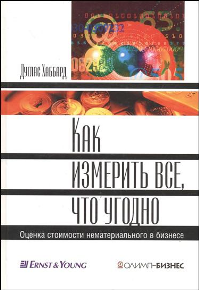
**Использование модели Раша в пересчете баллов ЕГЭ**

С моделями Раша я впервые столкнулся когда читал книгу Дугласа Хаббарда [«Как измерить всё, что угодно»](http://baguzin.ru/wp/?p=2511). Как нередко бывает, новый метод не произвел на меня особого впечатления. К тому же в книге описывалось, что основное применение метод находит в педагогике – области, не очень мне близкой. Однако, в связи с тем, что в этом году мой сын сдает ЕГЭ, я кое-что почитал на эту тему в Интернете. Каково же было мое изумление, когда оказалось, что в основе формирования баллов ЕГЭ лежит [политомическая модель Раша](http://en.wikipedia.org/wiki/Polytomous_Rasch_model). Тема меня заинтересовала, и я решил изучить метод подробнее. Начну с фрагмента упомянутой книги Дугласа Хаббарда



**Как стандартизировать любую оценку: модели Раша**

Выбирая из огромного разнообразия имеющихся статистических методов те, о которых следовало бы рассказать в этой книге, я старался не ограничиваться уже ранее знакомыми мне. В частности, для меня новинкой стали приемы дидактического тестирования (educational testing), практически неизвестные специалистам по измерению, работающим в других областях. Именно таким приемам посвящена книга с многозначительным названием «Objective Measurement («Объективное измерение»). Подобная публикация могла быть всесторонним исследованием проблем измерения, одинаково интересным для астронома, инженера-химика и программиста, но в ней говорится лишь об оценке способностей человека и дидактическом тестировании. Это все равно, как если бы вы увидели старую карту с названием «Карта мира», на которой на самом деле изображен богом забытый островок в Тихом океане и которая была составлена людьми, не знавшими, что населяют лишь крошечную часть огромной планеты. Один специалист по дидактическому тестированию как-то рассказал мне об «инвариантном сравнении» — характерной черте измерения, которая, по его словам, «настолько фундаментальна, что ее можно считать на 10% основой статистики». Другой эксперт в этой же области уверял, что это суть физики. Из всех физиков и статистиков, которых я позднее спрашивал об «инвариантном сравнении», что-то слышал о нем только один. По всей видимости, то, что специалисты по дидактическому тестированию считают общей для всех фундаментальной основой, на самом деле имеет значение только для них. Справедливости ради отмечу, что, наверное, кто-то скажет то же самое и о книге, претендующей научить читателя измерять все, что угодно.

Но, вообще говоря, у специалистов по дидактическому тестированию есть чему поучиться. Ведь им приходится решать все проблемы по оценке возможностей человека — большую группу задач измерения того, что многие компании считают неизмеряемым. Концепция инвариантного сравнения имеет отношение к ключевой проблеме проведения многих тестов на определение способностей человека, например теста на IQ. Принцип инвариантного сравнения гласит, что если один инструмент измерения говорит, что А больше, чем В, то и другой инструмент должен продемонстрировать то же самое. Иными словами, результаты сравнения А и В должны быть одинаковыми, какие бы инструменты измерения ни использовались. Для физика это очевидно настолько, что даже не заслуживает упоминания. Казалось бы, если на одних весах А весит больше, чем В, то и другие весы должны показывать то же самое, даже если первые весы пружинные, а вторые — цифровые. Аналогичные результаты должны наблюдаться с IQ-тестами или любыми другими тестами, оценивающими человеческую эффективность. Однако при проведении таких испытаний может произойти нечто совсем иное: результаты теста на IQ с одним набором вопросов могут сильно отличаться в выводах от результатов аналогичного теста с другим набором вопросов. Поэтому один тест может показать, что Боб умнее Шерри, а второй — что Шерри умнее Боба.

