенциально способен преодолеть. При этом результат определяется не количеством удачных попыток, а взятой высотой.

Постановка проблемы

Основная идея адаптивного тестирования заключается в том, чтобы получить максимум информации об уровне подготовленности испытуемого путём подбора наиболее подходящих для этого заданий. За счёт этого можно существенно повысить точность и надёжность педагогического измерения или при той же точности сократить время тестирования. Так, технология адаптивного тестирования позволяет корпорации Microsoft при сертификации специалистов уменьшать количество заданий теста на 60%, существенно сокращая время тестирования[2].

Однако количество информации, полученное из ответов на задания теста, зависит не только от соответствия уровня их сложности подготовленности испытуемого, но и от выбранной математической модели педагогического измерения. Закономерно возникает вопрос: какая из моделей работает наиболее эффективно именно в условиях адаптивного тестирования? Данная статья является попыт-

1 *Аванесов В.С.* Применение тестовых форм в Rasch Measurement //Педагогические измерения. №4, 2006.

2 http://www.mobukom.ru/cit/mcp/adaptive.html

кой автора дать ответ на этот вопрос.

Анализ информационной функции базовых моделей педагогического измерения

Чем больше информации, тем точнее наши сведения, т.е. меньше ошибка. Другими словами, увеличение количества информации означает повышение эффективности тестирования, так как сокращает время тестирования при равной точности педагогического измерения. В Item Response Theory (IRT) количеством информации[3] называют величину, обратную дисперсии ошибок, а информационной функцией — соответствующую аналитическую зависимость:

$$I = \frac{1}{D} = \frac{1}{\sigma^2}. \qquad (1)$$

Для трёх базовых моделей IRT[4,5] количество информации рассчитывается по формуле:

$$I = \frac{1}{\sigma_{\theta_i}^2} = \qquad (2)$$

$$= \sum_{j=1}^{m} a_j{}^2 \left[\left(\frac{1-P_{ij}}{P_{ij}} \right) \left(\frac{P_{ij}-c_j}{1-c_j} \right)^2 \right],$$

где σ_{θ_i} – стандартная ошибка уровня подготовленности i-го испытуемого; m — количество тестовых заданий; P_{ij} — вероятность правильного ответа i-го тестируемого на j-е задание; a_j и c_j — дифференцирующая способность и параметр коррекции на угадывание правильного ответа j-го задания.

Для одного задания выражение (2) примет вид:

$$I = a^2 \left[\left(\frac{1-P}{P} \right) \left(\frac{P-c}{1-c} \right)^2 \right]. \qquad (3)$$

В модели Г.Раша a_j =1, а c_j = 0, что приводит к следующей аналитической зависимости:

$$I = 1^2 \left[\left(\frac{1-P}{P} \right) \left(\frac{P-0}{1-0} \right)^2 \right] = \qquad (4)$$

$$= (1-P)P = \left(1 - \frac{e^{\theta-\beta}}{1+e^{\theta-\beta}} \right) \frac{e^{\theta-\beta}}{1+e^{\theta-\beta}}.$$

Из формулы (4) следует, что количество информации максимально при вероятности правильного ответа P = 0,5:

$$I'(P) = \left((1-P)P\right)' = 1-2P = 0,$$

$$P = 0,5.$$

Максимум количества информации равен I_{max} = 0,25 (рис. 1) и соответствует равенству уровня подготовленности испытуемого и уровня трудности задания $\theta = \beta$ (или $\theta - \beta = 0$):

$$I_{\max} = \left(1 - \frac{e^0}{1+e^0} \right) \frac{e^0}{1+e^0} =$$

$$= \left(1 - \frac{1}{1+1} \right) \frac{1}{1+1} = 0,25.$$

Информативность заданий, существенно отличающихся по уровню трудности от уровня подготовленности испытуемого

3 Количество информации — показатель, характеризующий уменьшение неопределённости состояния системы.

4 *Baker F.B.* The Basics of Item Response Theory. 2 ed., ERIC Clearinghouse on Assessment and Evaluation, Madison, Wisconsin, 2001. 172 p.

5 *Деменчёнок О.Г.* Компьютерная программа для подбора параметров основных моделей IRT. // Педагогические измерения, № 2, 2008.

