

**Анализ результатов
интернет-олимпиады по
физике и ЕГЭ по физике и
математике**

Монахов Вадим Валериевич
Физический факультет СПбГУ

Цель работы

Анализ заданий олимпиад и ЕГЭ с точки зрения современных теорий измерения способностей

Теория Раша

IRT - Items Response Theory

За все задания одинаковые баллы

D – сложность задания, I - способности

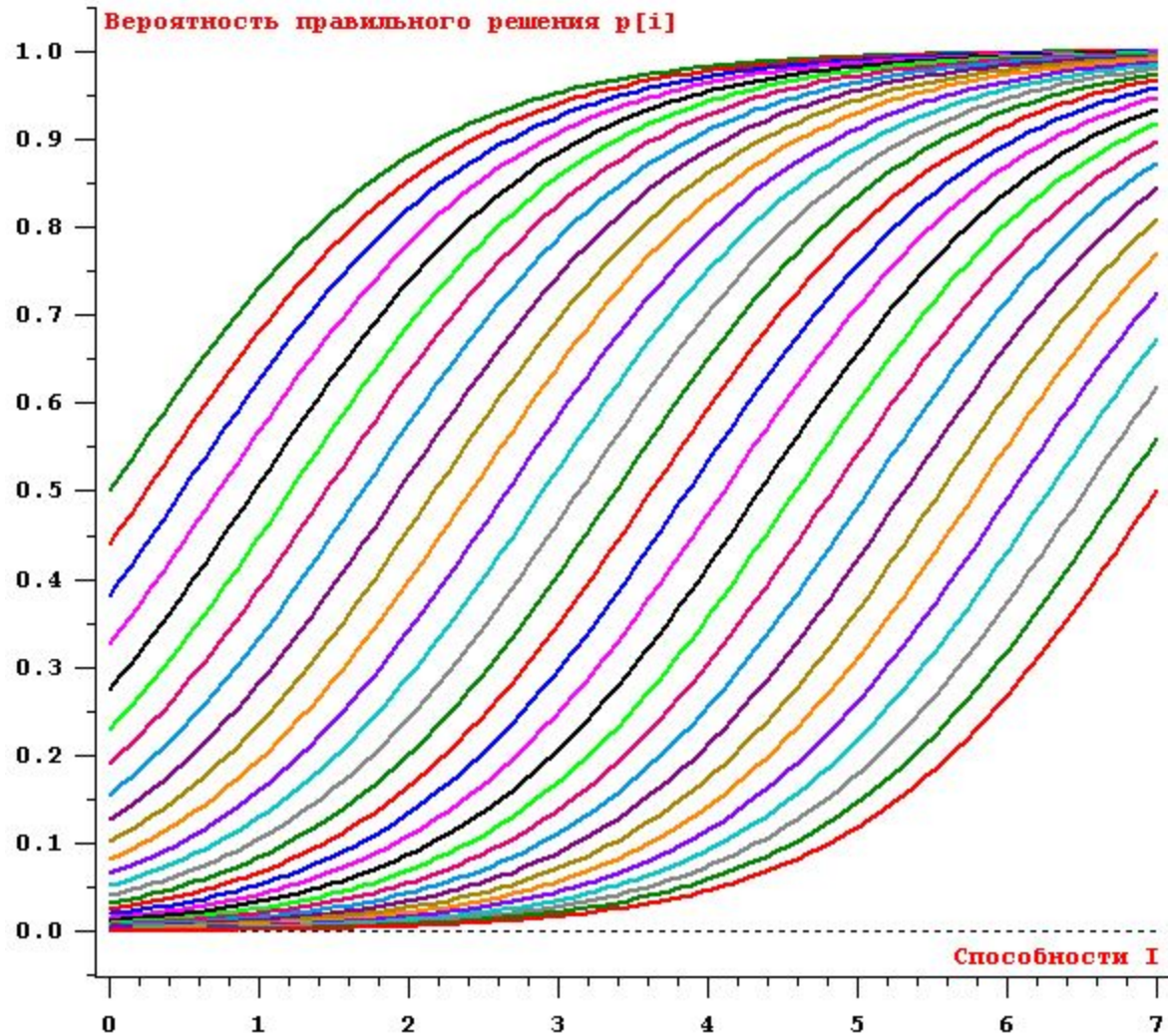
$$p = \frac{1}{1 + \exp(D - I)} \quad \text{– вероятность правильного решения}$$

$$q = 1 - p = \frac{\exp(D - I)}{1 + \exp(D - I)} \quad \text{– доля неверных ответов}$$