Другое проявление той же проблемы наблюдается, когда разным экспертам приходится оценивать способности больших коллективов. Если аттестуемых слишком много, то их делят на группы, и каждый человек имеет дело с разным составом судей. Возможно, один эксперт оценивает одно качество одного аттестуемого, а другой обращает внимание совсем на другое или же разным людям даются задачи разной степени сложности. Предположим, например, что вы хотите оценить квалификацию менеджеров проекта по тому, как они справятся с порученными им программами. Если таких менеджеров несколько, то вам потребуется не один судья. Экспертами могут стать непосредственные начальники аттестуемых менеджеров как люди, наиболее знакомые с их работой. Отметим, что проекты могут различаться по сложности. Предположим далее, что все менеджеры, независимо от того, какие проекты им поручены и кому они подчиняются, должны конкурировать за один фонд бонусов или за одни и те же карьерные места. Те, кого будет оценивать строгий судья или кому поручены более сложные проекты, окажутся в невыгодном положении по сравнению с другими. Сравнение разных менеджеров не будет инвариантным (то есть независимым от того, кто его проводит и по каким критериям). На самом деле, доминирующими решающими факторами, определяющими рейтинг менеджеров, могут оказаться условия, не контролируемые ими.

Решение этой проблемы в 1961 г. предложил датский статистик Георг Раш. Он разработал метод предсказания вероятности того, что испытуемый правильно ответит на вопрос бинарного типа «верно/неверно» на основе 1) процента в генеральной совокупности других респондентов, ответивших на этот вопрос правильно, и 2) процента других вопросов, на которые данный испытуемый уже ответил правильно. Выполнение тестов человеком, не знакомым с предлагаемым набором вопросов, можно предсказать с погрешностью, поддающейся вычислению, даже если давать испытуемым разные тесты.

Во-первых, Раш рассчитал вероятность того, что случайно выбранный из группы человек ответит на вопрос правильно. Она просто равна доле тех, кто дал верный ответ, в общем числе ответивших на данный вопрос. Этот показатель называется у Раша уровнем трудности задания (item difficulty). Затем Раш рассчитал логарифм отношения вероятностей правильного и неправильного ответов. Если трудность задания была 65%, то это означает, что 35% респондентов ответили правильно, а 65% — неправильно. Отношение вероятности ответить правильно к вероятности ответить неправильно — 0,55, а натуральный логарифм 0,55 составляет -0,62. При желании можно записать это в виде формулы в Excel: =ln(А1/(1 – А1)), где А1 — ячейка со значением вероятности правильного ответа (рис. 1).

Формула2.bmp

Рис. 1. Вид функции , где

Затем Раш проделал ту же процедуру с вероятностью, что этот человек ответит правильно на любой вопрос. Поскольку данный респондент давал правильные ответы в 82% случаев, соответствующий логарифм составил ln(0,82/0,18), или 1,52. Наконец, Раш сложил значения двух логарифмов и получил: (-0,62) + 1,52 = 0,9. Чтобы снова преобразовать это в вероятность, можно записать следующую формулу в Excel: = 1/(1/ехр(0,9) + 1).

В результате получится 71%. Это означает, что есть 71%-ная вероятность того, что данное лицо ответит на этот вопрос верно, учитывая трудность задания и правильность его ответов на другие вопросы. При большом числе вопросов и (или) большом числе испытуемых мы обнаружим следующее: когда уровень трудности задания 70% (то есть, вероятность получить правильный ответ 30%), около 30% людей ответят на этот вопрос верно; когда уровень трудности задания 80%, около 20% людей ответят на данный вопрос верно, и т.д. Таким образом, модели Раша — просто еще один способ калибровки вероятностей.

Мэри Лунц из чикагской компании Measurement Research Associates Inc. применила модели Раша к решению важной задачи в области общественного здравоохранения, порученной ей Американским обществом клинической патологии (American Society of Clinical Pathology). Использовавшийся этим обществом ранее порядок сертификации патологов давал большую погрешность, которую необходимо было уменьшить. Каждый кандидат должен был разобраться в одном или двух случаях, и каждый его ответ оценивался одним или несколькими экспертами. Практически невозможно одному эксперту оценить все задания, как невозможно гарантировать, что все они будут одинаковой сложности. Раньше получение кандидатом сертификата почти целиком зависело от того, какой экзаменатор ему попадется, и какой случай придется разбирать на экзамене. Иными словами, снисходительные экзаменаторы могли пропустить некомпетентных кандидатов. Лунц рассчитала стандартные очки Раша для каждого эксперта, экзаменационного задания, а также кандидата для всех уровней квалификации.

