

Математическая модель шкалирования результатов

Методика шкалирования результатов ЕГЭ в 2010 году преимущественна с соответствующей методикой прошлых лет. Важным фактором в принятии решения при выборе методики 2010 года является необходимость обеспечения сопоставимости результатов ЕГЭ 2009 и 2010 годов.

В 2008 году была утверждена методика, основные элементы которой используются в представленном ниже описании. В качестве основной модели при обработке результатов ЕГЭ несколько последних лет используется вероятностная модель Partial Credit, принадлежащая к семейству моделей Теории Моделирования и Параметризации Тестов (ТМПТ, в англоязычной литературе – Item Response Theory, IRT). Данная модель позволяет обрабатывать результаты тестирования по вариантам, составленным из заданий, оцениваемых разным числом баллов.

Выбор какой-либо из моделей ТМПТ означает установление зависимости оценки выполнения каждой из ступеней задания от измерительных параметров задания и параметров подготовленности испытуемого.

Главное свойство ТМПТ-моделей состоит в том, что они дают оценки не ожидаемых результатов испытуемого в конкретном тесте, а параметра подготовленности испытуемого. Благодаря этому ТМПТ-модели позволяют получить близкие (с точностью до ошибки измерения) оценки уровней подготовленности одинаково подготовленных участников.

Оценки параметров подготовленности участников тестирования располагаются на интервальной шкале. Выбор модели Partial Credit обусловлен двумя причинами.

Во-первых, в рамках этой модели оценки уровней подготовленности испытуемых обладают наименьшей ожидаемой ошибкой для подготовленности основной массы участников.

Во-вторых, в данной вероятностной модели оценка параметра зависит только от первичного балла участника. Это приводит к тому, что при выполнении одного варианта оценка уровня подготовленности участника тем больше, чем больше данный участник набрал первичных баллов. Процедура вычисления тестового балла состоит из четырех основных этапов. Первый этап заключается в независимой обработке результатов тестирования по каждому из вариантов теста. Под обработкой понимается вычисление оценок уровня подготовленности участников тестирования и оценок измерительных параметров категорий тестовых заданий всех вариантов КИМ.

В настоящее время для получения этих оценок используется статистический метод условного максимального правдоподобия. Данный метод позволяет реализовать раздельную оценку параметров подготовленности участников тестирования и измерительных параметров тестовых заданий. После проведения первого этапа оценки показателей подготовленности участников тестирования располагаются на разных шкалах. То есть, для каждого варианта КИМ образуется своя шкала и корректно сравнивать между собой уровни подготовленности участников тестирования можно только если полученные результаты приведены к единой для всех участников тестирования шкале. Именно в этом и заключается второй этап процедуры вычисления тестового балла.

При обработке результатов ЕГЭ сразу после второго этапа производится усреднение оценок параметров подготовленности участников тестирования, набравших одинаковый первичный балл. Полученная усредненная оценка приписывается всем этим участникам ЕГЭ.

Эта процедура устанавливает однозначное монотонное соответствие первичных и тестовых баллов. То есть при одинаковых первичных баллах участники тестирования получают одинаковые тестовые баллы вне зависимости от того, какой из вариантов КИМ они выполняли, и большему первичному баллу соответствует больший тестовый балл. На

третьем этапе в 2010 году при проведении шкалирования оценки параметров подготовленности подвергаются сдвигу, объединяющему шкалы 2009 г. и 2010 г.

Четвертый этап состоит в кусочно-линейном преобразовании оценок уровней подготовленности участников тестирования из исходной шкалы в стобалльную шкалу. Формула перевода шкалы логитов в шкалу тестовых баллов в 2010 году выглядит так же, как и в 2008 и 2009 году:

$$t = \begin{cases} 0, & \text{если } \theta < \theta_{\min}, \\ \text{округл} \left(\frac{6\theta_{\max} + 88\theta - 94\theta_{\min}}{\theta_{\max} - \theta_{\min}} \right), & \text{если } \theta_{\min} \leq \theta \leq \theta_{\max}, \\ 100, & \text{если } \theta > \theta_{\max} \end{cases}$$

где t – тестовый балл, θ – оценка уровня подготовленности участника ЕГЭ в логитах, θ_{\min} – оценка в логитах, соответствующая одному первичному баллу в 2009 году, θ_{\max} – оценка в логитах, соответствующая первичному баллу, на единицу меньшему максимально возможному в 2009 году. Значения уровня подготовленности, полученные в 2010 году и лежащие вне интервала $(\theta_{\min}, \theta_{\max})$, линейно отображаются на интервалы (0, 6) и (94, 100). Округление проводится до целого числа. Нулевому первичному баллу ставится в соответствие тестовый балл 0, а максимально возможному первичному баллу ставится в соответствие тестовый балл 100.

Описанная выше методика используется для всех предметов, по которым проводится ЕГЭ, кроме иностранных языков и математики, в связи со значительными изменениями в структуре КИМ по этим предметам.

Методика шкалирования результатов ЕГЭ по иностранным языкам в 2010 году

Для иностранных языков рекомендуется определение тестового балла на основе вычисления доли выполненных заданий КИМ. Вычисляется процент выполнения заданий варианта который и является тестовым баллом. После установления соответствия производится округление тестовых баллов до ближайшего целого.

Методика шкалирования результатов ЕГЭ по математике в 2010 году

Предлагаемая методика шкалирования результатов ЕГЭ по математике основывается на вычислении процентильных рейтингов участников экзамена 2009 и 2010 года (под процентильным рейтингом участника ЕГЭ понимается процентная доля участников, результат которых не лучше результата данного участника) и отождествлении тестовых баллов участников двух лет, имеющих одинаковое положение на порядковой процентильной шкале.

Методика включает два пункта.

1. Присвоение каждому участнику ЕГЭ по математике 2010 года процентильного рейтинга. С целью исключения инверсий балла в качестве базы расчета процентильного рейтинга предлагается использовать первичный балл.
2. Определение тестового балла участника ЕГЭ 2010 г. на шкале тестовых баллов в 2009 г. Если первичный балл участника равен 0, участнику присваивается тестовый балл 0. Если первичный балл участника равен 30 (максимальное возможное значение), участнику присваивается тестовый балл 100.

Если первичный балл участника принял значение от 1 до 29, то среди участников ЕГЭ по математике 2009 г. выбираются двое, имеющие наиболее близкие к рассматриваемому участнику процентильные рейтинги (у одного – не больше, у другого – не меньше), и затем тестовый балл рассматриваемого участника 2010 г. вычисляется усреднением тестовых баллов двух участников 2009 г. с помощью линейного преобразования.

Описанная методика исключает инверсии баллов 2010 года по сравнению с первичным, а также перерасчет баллов участников ЕГЭ 2009 года.

Литература:

1. В.А. Дюк Компьютерная психодиагностика, С-Пб., 1994
2. М.Б. Чельшкова. Теория и практика конструирования педагогических тестов. Москва, Логос, 2002.
3. Ю.Н. Тюрин, А.А. Макаров. Анализ данных на компьютере, М., Инфра-М, 2003.
4. Кендалл М. Методы ранговой корреляции.— М.: Статистика, 1974.
5. Ю.М. Нейман, В.А. Хлебников. Введение в теорию моделирования и параметризации педагогических тестов - Москва, 2000.