$$\frac{q}{p} = \exp(D - I); \quad \ln \frac{q}{p} = D - I$$

$$I = D - \ln \frac{q}{p} \quad \text{– } I \text{ не зависит от трудности } D$$

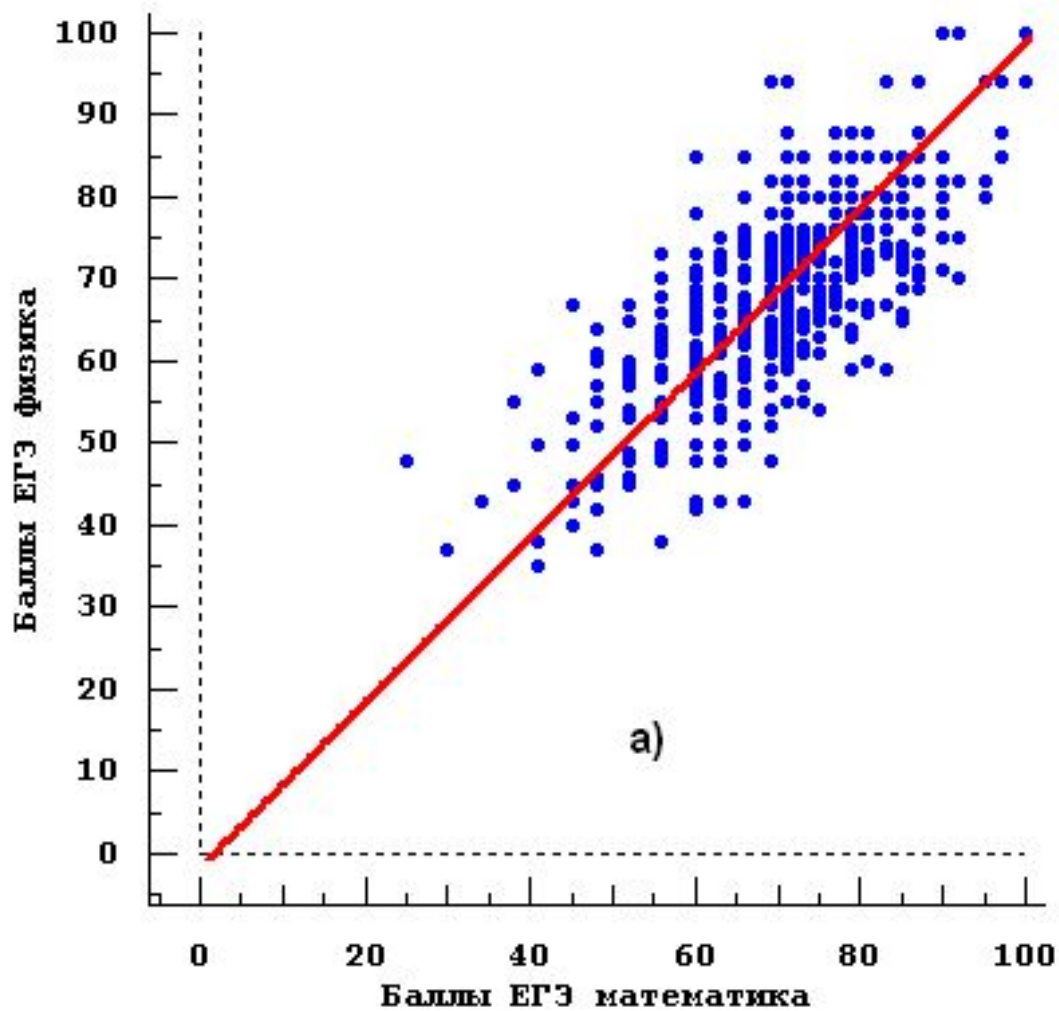
Теория Раша



Проверка IRT для олимпиад и ЕГЭ (предварительные результаты)

- Подтверждается основное положение IRT – статистический характер вероятности выполнения заданий олимпиад и ЕГЭ.
- Теория Раша может быть принята только как очень грубое приближение.
- Теория Бирнбаума, учитывающая дифференцирующие способности заданий, лучшее приближение, но также не очень хорошо описывает результаты.