В результате появилась возможность предсказать, пройдет ли кандидат экзамен у среднего экзаменатора при случае средней сложности, или у снисходительного эксперта при легком случае, или, наоборот, у строгого экзаменатора при очень сложном случае. И теперь (наверняка не слишком скоро для самих кандидатов) наконец появилась возможность полностью устранить при сертификации отклонения, связанные с характером экзаменатора или сложностью задания.

Интересное направление применения статистики Раша — **оценка сложности прочтения того или иного текста**. Доктор Джек Стеннер, президент и основатель компании MetaMetrics, Inc., использовал модели Раша для разработки схемы «Lexile»— способа оценки умения читать и писать, а также анализа сложности текста для восприятия. Система «Lexile» позволяет оценивать навыки чтения, письма, выполненные тесты, тексты и способности студентов, впервые сделав возможным сравнение наиболее распространенных языков. Располагая персоналом всего в 56 человек, компания MetaMetrics добилась в этой сфере гораздо больше, чем любая другая государственная или частная организация. Так:

* баллы «Lexile» используются во всех основных тестах на умение читать. Свои показатели «Lexile» знают около 20 млн. американских учащихся;
* система «Lexile» использована для классификации по сложности восприятия около 100 тыс. книг и десятков миллионов журнальных статей;
* программы обучения чтению по учебникам некоторых авторов основаны на системе «Lexile»;
* на систему «Lexile» переходят все новые учебные заведения штатов и местные учебные заведения.

Показатель 100 баллов по системе «Lexile» означает текст первого уровня сложности, а 1700 — это уровень сложности текста решений Верховного суда, научных журналов и т. п. Компания MetaMetrics считает, что читатель, набравший 600 баллов, сумеет на 75% понять содержание текста, уровень сложности которого составляет также 600 баллов.

**Теоретические основы использования модели Раша при тестировании**[[1]](#footnote-1)

Классическая теория тестирования, несмотря на хорошо разработанный математический аппарат, прозрачность и ясность получаемых выводов, имеет принципиальные недостатки. В частности, тестовые баллы испытуемых зависят от трудности заданий в тесте, а трудность задания зависит от выборки испытуемых. Большим недостатком классической теории является нелинейность тестовых баллов испытуемых. За рубежом уже несколько десятилетий развивается современная теория тестирования – Item Response Theory (IRT), являющаяся частью более общей теории латентно-структурного анализа. Отдельно следует указать теорию Георга Раша (G. Rasch), которую иногда называют однопараметрической (теорией) IRT.

Перечислим преимущества IRT перед классической теорией тестов:

* использование модели Раша превращает измерения, выполненные в порядковых шкалах, в линейные измерения, в результате качественные данные анализируются с помощью количественных методов;
* мера измерения параметров модели Раша является линейной, что позволяет использовать широкий спектр статистических процедур для анализа результатов измерений;
* оценка трудности тестовых заданий не зависит от выборки испытуемых, на которых она была получена;
* оценка уровня подготовленности испытуемых не зависит от используемого набора тестовых заданий;
* неполнота данных (пропуск некоторых комбинаций испытуемый – тестовое задание) не является критичным.

ЛАТЕНТНЫЙ ПАРАМЕТР – это свойство личности, недоступное для прямого наблюдения. Латентными параметрами являются, например, чувство патриотизма, толерантность, уровень знаний, и т. п. О величине латентного параметра можно судить по ее индикатору (индикаторной переменной). Главное достоинство индикатора – его доступность для прямого наблюдения. Измеряя значение индикатора, мы можем судить о значении латентного параметра, с которым он связан. Например, индикатором может являться тестовое задание. Значением индикатора является числовое (символьное) выражение реакции испытуемого, на это тестовое задание. По этому индикатору мы можем судить об уровне знаний, соответствующих данному тестовому заданию.