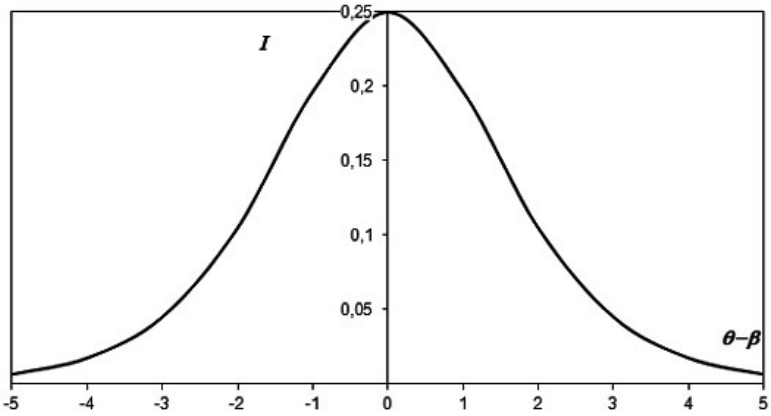

Рис. 1. Зависимость количества информации от разности уровня подготовленности испытуемого и уровня трудности задания для модели Раша

(правый и левый конец графика на рис. 1.), почти нулевая. Действительно, способность испытуемого решать очень простые задания (θ много больше β) или неудачи в решении заданий повышенной сложности (θ значительно меньше β) мало информативны, так как не дают возможности уточнить уровень подготовленности тестируемого.

Для двухпараметрической модели (c_j = 0) выражение (3) примет вид:

$$I = a^2\left[\left(\frac{1-P}{P}\right)\left(\frac{P-0}{1-0}\right)^2\right] =$$

$$= a^2(1-P)P = a^2\left(1-\frac{e^{a(\theta-\beta)}}{1+e^{a(\theta-\beta)}}\right)\cdot$$

$$\cdot\frac{e^{a(\theta-\beta)}}{1+e^{a(\theta-\beta)}} \qquad (5)$$

Нетрудно заметить, что и в этом случае максимум количества информации достигается при $P = 0{,}5$ и равенстве уровня подготовленности испытуемого и уровня трудности задания $\theta = \beta$ (рис. 2):

$$I_{\max} = a^2\left(1-\frac{e^0}{1+e^0}\right)\frac{e^0}{1+e^0} = \qquad (6)$$

$$= a^2\left(1-\frac{1}{1+1}\right)\frac{1}{1+1} = 0{,}25a^2.$$

Однако максимальное значение равно $0{,}25a^2$, т.е. в зависимости от величины дифференцирующей способности задания a максимум может оказаться больше соответствующего значения для модели Раша (при $a > 1$) или меньше его (при $0 < a < 1$).

Например, при дифференцирующей способности задания $a = 2$ максимум количества информации равен $I_{max} = 0{,}25 \cdot 2^2 = 1$, а при $a = 0{,}5$

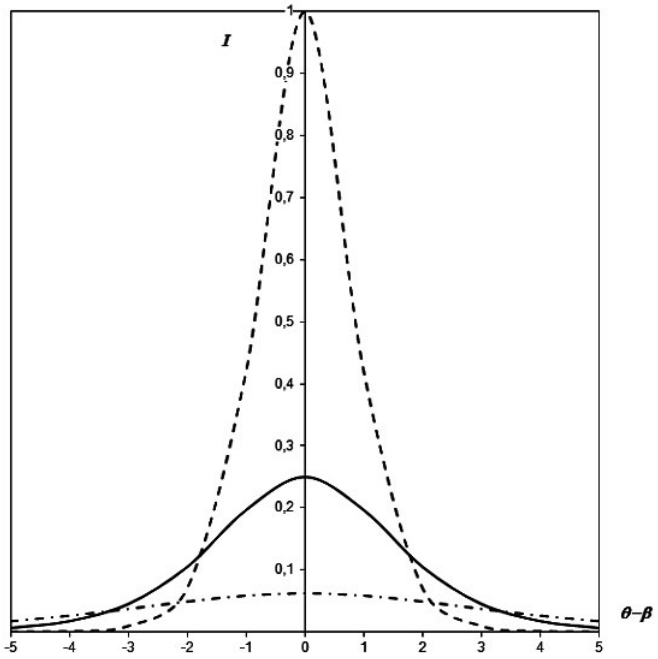

Рис. 2. Зависимость количества информации от разности уровня подготовленности испытуемого и уровня трудности задания для двухпараметрической модели:
– – – – – – – при a = 1, — – — – — при a = 2, — · — · — — при a = 0,5

I_{max} = 0,25·0,5^2 = 0,0625. Таким образом, количество информации существенно зависит от дифференцирующей способности задания. При адаптивном тестировании достигается примерное равенство уровня подготовленности испытуемого и уровня трудности задания $q \approx b$, а количество информации близко к максимуму. Если выбирать задания с дифференцирующей способностью более единицы $a > 1$, двухпараметрическая модель окажется эффективнее модели Раша.