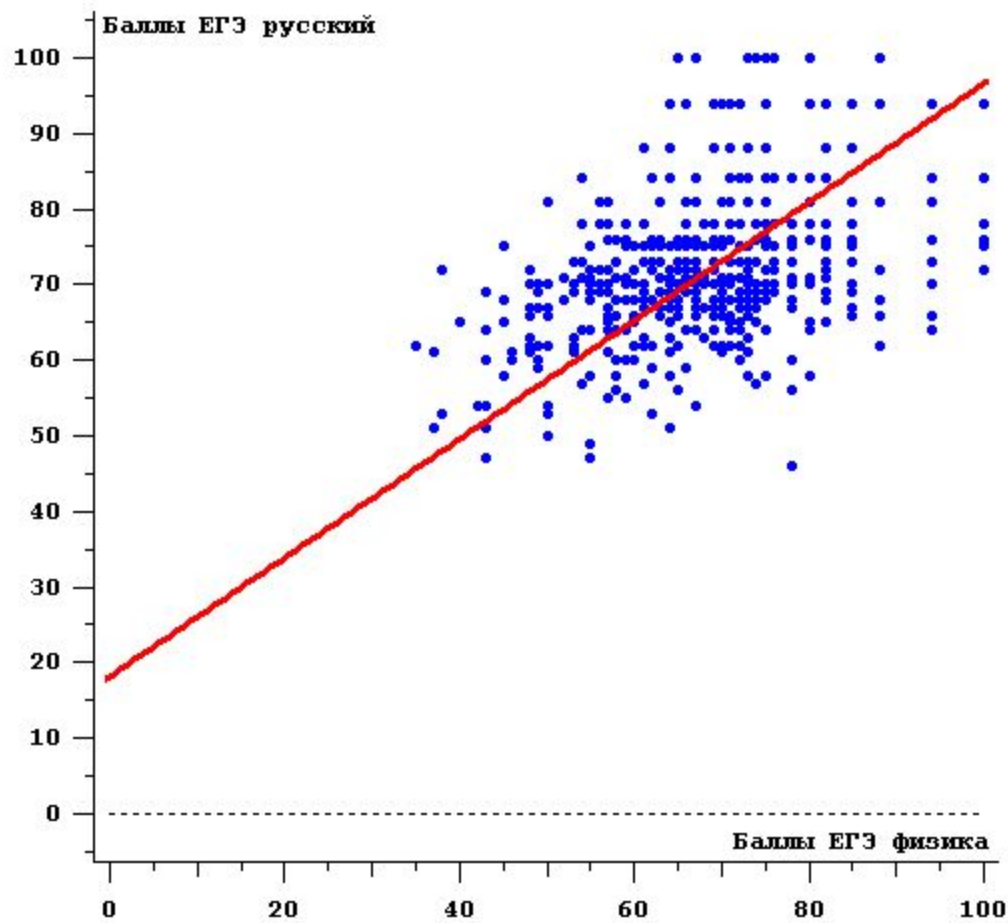
Аппроксимация результатов ЕГЭ 2010 года



Аппроксимация результатов ЕГЭ по математике и физике

- Анализ для 529 человек.
- $y=ax+b$; $a=1.00\pm 0.03$, $b=-1.7\pm 2$
- **Очень высокая корреляция: $R=0.73$**
- **Разброс $\sigma = 6.1$ балла** – если считаем, что он одинаков для ЕГЭ по физике и по математике.
- **Зона статистически значимого разброса $\pm 2\sigma = \pm 12.2$ балла** (вероятность 95%)

Сравнение результатов ЕГЭ по физике и ЕГЭ по русскому языку



Аппроксимация результатов ЕГЭ по физике и русскому

- $y=ax+b$; $a=0.78\pm0.2$, $b=18\pm15$
- **Корреляция: $R=0.42$**
- **Разброс $\sigma = 12$ баллов.**
- **Зона статистически значимого разброса $\pm 2\sigma = \pm 24$ балла (вероятность 95%).**

Прием в вузы на основании ЕГЭ по математике, физике и русскому

- Разброс ± 12.2 балла не позволяет достоверно различать способности в области уже 25 баллов – в т.ч. **75-100 баллов**.
- Прием по сумме баллов ЕГЭ по математике и физике улучшает точность измерения физико-математических способностей примерно в $\sqrt{2}$ раз.
- Добавление баллов по русскому языку – заметно ухудшает точность, особенно в области высоких баллов: зона разброса ± 24 балла превращает поступление в рулетку.