ИНДИКАТОР – это некоторое средство воздействия (вопрос, тестовое задание), связанное с определенным латентным параметром, реакция на который, доступна для непосредственного наблюдения. Допустим, нас интересует латентный параметр «Уровень знаний по физике». Для этого мы создаем КОНСТРУКТ – систему индикаторов, позволяющих оценить латентный параметр. В нашем примере конструктом является тест по физике, а индикаторами – тестовые задания.

ОСНОВНЫЕ ДОПУЩЕНИЯ IRT:

1. существуют латентные (скрытые) параметры личности, недоступные для непосредственного наблюдения. В тестировании это уровень подготовленность испытуемого и уровень трудности задания;
2. существуют индикаторные переменные, связанные с латентными параметрами, доступные для непосредственного наблюдения. По значениям индикаторных переменных можно судить о значениях латентных параметров;
3. оцениваемый латентный параметр должен быть одномерным. Это означает, что, например тест, должен измерять знания только в одной, четко заданной, предметной области. Если условие одномерности не выполняется, то необходимо переработать тест, удалив задания, нарушающие его гомогенность.

ОСНОВНОЙ ЗАДАЧЕЙ IRT является переход от индикаторных переменных к латентным параметрам. В IRT устанавливается связь между двумя множествами значений латентных параметров. Первое множество составляют значения латентного параметра, определяющего уровень подготовленности испытуемых θi, где i - номер испытуемого, изменяющийся в интервале от 1 до N (N – количество испытуемых). Второе множество составляют значения латентного параметра, характеризующего трудность j-го задания βj. Индекс j меняется в пределах от 1 до M, где M – количество заданий в тесте.

Георг Раш предположил, что уровень подготовленности испытуемого θi и уровень трудности задания βj размещены на одной шкале и измеряются в одних и тех же единицах – логитах. Аргументом функции успеха испытуемого является разность θi – βj. Если эта разность положительна и велика, то соответственно высока вероятность достижения успеха i-го испытуемого в j-м задании. Если же эта разность отрицательна и велика по модулю, то вероятность достижения успеха i-го испытуемого в j-м задании будет низкой.

В качестве математической модели, связывающей успех испытуемого с уровнем его подготовленности и трудностью задания в модели Раша выбирается функция:

*(1) Pj(θ) =*

Поскольку модель Раша описывает вероятность успеха испытуемого как функцию одного параметра (θi – βj), то иногда ее называют однопараметрической моделью IRT. Взаимодействие двух множеств θi и βj образует данные, обладающие свойством «совместной аддитивности» (conjoint additivity). Правильное использование модели Раша позволяет отделить оценки испытуемых от оценок трудности заданий и наоборот. Это свойство Rasch Measurement носит название separability parameter estimates – «независимость оценок заданий от испытуемых и оценок испытуемых от параметров заданий». На рис.2 показаны три характеристические кривые согласно уравнению (1) с трудностями заданий -2, 0 и +2 логита (первое самое легкое, второе – среднее, третье самое трудное). Из приведенных зависимостей видно, что чем выше уровень подготовленности θ испытуемого, тем выше вероятность успеха в том или ином задании. Например, для испытуемого с θ=0 вероятность правильно ответить на первое задание близка к единице, на второе равна 0,5 и на третье почти равна нулю. Отметим, что в точках, где θ = β вероятность правильного ответа равна 0,5. То есть, если трудность задания равна уровню подготовленности испытуемого, то он с равной вероятностью может справиться или не справиться с этим заданием.