По трёхпараметрической модели вероятность правильного ответа тестируемого равна[6]:

$$P = c + (1-c)\frac{e^{a(\theta-\beta)}}{1+e^{a(\theta-\beta)}} = \\ = c + (1-c)\frac{1}{1+e^{-a(\theta-\beta)}}. \quad (7)$$

Выясним влияние параметра коррекции на угадывание правильного ответа на количество информации. Используя уравнения (3) и (7), проведём расчёты при фиксированном значении дифференцирующей

6

Partchev I. A visual guide to item response theory. Jena: Friedrich-Schiller-Universitat. 2004. 61 p.

способности задания a = 1 и значениях параметра коррекции на угадывание правильного ответа c = 0; 0,2 и 0,4 (результаты приведены на рис. 3).

Расчёты показывают, что увеличение параметра коррекции на угадывание правильного ответа снижает информативность ответа:

- при c = 0 максимальное значение количества информации равно 0,25;
- при c = 0,2 I_{max} = 0,17;
- при c = 0,4 I_{max} = 0,11.

Коррекция на угадывание правильного ответа во многом «съедаёт» повышение информативности ответа за счёт высокой дифференцирующей способности задания (рис. 4).

Так, при значении дифференцирующей способности a = 2:

- при c = 0 максимальное значение количества информации равно 1;
- при c = 0,2 I_{max}=0,68;
- при c = 0,4 I_{max}=0,44.

Таким образом, усложнение модели измерения путём введения третьего параметра — коррекции на угадывание правильного ответа — снижает эффективность измерения по сравнению с двухпараметрической моделью.

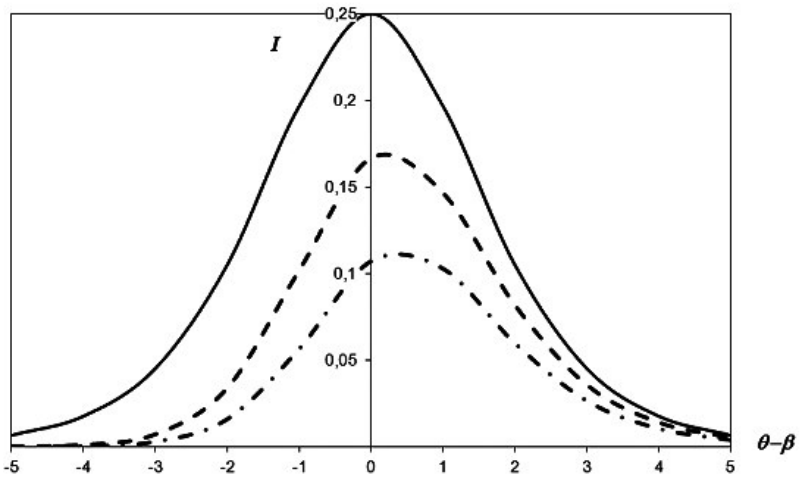

Рис. 3. Зависимость количества информации от разности уровня подготовленности испытуемого и уровня трудности задания для трёхпараметрической модели при a = 1:
–––––––– — c = 0; — – — – — – — c = 0,2; — · — · — · — — c = 0,4