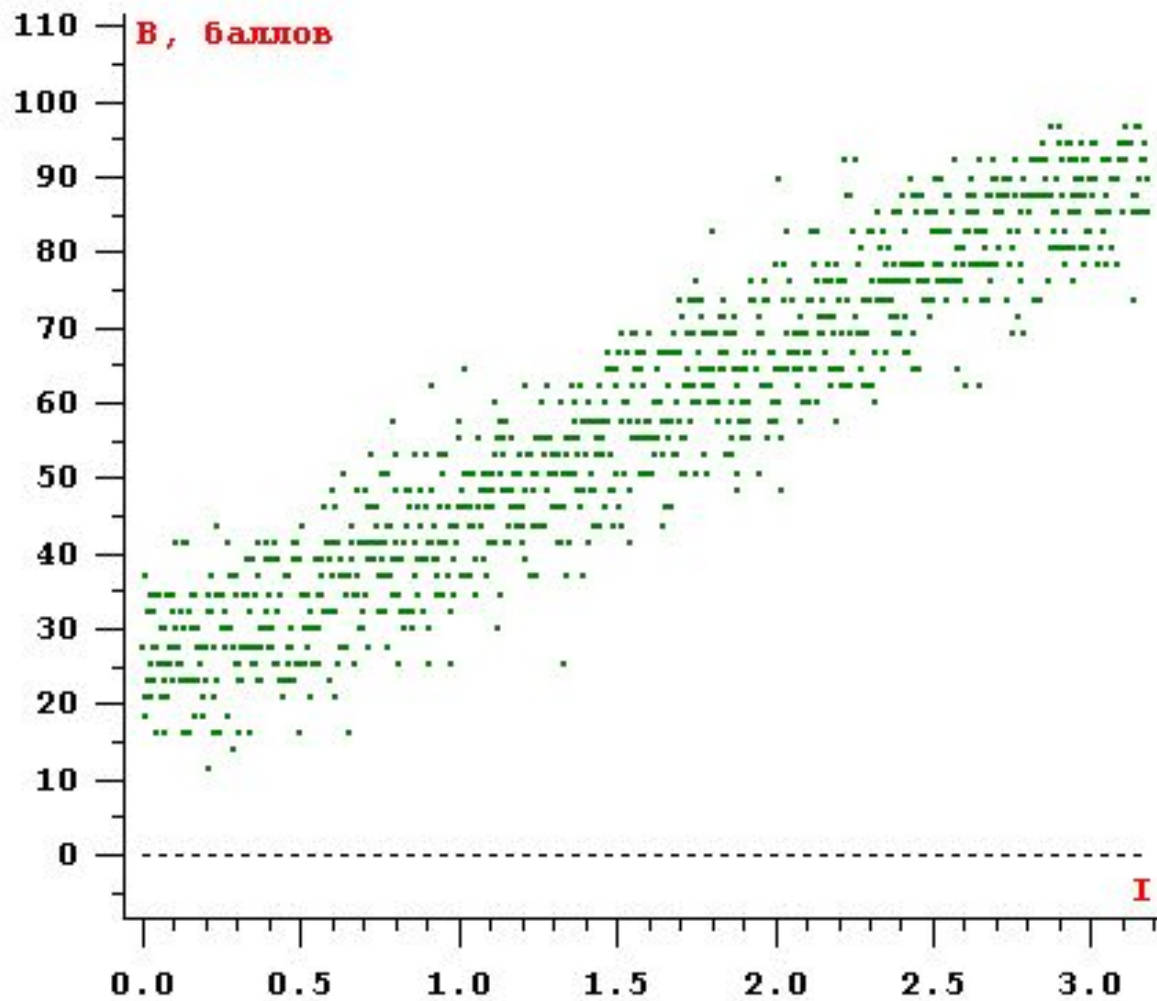
Сложность заданий ЕГЭ и интернет-олимпиады по физике

- $D = \ln \frac{q}{p}$ - для всех участников
- ЕГЭ часть А: от **-2.5** до **0.62**
- ЕГЭ часть В: от **-1.5** до **0.66**
- ЕГЭ часть С: от **-0.06** до **0.68**
- Отборочные туры: от **-1.1** до **2.3**
- **Очный тур:** от **0.67** до **2.76**