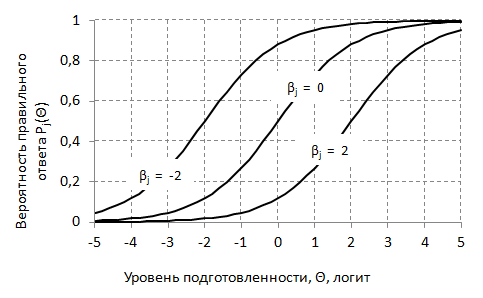


Рис. 2. Характеристические кривые заданий в модели Раша

**Методика шкалирования результатов ЕГЭ в 2012 году[[2]](#footnote-2)**

Методика шкалирования результатов ЕГЭ, используемая в 2012 г., соответствует методике 2011 года. При этом реализуется поэтапное установление соответствия тестовых и первичных баллов для каждого общеобразовательного предмета, по которому проводится ЕГЭ.

**I этап.** Сначала в диапазоне первичных баллов от нуля до максимального первичного балла ПБmax для каждого общеобразовательного предмета ЕГЭ выбираются два значения первичных баллов: ПБ1 и ПБ2, разделяющие группы участников с различным уровнем подготовки по данному предмету.

Величина ПБ1 выбирается как наименьший первичный балл, получение которого свидетельствует об усвоении участником экзамена основных понятий и методов по соответствующему общеобразовательному предмету. Он определяется на основе экспертизы демонстрационного варианта по данному общеобразовательному предмету специалистами общего образования, ссузов и вузов различного профиля из разных субъектов РФ. Экспертиза осуществляется с учетом уровня сложности каждого задания и значимости проверяемого им содержания, умения, навыка, способа деятельности в контексте общеобразовательного предмета. При этом требования к значению ПБ1 соответствуют требованиям, которые использовались при определении ПБ1 прошлого года (для обеспечения эквивалентности шкал двух лет). Величина ПБ2 определяется профессиональным сообществом как наименьший первичный балл, получение которого свидетельствует о высоком уровне подготовки участника экзамена, а именно, о наличии системных знаний, овладении комплексными умениями, способности выполнять творческие задания по соответствующему общеобразовательному предмету.

Если спецификация экзаменационного варианта не изменилась по сравнению с прошлым годом, то ПБ1 и ПБ2 также остаются неизменными. Если же структура экзаменационной работы или сложность заданий контрольных измерительных материалов поменялись, то устанавливаются новые значения ПБ1 и ПБ2 с учетом имеющихся изменений.

**II этап.** Первичным баллам ПБ1 и ПБ2 ставятся в соответствие тестовые баллы ТБ1 и ТБ2 по каждому общеобразовательному предмету. Для математики и русского языка (обязательных для сдачи ЕГЭ предметов) в качестве величин ТБ1 выбираются минимальные тестовые баллы ЕГЭ 2012 г., установленные распоряжениями Рособрнадзора: математика – 24, русский язык – 36. Данные значения совпадают с минимальными баллами ЕГЭ 2011 г. по математике и русскому языку соответственно. Тестовые баллы ТБ2 для русского языка и математики, а также тестовые баллы ТБ1 и ТБ2 для остальных общеобразовательных предметов (предметов по выбору) устанавливаются равными аналогичным баллам 2011 года. В таблице 1 представлены значения ПБ1 и ПБ2, ТБ1 и ТБ2 на 2012 год.

Необходимо подчеркнуть, что минимальные тестовые баллы, свидетельствующие об освоении школьного курса по предметам по выбору, устанавливаются Рособрнадзором после проведения основного экзамена и получения статистических данных по использовавшимся вариантам ЕГЭ 2012 года и по некоторым предметам могут не совпадать с ТБ1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Предмет** | **ПБ1** | **ТБ1** | **ПБ2** | **ТБ2** |
| Русский язык | 17 | 36 | 54 | 73 |
| Математика | 5 | 24 | 15 | 63 |
| Обществознание | 15 | 39 | 48 | 72 |
| История | 13 | 32 | 46 | 72 |
| Физика | 12 | 39 | 33 | 62 |
| Химия | 14 | 36 | 58 | 80 |
| Биология | 17 | 36 | 60 | 79 |
| География | 14 | 37 | 44 | 69 |
| Информатика | 8 | 40 | 35 | 84 |
| Иностранные языки | 16 | 20 | 65 | 82 |
| Литература | 8 | 32 | 36 | 73 |

**III этап.** По каждому общеобразовательному предмету определяется соответствие между первичным баллом и тестовым баллом на основе следующей процедуры. Первичному баллу 0 ставится в соответствие тестовый балл 0, а максимальному первичному баллу ПБmax ставится в соответствие тестовый балл 100. Все промежуточные первичные баллы между 0, ПБ1, ПБ2 и ПБmax переводятся в тестовые, пропорционально распределенные между соответствующими значениями тестовых баллов: 0, ТБ1, ТБ2 и 100. На рис. 3 представлена получаемая зависимость.