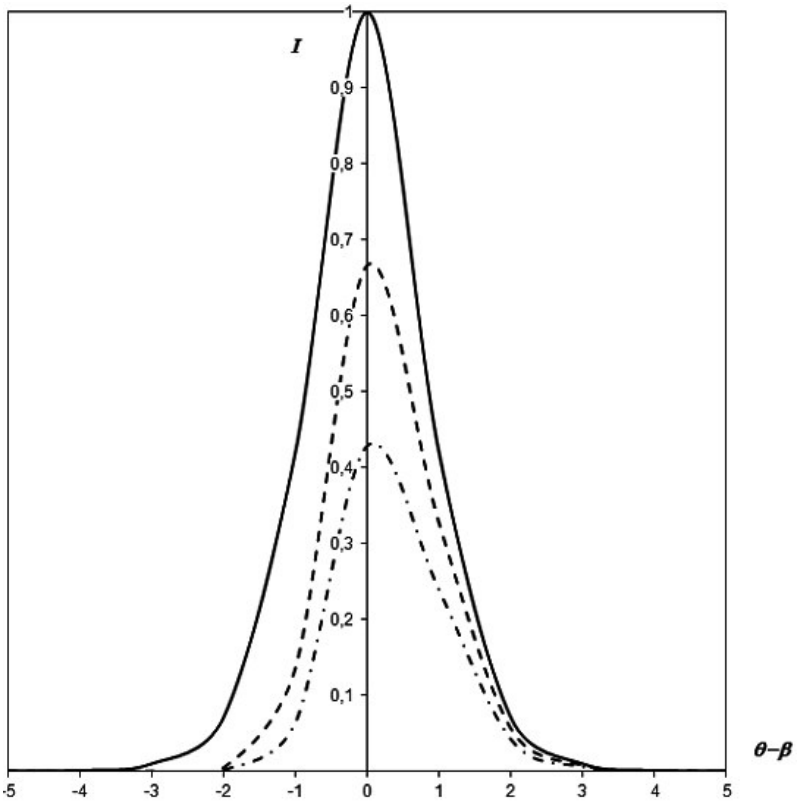

Рис. 4. Зависимость количества информации от разности уровня подготовленности испытуемого и уровня трудности задания для трёхпараметрической модели при $a = 2$: ––––––––– — $c = 0$; — – — – — – — $c = 0{,}2$; — · — · — · — — $c = 0{,}4$

Анализ информационной функции Partial Credit Model

Рассмотренные выше математические модели педагогических измерений применимы только в тех случаях, когда результат выполнения тестового задания оценивается нулём («неправильно») или одним баллом («правильно»). Эти модели лишены возможности учёта частично или не полностью правильных ответов.

От этого ограничения свободны модели с градацией степени правильности ответа. В таких моделях за одно и то же задание можно получить разное количество баллов, в зависимости от полноты и правильности решения.

Partial Credit Model (PCM) — это наиболее известная модификация модели Раша для тестовых заданий с градацией степени правильности ответа. Эта модель выражается аналитической зависимостью[7]:

$$\pi_{ijx} = \frac{e^{\sum_{k=0}^{x}(\theta_i - \beta_{jk})}}{\sum_{l=0}^{x_{\max j}} e^{\sum_{k=0}^{l}(\theta_i - \beta_{jk})}}, \qquad (8)$$

где p_{ijx} — вероятность достижения тестируемым результата x_{ij} (т.е. того, что тестируемый i выполнит ровно x шагов и получит x баллов в задании j); $x = 0, 1 \ldots \ldots x_{ij} \ldots x_{max\,j}$ — количество шагов; $x_{max\,j}$ — максимально возможное количество баллов за задание j; $\beta_{j0} = 0$, $\sum_{n=0}^{0}(\theta_i - \beta_{j0}) = 0$.

Количество информации для Partial Credit Model[8]:

$$I = \frac{1}{\sigma_{\theta i}^{\,2}} = \sum_{j=1}^{m}\left(\sum_{l=1}^{x_{\max j}} l^2 \cdot \pi_{ijl} - \left(\sum_{l=1}^{x_{\max j}} l \cdot \pi_{ijl}\right)^2\right) \qquad (9)$$

или для одного тестового задания:

$$I = \sum_{l=1}^{x_{\max j}} l^2 \cdot \pi_{ijl} - \left(\sum_{l=1}^{x_{\max j}} l \cdot \pi_{ijl}\right)^2. \qquad (10)$$

Например, для заданий, максимально оцениваемых двумя и тремя баллами, уравнение (10) принимает вид:

$$I = \pi_{ij1} + 4\pi_{ij2} - \left(\pi_{ij1} + 2\pi_{ij2}\right)^2, \qquad (11)$$

$$I = \pi_{ij1} + 4\pi_{ij2} + 9\pi_{ij3} - \left(\pi_{ij1} + 2\pi_{ij2} + 3\pi_{ij3}\right)^2. \qquad (12)$$

Количество информации для анализа заданий с большим количеством градаций степени правильности ответа находится аналогичным образом.

Начнём анализ с заданий, максимальная оценка за которые равна двум баллам. Количество информации существенно зависит от близости уровней трудности шагов задания и уровня подготовленности испытуемого: чем они ближе, тем больше получаемое при тестировании количество информации (рис. 5).

Теоретический максимум близок к 0,67:

- при уровнях трудности первого и второго шага $b_1 = -0{,}1$ и $b_2 = 0{,}1$ максимальное значение количества информации I_{max} равно 0,64;
- при $\beta_1 = -1$ и $\beta_2 = 1$ $I_{max} = 0{,}41$;
- при $\beta_1 = 1$ и $\beta_2 = 4$ $I_{max} = 0{,}31$.