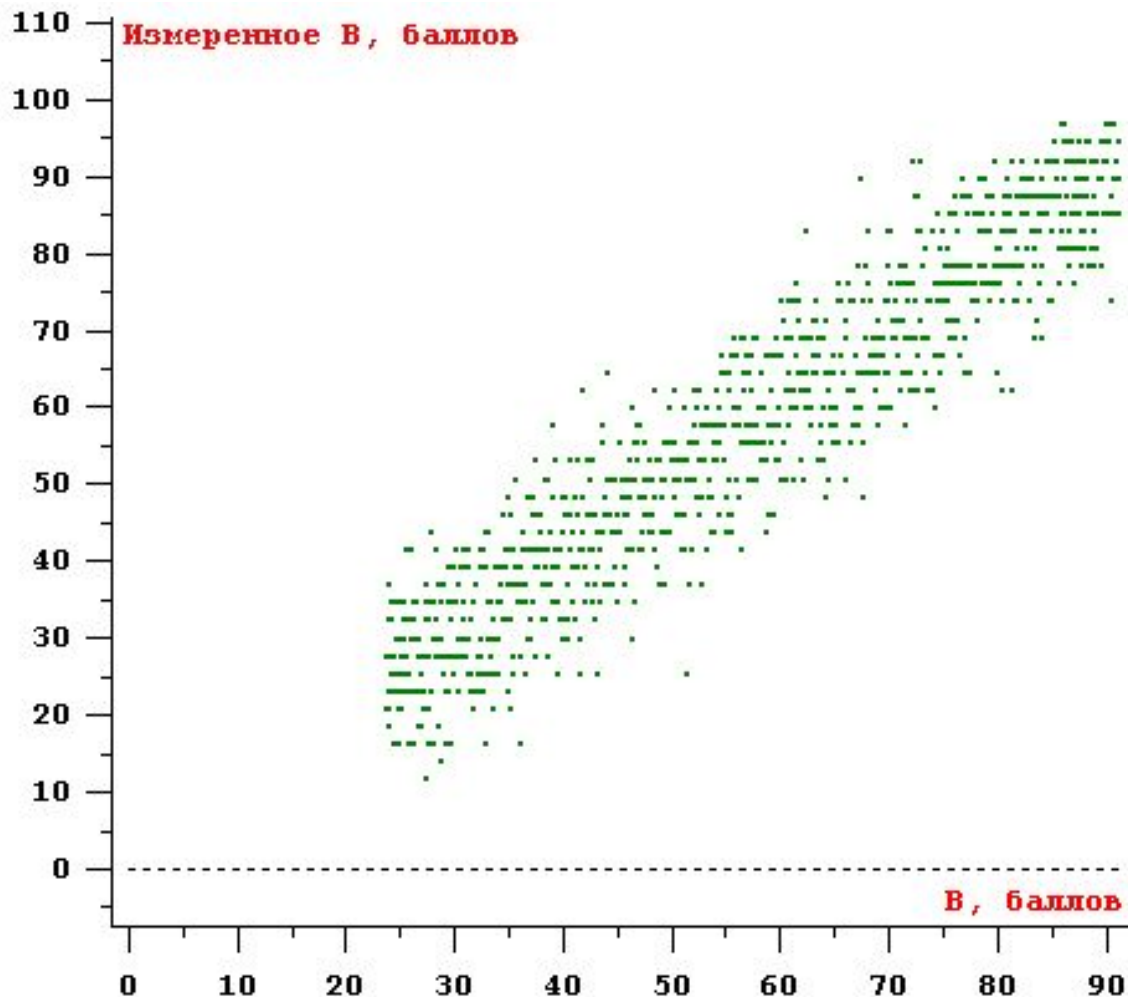
Сложность заданий ЕГЭ и интернет-олимпиады по физике

- Диапазон сложности ЕГЭ сдвинут в область **простых заданий** даже по сравнению с самыми простыми заданиями отборочных туров олимпиады.
- Сложность **самого простого** задания очного тура соответствовала **самому сложному** заданию ЕГЭ.