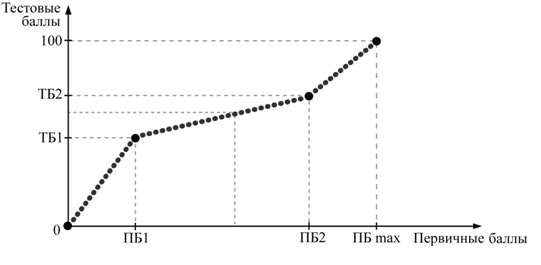


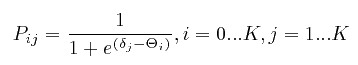
Рис. 3. Соответствие между тестовыми и первичными баллами

Если промежуточные первичные баллы соответствуют дробным значениям тестовых, то производится округление тестового балла до ближайшего большего целого числа.

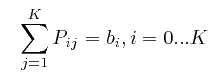
Указанная процедура позволяет согласовывать тестовые баллы одинаково подготовленных участников 2011 и 2012 гг. и обеспечивает сравнительную сопоставимость результатов экзамена по годам.

**Шкала результатов ЕГЭ на основе политомической модели Раша[[3]](#footnote-3)**

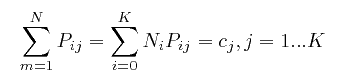
В основе методики шкалирования результатов ЕГЭ лежит политомическая модель Раша. Суть ее заключается в нехитром предположении, что при тестировании большого количества человек вероятность правильного ответа i-м испытуемым на j-й вопрос теста зависит от «трудности» δj этого вопроса и от«подготовленности» Θi этого испытуемого следующим образом:

(3) 

где K — это количество вопросов в тесте (или общее количество первичных баллов, если говорить о ЕГЭ). Величина подготовленности испытуемого определяется только количеством правильно решенных им заданий (от 0 до K, всего K+1 градаций) и не зависит от их конкретного набора. При достаточно большом количестве тестируемых количество *bi* человек, набравших *i* первичных баллов (то есть, количество испытуемых с уровнем подготовленности Θi) будет равно сумме вероятностей правильного ответа на каждое из K заданий:

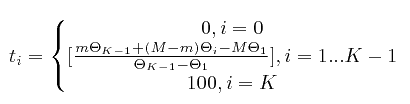
(4) 

В свою очередь, количество *cj* решивших задание с трудностью *δj* будет равно сумме вероятностей правильного ответа для каждого из N участников тестирования (каждый из которых характеризуется одним из K+1 уровней подготовленности):

(5) 

где *Ni* — количество человек, набравших *i* первичных баллов и характеризующихся одним и тем же уровнем подготовленности *Θi*.

Таким образом, у нас получается 2K+1 нелинейных уравнений с 2K+1 неизвестными — δ1, δ2, ... δK, Θ0, Θ1, ... ΘK. Решив эти уравнения, мы получим значения трудности каждого задания и, что самое важное, значения подготовленности для каждого из K+1 возможных значений набранного первичного балла. Именно эти значения подготовленности Θi непосредственно переводятся затем в шкалу тестовых баллов ЕГЭ ti по формуле:

(6) 

где m — это количество тестовых баллов, соответствующее одному набранному первичному баллу, M — количество тестовых баллов, соответствующих K-1 первичным баллам. Эти два числа — единственное, что непонятно в методике. В 2008 году они были равны соответственно 6 и 94 для ЕГЭ по всем предметам. Но в 2009 году они уже были другие, причём для всех предметов разные. Кто и на каком основании их теперь определяет — неизвестно[[4]](#footnote-4).