Увеличение количества градаций степени правильности ответа, равного максимальному баллу задания, повышает информативность (рис. 6).

Для трёхбалльного задания теоретический максимум информации близок к 1,25:

- при уровнях трудности первого, второго и третьего шага соответственно $\beta_1 = -0{,}1$; $\beta_2 = 0$ и $\beta_3 = 0{,}1$ максимальное значение

7 *Wright B.D., Masters G.N.* Rating Scale Analysis: Rasch Measurement. Chicago: Mesa Press, 1982. 204 p.

8 Там же.

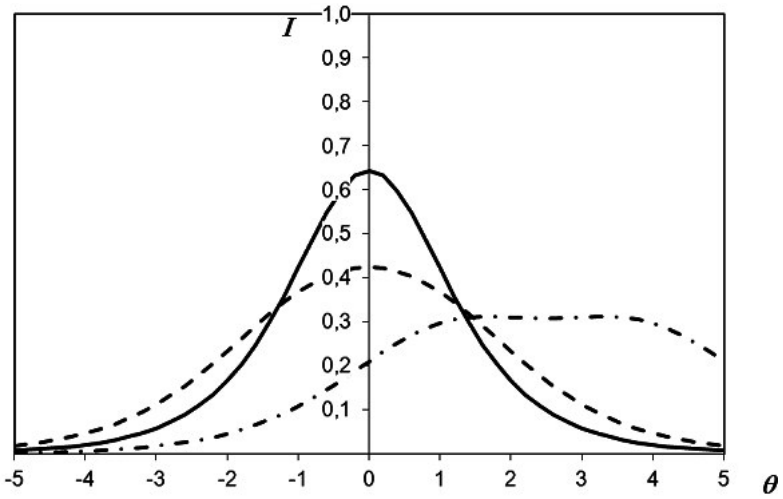

Рис. 5. Зависимость количества информации двухбалльного задания РСМ от уровня подготовленности испытуемого: –––––––– — при $\beta_1 = -0{,}1$ и $\beta_2 = 0{,}1$; — – — – — — при $\beta_1 = -1$ и $\beta_2 = 1$; — · — · — · — — при $\beta_1 = 1$ и $\beta_2 = 4$

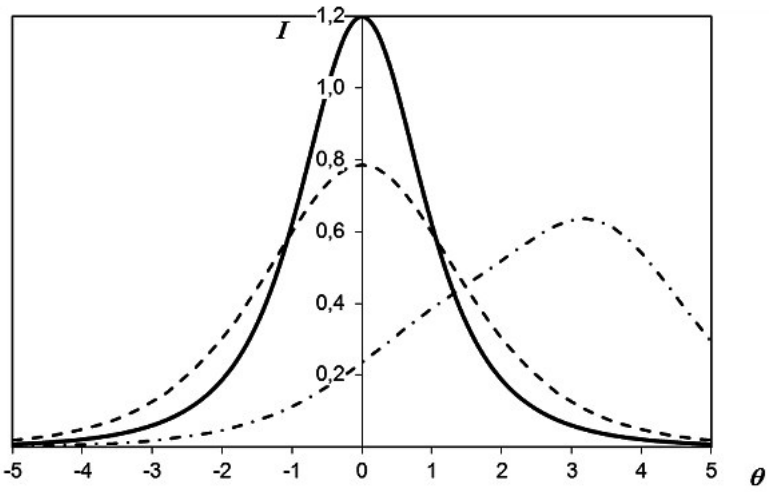

Рис. 6. Зависимость количества информации трёхбалльного задания РСМ от уровня подготовленности испытуемого: –––––––– — при $\beta_1 = -0{,}1$; $\beta_2 = 0$ и $\beta_3 = 0{,}1$; — – — – — при $\beta_1 = -1$, $\beta_2 = 0$ и $\beta_3 = 1$; — · — · — · — при $\beta_1 = 1$, $\beta_2 = 3$ и $\beta_3 = 4$

количества информации I_{max} равно 1,2;

- при $\beta_1 = -1$, $\beta_2 = 0$ и $\beta_3 = 1$ $I_{max} = 0{,}79$;
- при $\beta_1 = 1$, $\beta_2 = 3$ и $\beta_3 = 4$ $I_{max} = 0{,}64$.

Теоретически количество градаций степени правильности ответа не ограничено, но в практике педагогического тестирования это число обычно не превышает четырёх. Поэтому были проведены расчёты для четырёх балльного задания (рис. 7).

