С.Г. Кисельгоф
Национальный исследовательский университет «Высшая школа экономики»

МОДЕЛЬ ВЫБОРА ВУЗОВ АБИТУРИЕНТАМИ ПРИЕМНОЙ КАМПАНИИ В РОССИИ

Введение

В статье будет рассмотрена модель приемной кампании в российских государственных вузах. Для целей моделирования сделано предположение о том, что все вузы и все абитуриенты разделены на некоторое количество категорий, в зависимости от их «качества»; причем деление на категории известно всем участникам в модели. Выбор абитуриента в рамках приемной кампании ограничен пятью вузами. При подаче документов абитуриент ориентируется на качество вуза и ожидаемую вероятность поступления в этот вуз. При предположении о квадратичных функциях полезности абитуриента и экспоненциальном падении ожидаемой вероятности поступления при росте качества вуза оказывается, что абитуриент будет всегда подавать документ в вуз, поступление в который он считает гарантированным («качество» такого вуза считаем соответствующим качеству абитуриента), а также, в зависимости от параметров, в один или два вуза качеством выше на три ступени, от одного до трех вузов качеством выше на две ступени и один вуз на одну ступень выше. В результате при существующей системе приема (которая подробно описана ниже) сильно пострадают вузы уровня выше среднего: они недоберут студентов, так как сильные абитуриенты уйдут в наиболее привлекательные вузы, а более слабые будут лишены возможности перейти на освободившиеся места из-за ограниченного количества шагов. Кроме того, возможны несправедливые ситуации, когда относительно более слабые абитуриенты получат места в более сильных вузах, чем относительно более сильные.

В работе будет опущено подробное описание существующего механизма, поскольку автор предполагает, что механизм знаком большинству читателей.

Автор выражает благодарность Ф.Т. Алескерову за всестороннюю поддержку, обсуждения и ценные замечания, а также К.С. Сорокину за важные содержательные комментарии и помощь в подготовке текста.
Автор благодарит Международную лабораторию анализа и выбора решений НИУ ВШЭ за финансовую поддержку.

Подробное описание приведено в препринте, подготовленном автором по представляемой работе.

1. Математическая постановка задачи

1.1. Описание ситуации

Рассматривается прием абитуриентов на одну группу специальностей в государственные вузы в соответствии с описанными выше правилами.

- A – множество всех абитуриентов.
- B – множество всех вузов.
- A_i – множество абитуриентов уровня подготовки i, $i = \overline{\{1.M\}}$. Таким образом, абитуриенты разбиты на M категорий по уровню подготовки. Чем выше номер категорий, тем выше уровень подготовки абитуриента (результат ЕГЭ).
- B_j – множество вузов качества j, $j = i = \overline{\{1.M+1\}}$. Таким образом, вузы разбиты на $M + 1$ категорию по качеству (репутации). Чем выше номер категории, тем выше качество образования входящих в нее вузов.

Обозначим число абитуриентов в каждой группе через k_i, а число вузов в каждой категории через n. Число мест в каждом вузе примем одинаковым и равным L.

1.2. Предпочтения вузов

Все вузы имеют одинаковые предпочтения на множестве абитуриентов, устроенные следующим образом: группы абитуриентов упорядочены по предпочтительности зачисления в вуз. При этом абитуриенты, сравнить которых между собой вузы не могут, попадают в одну категорию A_i по уровню подготовки. Внутри группы с одинаковым уровнем подготовки вуз не может сравнить абитуриентов, т.е. считает их одинаковыми.

Предположение об одинаковых предпочтениях вузов является жизнеспособным в том случае, если мы рассматриваем прием только на одну группу специальностей, на которой в разных вузах требуется одинаковый набор результатов ЕГЭ, и, следовательно, абитуриенты имеют одинаковую сумму баллов с точки зрения любого вуза.

1.3. Предпочтения абитуриентов

- Вузы из более высокой группы предпочтительнее вузов из более низкой группы для любого абитуриента.

- Каждый абитуриент ранжирует вузы внутри группы вузов одинакового качества в индивидуальном порядке.

- Полезность от поступления в вуз $b \in B_j$ для некоторого абитуриента a оценивается в $(j+s_a)^2, \; 0 < s_a \le 1$.

- s_a определяется местом вуза в ранжировке данного абитуриента. А именно, если вуз является лучшим для абитуриента в данной группе, то $s_a = 1$. Если вуз является l-ым по предпочтительности в этой группе, то $s_a = \frac{n+1-l}{n}$. То есть полезность s_a задается таким образом, чтобы любой вуз категории $j + 1$ был заведомо предпочтительнее любого вуза из категории j.

- Кроме полезности от поступления, вузы оцениваются абитуриентами по вероятности поступления. Вероятность поступления зависит от уровня подготовки абитуриента (прямо) и качества вуза (обратно):

$$p = \begin{cases} 2^{i-j}, & j \ge i, \\ 1, & j < i. \end{cases}$$

Таким образом, абитуриенту $a \in A_i$ невыгодно подавать документы в вуз $b \in B_j$, если $i > j$, так как это будет заведомой потерей по сравнению с гарантированным вузом из категории $j = i$.

1.4. Выбор вуза абитуриентом

Каждый абитуриент должен выбрать набор из пяти вузов. Для каждого набора оценивается ожидаемая полезность, абитуриент выбирает набор с наибольшей ожидаемой полезностью. Будем считать, что все абитуриенты придерживаются следующего разумного принципа: «если я зачислен сразу в несколько вузов, то выбираю лучший, т.е. приносящий наибольшую полезность от поступления».