Итак, например, для определения шкалы перевода баллов ЕГЭ по математике в 2009 году (всего K=30 первичных баллов) нужно:

1. Собрать и обработать результаты всех выпускников;
2. Определить общее количество человек, получивших каждый конкретный первичный балл (c1...c30) и общее количество человек, набравших конкретное количество первичных баллов (b0...b30);
3. Решить систему из 2K+1 = 61-го нелинейного уравнения;
4. Перевести значения подготовленности Θ0...Θ30 в тестовые баллы t0...t30, причем пока всё перечисленное не будет сделано, можно утверждать только, что t0 = 0 и t30 = 100.

Процедура шкалирования результатов остается той же самой для ЕГЭ по любому предмету. Отличается только количество коэффициентов и уравнений. Так, по русскому языку всего 60 первичных баллов, а значит, уравнений будет 121.

**Анализ результатов ЕГЭ по некоторым предметам в 2011-м году[[5]](#footnote-5)**

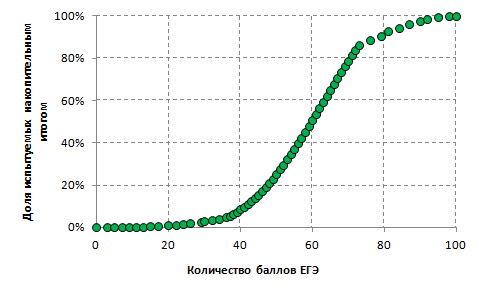


Рис. 4. Результаты ЕГЭ по русскому языку в 2011-м году. Сравните с рис. 3. Например, если вы набрали 81 балл, то попали в верхние 10% испытуемых; 60 баллов соответствует половине, тех кто сдавал экзамен.

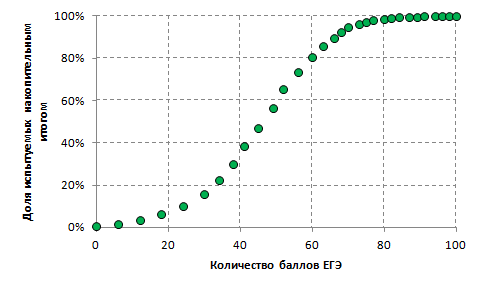


Рис. 5. Результаты ЕГЭ по математике в 2011-м году. Видно, что математика представляла большие трудности, чем русский язык. Так набранные 60 баллов позволяли вам попасть в верхние 20% испытуемых.

С любопытной критикой расчета баллов ЕГЭ можно ознакомиться в работе [«Анализ системы перевода первичного балла (i) в тестовый балл ЕГЭ (t)»](http://www.google.ru/url?sa=t&rct=j&q=&esrc=s&source=web&cd=10&ved=0CGgQFjAJ&url=http%3A%2F%2Fdistant.msu.ru%2Fpluginfile.php%2F14770%2Fmod_folder%2Fcontent%2F1%2F%25D0%25A1%25D0%25B5%25D0%25BA%25D1%2586%25D0%25B8%25D1%258F%25201-1%2F%25D0%2590%25D0%25BD%252)

1. Цитируется по книге Владимира Сергеевича Кима «[Тестирование учебных достижений](http://clipperkim.narod.ru/test/monotest/src/contents.html)». [↑](#footnote-ref-1)
2. На основании [Приложения 1](http://www.fipi.ru/binaries/1309/shkal2012.doc) к приказу ФИПИ (Федерального института педагогических измерений) от 18 апреля 2012 г. № 13-П [↑](#footnote-ref-2)
3. Цитируется по [Перевод первичных баллов ЕГЭ в шкалу тестовых баллов](http://live.mephist.ru/?mid=1273162348) [↑](#footnote-ref-3)
4. Заметка написана в мае 2010 г. Теперь уже известно – см. выше раздел «Методика шкалирования результатов ЕГЭ в 2012 году» [↑](#footnote-ref-4)
5. По данным с сайта <http://4ege.ru/novosti-ege/1532-shkala-perevoda-ballov-ege.html> [↑](#footnote-ref-5)