Для четырёхбалльного задания теоретический максимум близок к 2:

- при $\beta_1 = -0{,}2$; $\beta_2 = -0{,}1$; $\beta_3 = 0{,}1$ и $\beta_4 = 0{,}2$ $I_{max} = 1{,}8$;
- при $\beta_1 = -2$, $\beta_2 = -1$, $\beta_3 = 1$ и $\beta_4 = 2$ $I_{max} = 0{,}65$;
- при $\beta_1 = -1$, $\beta_2 = 1$, $\beta_3 = 3$ и $\beta_4 = 5$ $I_{max} = 0{,}5$.

Таким образом, увеличение числа градацией степени правильности ответа на тестовое задание в модели РСМ существенно повышает информационную ценность ответов по сравнению с базовой моделью Раша:

- максимум количества информации для двухбалльного задания больше максимума информационной функции модели Раша в 2,68 раза;
- для трёхбалльного задания — в 5 раз;

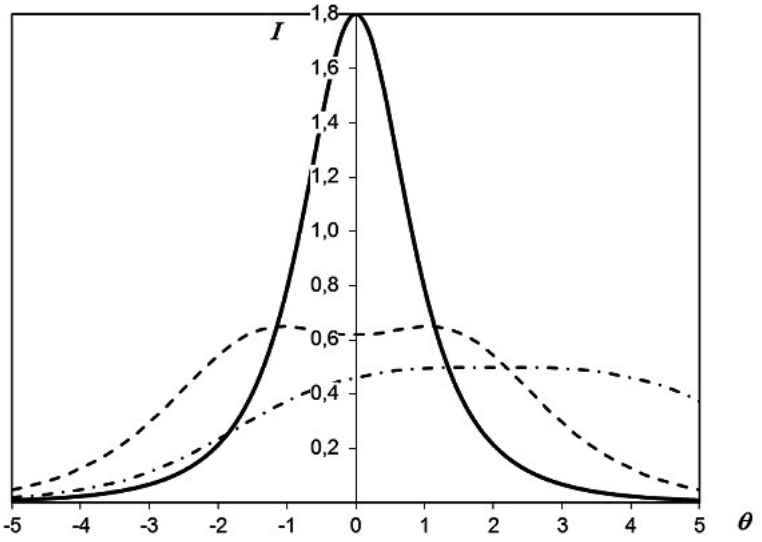

Рис. 7. Зависимость количества информации четырёхбалльного задания РСМ от уровня подготовленности испытуемого:
------ — при $\beta_1 = -0{,}2$; $\beta_2 = -0{,}1$; $\beta_3 = 0{,}1$ и $\beta_4 = 0{,}2$;
— – — – — при $\beta_1 = -2$, $\beta_2 = -1$, $\beta_3 = 1$ и $\beta_4 = 2$;
— · — · — · — при $\beta_1 = -1$, $\beta_2 = 1$, $\beta_3 = 3$ и $\beta_4 = 5$

• для четырёхбалльного задания — в 8 раз.

Необходимо сделать важное замечание: теоретический максимум количества информации может быть получен при почти совпадающих значениях уровня подготовленности испытуемого и уровней трудности шагов задания — различие не более 0,0001 логита. Например, для двухбалльного задания $\theta = 0$, $\beta_1 = -0{,}0001$ и $\beta_2 = 0{,}0001$. Уровни трудности шагов взаимно независимы, так как знание одного из этих уровней не даёт возможности однозначно определить остальные уровни. Поэтому мы вправе считать различие между уровнями β_1 и β_2 случайной величиной. Тогда вероятность практического совпадения значений уровня подготовленности испытуемого и уровней трудности шагов задания близка к нулю, т.е. теоретический максимум количества информации не может быть отправной точкой для оценки эффективности адаптивного тестирования. Но даже с учётом этого замечания модель PCM превосходит модель Раша, поскольку даже при значительном расхождении уровня подготовленности испытуемого и уровней трудности заданий количество информации больше максимума информационной функции модели Раша в 1,7–3,3 раза (см. табл. 1).

