

2007 г.

## Методика шкалирования результатов единого государственного экзамена

### 1. Введение

1.1. Оценка уровней подготовленности участников единого государственного экзамена производится в результате обработки ответов на задания контрольных измерительных материалов (КИМ) с использованием современных научно обоснованных методов.

1.2. Сбор первичной информации об ответах испытуемых на задания контрольных измерительных материалов (КИМ) ЕГЭ осуществляется в соответствии с требованиями утвержденной инструкции по проведению ЕГЭ в 2007 г.

1.3. КИМ ЕГЭ содержат задания типов **А, В, С**.

Задания типа **А** (задания с выбором ответа) предполагают выбор одного ответа (номера ответа) из группы предложенных вариантов ответов;

Задания типа **В** (задания с кратким ответом) предполагают ответ в виде набора символов (цифр или букв);

Задания типа **С** (задания с развернутым ответом) предполагают произвольную запись развернутого ответа на задание.

1.4. Оценка уровней подготовленности участников ЕГЭ производится отдельно по каждому предмету.

1.5. По каждому предмету используется некоторое множество вариантов КИМ одинаковой структуры, содержащие близкие по сложности задания с одинаковым номером в различных вариантах КИМ.

1.6. Ввод в ПЭВМ сведений об ответах участников ЕГЭ на задания КИМ с бланков ответов производится автоматизировано.

### 2. Определения

2.1. **Первичным баллом** участника ЕГЭ за задание по варианту КИМ ЕГЭ называется балл, полученный конкретным участником ЕГЭ за

решение соответствующего задания. Первичные баллы за задания типа **A** и **B** выставляются на основе компьютерной проверки правильности ответа, первичные баллы за задания типа **C** определяются экспертами в соответствии инструкциям по проверке заданий с развернутым ответом. Если участник ЕГЭ не приступал к выполнению задания, то его первичный балл за это задание будет равен нулю.

2.2. **Первичным баллом участника ЕГЭ** по варианту КИМ ЕГЭ называется сумма первичных баллов участника ЕГЭ, полученных им за все задания полученного варианта КИМ.

2.3. **Шкалированием результатов ЕГЭ** называется процедура, в результате действия которой каждому из участников ЕГЭ на основе его первичного балла по варианту предмета приписывается определенное целое число, характеризующее его уровень подготовленности по предмету и находящееся в диапазоне от 0 до 100 включительно.

2.4. В качестве основы процедуры шкалирования результатов ЕГЭ принята методика, которая в зарубежной литературе обозначается как модель Partial Credit.

2.5. Модель Partial Credit позволяет работать с заданиями политомического типа. В рамках этой модели можно получать объективные оценки уровней подготовленности испытуемых и параметров заданий (подробно об этом можно узнать из работ [1], [3]).

2.6. **Тестовым баллом** участника ЕГЭ называется число, которое ему присваивается в результате шкалирования.

### **3. Процедура шкалирования**

3.1. После проведения ЕГЭ по конкретному предмету и ввода в ПЭВМ ответов участников ЕГЭ на задания КИМ формируется массив (матрица) первичных баллов по каждому варианту КИМ.

Варианты КИМ ЕГЭ обрабатываются независимо друг от друга.

3.2. Для каждого варианта КИМ ЕГЭ в рамках модели Partial Credit вычисляются оценки уровней подготовленности участников ЕГЭ и оценки параметров заданий.

3.3. С использованием метода условного наибольшего правдоподобия производится вычисление оценок уровней подготовленности испытуемых и параметров заданий. Указанный метод, в отличие от большинства остальных, позволяет реализовать раздельное независимое оценивание уровней подготовленности испытуемых и параметров заданий без каких-либо дополнительных предположений. При использовании указанного метода в качестве оценок параметров заданий берутся числа, доставляющие максимум некой функции, называемой функцией правдоподобия. Оценки уровней подготовленности испытуемых вычисляются аналогично. Более детально с данным методом можно ознакомиться, например, в [2].

3.4. Оценки уровней подготовленности участников ЕГЭ и оценки параметров заданий наносятся на единую, общую для всех вариантов КИМ ЕГЭ шкалу логитов с помощью общих (узловых) заданий. Для этого выбирается один из вариантов КИМ (базовый), на шкалу которого будут нанесены результаты по всем остальным вариантам. Для каждого из оставшихся вариантов КИМ ищется оценка сдвига относительно базового варианта путем минимизации различия характеристических функций общих для этих двух вариантов КИМ заданий. После того как оценка сдвига была определена, на эту величину сдвигаются оценки уровней подготовленности испытуемых и параметров заданий каждого рассматриваемого варианта КИМ.

3.5. Оценка уровня подготовленности испытуемых с конкретным первичным баллом рассчитывается как средняя величина по всем вариантам КИМ предмета.

3.7. Оценки уровней подготовленности участников ЕГЭ с помощью линейного преобразования переносятся с непрерывной неограниченной шкалы логитов на дискретную стобалльную шкалу тестовых баллов.

3.8. Вычисление тестового балла производится следующим образом. Испытуемым, набравшим максимально возможный первичный балл, ставится в соответствие 100 тестовых баллов. Испытуемым, набравшим ноль первичных баллов, ставится ноль тестовых баллов.

Тестовые баллы остальных испытуемых вычисляются по формуле:

$$t = 50 + \gamma \frac{\theta - \bar{\theta}}{\sigma},$$

$t$  – тестовый балл испытуемого,  $\theta$  – оценка уровня подготовленности испытуемого в логитах,  $\bar{\theta}$  – среднее значение оценок уровней подготовленности испытуемых в логитах,  $\sigma$  – стандартное отклонение оценок уровней подготовленности в логитах,  $\gamma$  – параметр, устанавливаемый индивидуально для каждого предмета так, чтобы тестовые баллы испытуемых находились в интервале от 0 до 100.

#### Список литературы

1. *Нейман Ю.М., Хлебников В.А.* Как оценивается уровень подготовленности учащихся по результатам единого государственного экзамена. – Москва.: Центр тестирования Минобразования России. 2004. – 44 с.
2. *Frank B. Baker*, Item response theory: parameter estimation techniques. - New York.: Marcel Dekker. 1992. – 496 p.
3. *Wright B.D., Masters G.N.* Rating scale analysis. - Chicago.: MESA PRESS. 1982. – 204 p.