Модельная зависимость баллов ЕГЭ от способностей участника



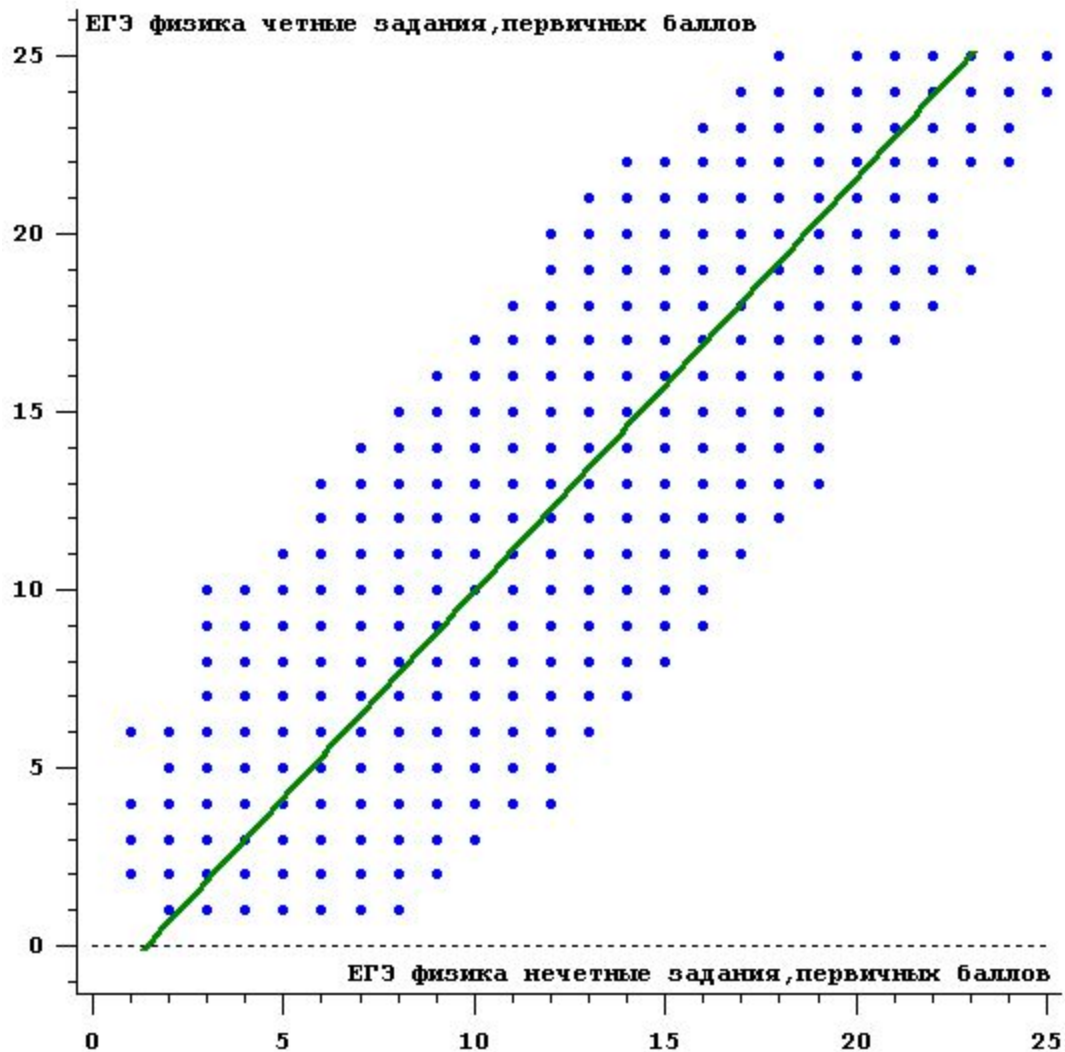
Зависимость набранных баллов от баллов при отсутствии статистического разброса