Таблица 1

	Количество информации				
	Модель Раша	Двухпараметрическая модель	Partial Credit Model		
			двухбалльное задание	трёхбалльное	четырёхбалльное
Максимальное значение	0,25	$0{,}25a^2$	0,67	1,25	2
Незначительное расхождение уровня подготовленности испытуемого и уровней трудности заданий (0,1 логита)	0,249	$\frac{a^2 e^{0,1a}}{\left(1+e^{0,1a}\right)^2}$	0,644	1,2	1,901
Значительное расхождение уровня подготовленности испытуемого и уровней трудности заданий (1 логит)	0,197	$\frac{a^2 e^{a}}{\left(1+e^{a}\right)^2}$	0,424	0,788	0,837

Конкурс моделей

Трёхпараметрическую модель можно исключить из конкурса моделей: как было показано выше, она уступает двухпараметрической модели по информационной ценности ответов. Для выявления наиболее подходящих моделей по критерию информативности тестирования сведём полученные данные в табл. 1.

Сначала сравним модель Раша и двухпараметрическую модель. Например, при дифференцирующей способности задания $a = 3$:

- максимум количества информации для двухпараметрической модели в 9 раз больше максимума для модели Раша $I_{max} = 0{,}25 \cdot 3^2 = 2{,}25$;
- при незначительном расхождении уровня подготовленности испытуемого и уровня трудности задания (0,1 логита) двухпараметрическая модель позволит получить информации в 8,8 раза больше, чем модель Раша;
- если уровни различаются на 1 логит, то количество информации для двухпараметрической модели больше в 2,1 раза.

Однако ситуация кардинально меняется при дифференцирующей способности задания менее единицы. Так, при дифференцирующей способности задания $a = 0{,}3$:

- максимум количества информации для двухпараметрической модели в 11 раз меньше максимума для модели Раша $I_{max} = 0{,}25 \cdot 0{,}3^2 = 0{,}0225$;
- при расхождении уровня подготовленности испытуемого и уровня трудности задания 0,1 логита информативность ответа по модели Раша в 11 раз больше, чем по двухпараметрической модели;
- если уровни различаются на 1 логит, то количество информации для двухпараметрической модели меньше в 9 раз.

Чтобы гарантировать эффективность двухпараметрической модели при адаптивном тестировании необходимы тестовые задания разных уровней трудности с высокой дифференцирующей способностью. Однако дифференцирующую способность заданий невозможно задать на этапе разработки теста — её можно определить только путём обработки результатов выполнения теста достаточно репрезентативной группой испытуемых.

Если гипотетически считать, что количество заданий теста не ограничено, то для каждого уровня подготовленности испытуемого можно выбрать достаточно заданий с высокой дифференцирующей способностью, что означает преимущество двухпараметрической модели. Но реальное распределение параметров заданий конкретного теста может оказаться таким, что преимущество получит модель Раша.

Влияние распределения параметров заданий можно наглядно проиллюстрировать результатами тестирования школьников Красноярского края по русскому языку и математике (данные предоставлены краевым государственным бюджетным специализированным учреждением «Центр оценки качества образования», г. Красноярск). Расчёты проведены с помощью компьютерной программы Estimate3PL (сайт www.asksystem.narod.ru). Большое количество испытуемых (свыше 22 тысяч) предопределило высокую точность оценки параметров тестовых заданий. Распределение параметров тестовых заданий и информационная функция теста приведены на рис. 8.

Тест по русскому языку состоит из 40 заданий низкого и среднего уровня трудности (параметры заданий обозначены маркерами на рис. 8*а*). Ввиду того, что, дифференцирующая способность тестовых заданий оказалась близка к единице (среднее значение $a_{ср}$ = 1,03), информационные функции теста, соответствующие модели Раша и двухпараметрической модели почти совпадают (сплошная и пунктирная линии на рис. 8*а*). В правой части графика обе линии близки к нулю, что означает низкую точность результатов тестирования хорошо подготовленных школьников.

Тест по математике тоже включает задания невысокого уровня трудности (всего 31 задание, рис. 8*б*). Однако дифференцирующая способность заданий оказалась выше, благодаря чему максимум информационной функции двухпараметрической модели почти в два раза превысил максимум информационной функции модели Раша. А справа и слева от максимума двухпараметрическая модель проигрывает (пунктирная линия ниже сплошной). Это означает, что участки диапазона измерений, на которых мало заданий и (или) задания имеют низкую дифференцирующую способность, являются для двухпараметрической модели весьма проблемными, так как информационная ценность ответов очень мала. Если уровень подготовленности испытуемого соответствует такому проблемному участку диапазона измерений, то двухпараметрическая модель получает значительно меньше информации, и точность измерения резко падает. У модели Раша точность снижается медленнее.

Partial Credit Model

Как показано выше, Partial Credit Model эффективнее модели Раша, так как в равных условиях позволяет получить больше информации о качестве измерения (см. табл. 1).

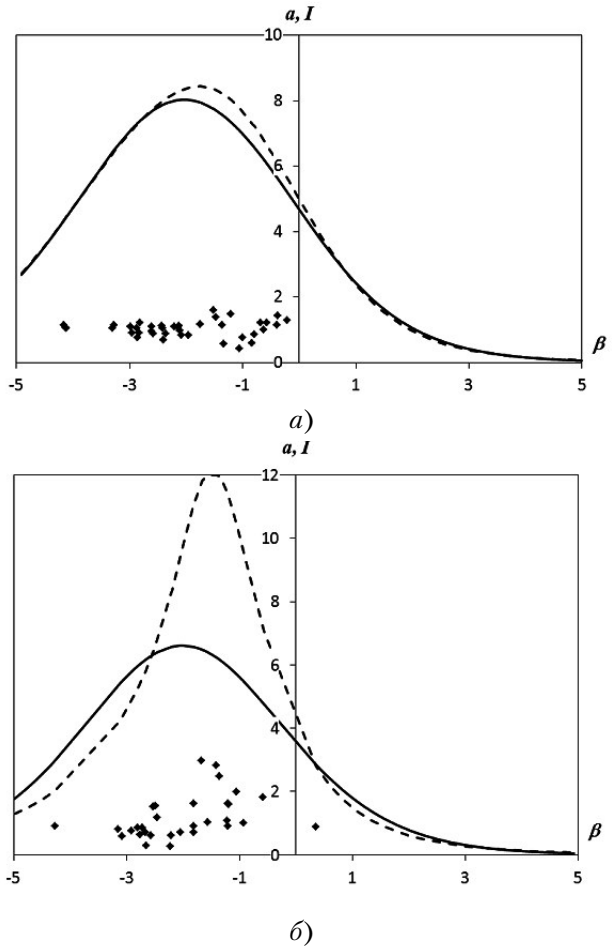

Рис. 8. Распределение параметров тестовых заданий и информационная функция тестов по русскому языку (а) и математике (б): • — уровень трудности β и дифференцирующая способность a тестового задания; ––––––––– — информационная функция теста по модели Раша; — – — – — – — информационная функция теста по двухпараметрической модели

Однако сопоставление РСМ и двухпараметрической модели не приводит к столь однозначному ответу (табл. 2):

• при наличии заданий нужного уровня трудности с высокой дифференцирующей способностью (a = 3) более предпочти-

тельной является двухпараметрическая модель. В этом случае и максимум количества информации, и количество информации при незначительном расхождении уровня подготовленности испытуемого и уровней трудности заданий двухпараметрической модели превышает аналогичные показатели модели PCM (выделено в табл.2 жирным шрифтом);

- при низкой дифференцирующей способности заданий (a = 0,3) двухпараметрическая модель безоговорочно проигрывает: получаемое количество информации в десятки раз ниже, чем у модели PCM;
- если уровень трудности задания существенно отличается от уровня подготовленности испытуемого (т.е. нет подходящих по уровню трудности заданий), то двухпараметрическая модель также проигрывает. Даже высокая дифференцирующая способность заданий в этом случае не помогает (нижняя строка табл. 2).

Выводы

1. Тип математической модели оказывает существенное влияние на эффективность адаптивного тестирования.

2. Наиболее эффективной в условиях адаптивного тестирования является двухпараметрическая модель, при условии достаточного количества заданий с высокой дифференцирующей

Таблица 2

	Количество информации				
	Двухпараметрическая модель		Partial Credit Model		
	a = 0,3	a = 3	Двухбалльное задание	трёхбалльное	четырёхбалльное
Максимальное значение	0,0225	**2,25**	0,67	1,25	2
Незначительное расхождение уровня подготовленности испытуемого и уровней трудности заданий (0,1 логита)	0,0225	**2,20**	0,644	1,2	1,901
Значительное расхождение уровня подготовленности испытуемого и уровней трудности заданий (1 логит)	0,0220	0,41	0,424	0,788	0,837

способностью, равномерно распределённых по всему диапазону измерения уровня подготовленности испытуемых.

3. Модель РСМ и модель Раша являются предпочтительными, если указанное выше условие не выполняется, т.е.:

- при малом количестве тестовых заданий;
- при наличии участков диапазона измерений, на которых мало заданий и (или) задания имеют низкую дифференцирующую способность.

4. По эффективности адаптивного тестирования трёхпараметрическая модель уступает двухпараметрической, а модель Раша — модели РСМ.