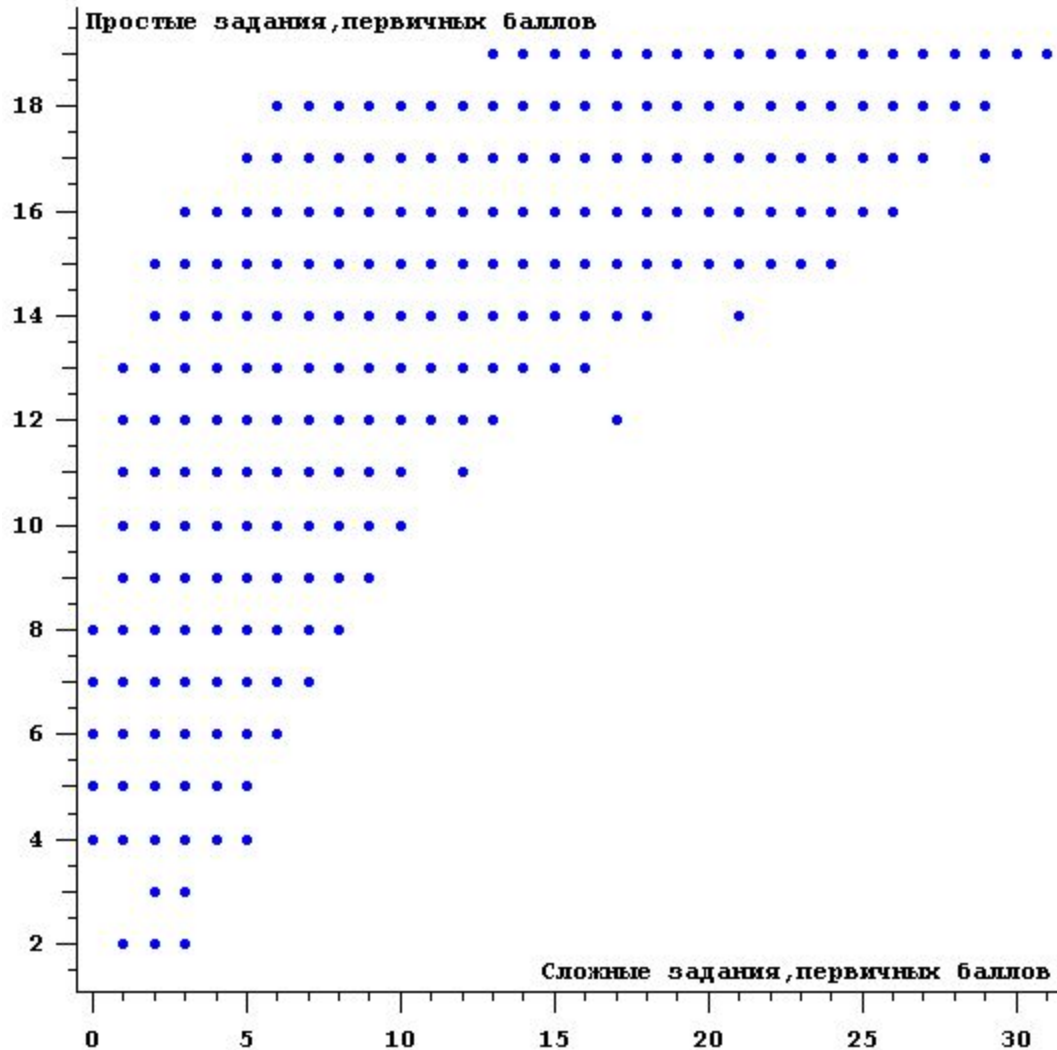
Результаты моделирования

- Среднеквадратичный разброс $\sigma = 7.0$ баллов
- Зона статистически значимого разброса $2\sigma = 14$ баллов (вероятность попадания 95%).
- Реальный разброс даже несколько меньше, чем рассчитанный по модели Раша.
- Для уменьшения разброса – либо заметно увеличивать число заданий (нереалистично), либо **уменьшать диапазон сложности заданий.**

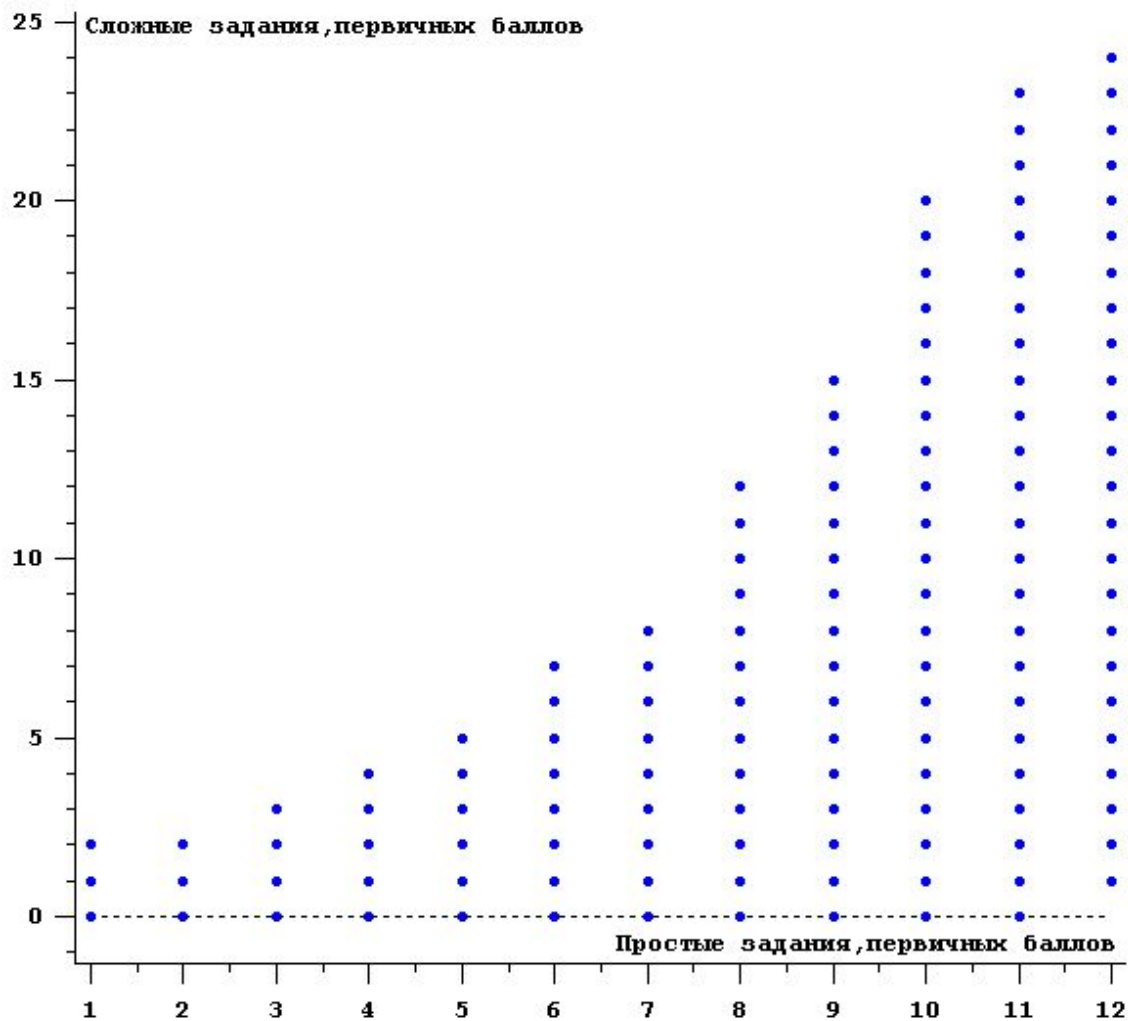
ЕГЭ Москва 2010, четные и нечетные задания



18 простых и 18 сложных заданий ЕГЭ



12 самых простых и 12 самых сложных заданий ЕГЭ



Способность выполнять простые и сложные задания

- Высокая способность выполнять простые задания соответствует очень широкому диапазону способности выполнения сложных заданий, от низкого до высокого.
- Т.е. способность **отлично** выполнять простые задания никоим образом **не гарантирует** способности выполнять **сложные** задания.
- Способность выполнять сложные задания **гарантирует** способность выполнять простые задания – но, как правило, не на высший балл.
- Сложность задания не является константой, не зависящей от способностей. Имеющиеся варианты теорий IRT не учитывают эту особенность – но рекомендуется, чтобы различие способностей участников не превышало 3 логитов.

Выводы, часть 1

- ЕГЭ по физике и математике обеспечивают высокую точность измерения способностей в области заданий низкой и средней сложности.
- Эти ЕГЭ в очень большой степени (>73%) проверяют **одни и те же способности**.
- Разброс результатов в значительной степени носит статистический характер и не может быть уменьшен без ввода дополнительных “туров” или без заметного сужения проверяемого диапазона способностей.

Выводы, часть 2

- Способность отлично выполнять простые задания никоим образом **не гарантирует** способности выполнять **сложные** задания.
- Поэтому участник, получивший 100 баллов по ЕГЭ, может не справиться со сложными заданиями олимпиадного уровня.
- Имеющаяся система поступления по результатам **ЕГЭ** оправданна для **большинства** вузов, но создает проблемы при приеме в **ведущие вузы**.

Выводы, часть 3

- Это связано с двумя причинами:
 - тем, что ширина зоны разброса перекрывает почти весь диапазон интересующих эти вузы баллов (70-100 баллов ЕГЭ).
 - тем, что ЕГЭ не проверяет способности в области высоких и сверхвысоких сложностей заданий, характерных для олимпиад.
- Имеющаяся система поступления по результатам **олимпиад** вполне оправданна при приеме в **ведущие** вузы.

Выводы, часть 4

- Разбиение ЕГЭ на **базовый** (с простыми заданиями - для получения аттестата зрелости) и **профильный** (с более сложными заданиями) позволит значительно улучшить качество ЕГЭ.
- Целесообразно **суммировать** результаты экзаменов по **математике** и **физике**, и **не суммировать** их с баллами по русскому языку. Более желательно, чтобы вуз мог назначать нижний **пороговый балл** по русскому языку (например, 60 баллов).

Предложение

- Обратиться в Минобрнауки с поддержкой варианта Стандарта образования для старшей школы, который разработан **Президиумом РАО**. В этом проекте предусмотрено два уровня сдачи ЕГЭ, базовый и профильный.
- Указать в этом обращении, что введение профильного уровня ЕГЭ не означает, что можно отказаться от приема по олимпиадам, так как ЕГЭ не сможет осуществить проверку в области высокой сложности заданий без понижения точности измерений.

Литература

- G.Rasch. On general laws and the meaning of measurement in psychology, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, IV. Berkeley, California: University of California Press, 1961, pp.321-333.
- Birnbaum A. Some Latent Trait Models and Their Use in Inferring an Examinee's Ability. In F.M. Lord and M.R.Novick. *Statistical Theories of Mental Test Scores*. Reading Mass.: Addison-Wesley, 1968. Ch.17-20. -p.397-479.
- Монахов В.В. Анализ результатов ЕГЭ по математике и физике и интернет-олимпиады по физике. Компьютерные инструменты в образовании, 2011, №1, с.50-57.
- Монахов В.В., Монахова С.В. Определение способностей учащихся с помощью олимпиад, тестов и компьютерных моделей // *Физическое образование в ВУЗах*. 2008. т. 14. № 3. С. 75-86.