Рассмотрим абитуриента a из группы A_i и его набор, состоящий из пяти вузов. Вузы в наборе упорядочены по предпочтительности для абитуриента a: первый вуз (категория j_1) самый лучший, последний (категория j_5) – самый худший.

Оценим ожидаемую полезность поступления в лучший вуз набора:

$$\overline{u_a}(j_1, s_1) = 2^{i-j_1}(j_1 + s_1)^2.$$

Если абитуриент не поступит в лучший вуз (а это произойдет с вероятностью $p = 1 - 2^{i-j_1}$), то он будет рассматривать следующий вуз в своем наборе.

Абитуриент считает события «поступил в вуз j_1» и «поступил в вуз j_2» независимыми, поэтому вероятность того, что абитуриент не поступил в первый вуз, но поступил во второй, оценивается как произведение вероятностей этих событий.

Тогда ожидаемая полезность поступления во второй вуз набора равна

$$\overline{u_a}(j_2, s_2) = (1-2^{i-j_1})2^{i-j_2}(j_2+s_2)^2.$$

При этом, естественно, выполняется условие $j_1 + s_1 > j_2 + s_2$.

Ожидаемая полезность поступления в третий, четвертый и пятый вузы в наборе вычисляется аналогичным образом. Итоговое выражение для ожидаемой полезности абитуриента a выглядит следующим образом:

$$\begin{aligned}\overline{u_a}(j_2, s_2) &= 2^{i-j_1}(j_1+s_1)^2 + (1-2^{i-j_1})(2^{i-j_2}(j_2+s_2)^2 + \\ &+(1-2^{i-j_2})(2^{i-j_3}(j_3+s_3)^2 + (1-2^{i-j_3})(2^{i-j_4}(j_4+s_4)^2 + \\ &+(1-2^{i-j_4})2^{i-j_5}(j_5+s_5)^2))).\end{aligned}$$

2. Какой выбор сделает абитуриент?

После того как сделаны предположения о функции полезности абитуриента, можно определить, в какие вузы он будет подавать заявления. Для поиска наилучшего набора вузов мы будем решать задачу максимизации приведенной выше функции полезности, пользуясь принципом динамического программирования. Ниже приводятся полученные результаты (их доказательство см. в [Кисельгоф, 2011]).

Первое, что может быть обнаружено при таком подходе, – это то, что в качестве самого слабого вуза в наборе все абитуриенты $a \in A_i$, при $i \geq 2$, будут выбирать свой любимый вуз из соответствующей группы B_i, поступление в который абитуриент считает гарантированным.

При дальнейшем анализе оказывается, что абитуриенты разных уровней подготовки не единодушны в своем выборе. Они выбирают различные наборы вузов в зависимости от своего уровня i и числа вузов в каждой категории n, т.е. степени неопределенности.

Параметр n, напомним, показывает число вузов в одном классе качества. Если это число мало, значит, общество имеет хорошо детализированные представления о качестве вузов. Если же это число, напротив, достаточно велико, то общество и, в частности, абитуриенты находятся в ситуации существенной неопределенности, так как вынуждены опираться только на свои личные впечатления при сравнении больших групп вузов. Теперь покажем, как уровень неопределенности влияет на выбор абитуриентов. Сначала рассмотрим общую картину выбора абитуриентов при разных n. Мы будем рассматривать ситуации начиная с $n = 3$.

Ситуация $n = 2$ – неинтересная и в некотором смысле вырожденная, так как предполагает очень высокий уровень определенности в представлениях абитуриентов о вузах.

Также заметим, что мы изначально не задаем фиксированное число уровней подготовки и групп абитуриентов, а лишь говорим, что это число одинаково. Ясно, что абитуриенты самых сильных групп оказываются в особом положении в силу ограниченности выбора, поэтому их поведение должно быть рассмотрено отдельно.

2.1. Выбор при разных уровнях неопределенности

Итак, были выявлены ступени неопределенности с разным поведением групп.

1. $\boldsymbol{n = 3; 4}$.

Вуз 1	Вуз 2	Вуз 3	Вуз 4	Вуз 5	$n = 3$	$n = 4$
$(i + 3, 1)$	$\left(i+3, \frac{n-1}{n}\right)$	$(i + 2,1)$	$(i + 1,1)$	$(i,1)$	$i \geq 6$	$i \geq 3$
$(i + 4,1)$	$(i + 3,1)$	$(i + 2,1)$	$(i + 1,1)$	$(i,1)$	$2 \leq i \leq 5$	$i = 2$

Если абитуриентам известно очень подробное разбиение вузов на группы по качеству образовательных услуг (каждая группа содержит три или четыре вуза одного качества), то абитуриенты с низким уровнем подготовки выберут свои любимые вузы из категорий от B_i до B_{i+4}. Абитуриенты с более высокими уровнями подготовки выберут два вуза из категории B_{i+3} и по одному вузу из категорий B_{i+2}, B_{i+1}, B_i.

Уже здесь проявляется закономерность, которая будет наблюдаться все время: при каждом конкретном n, чем слабее группа абитуриентов, тем более рискованный набор они выбирают. Например, при $n = 3$ абитуриенты уровня $i = 5$ выберут набор $(i + 4; i + 3; i + 2; i + 1; i)$, в то время как абитуриенты уровня $i = 6$ выберут менее рискованный набор $(i + 3; i + 3; i + 2; i + 1; i)$.

2. **5 ≤ *n* ≤ 14.** На второй ступени неопределенности абитуриенты разбиваются на три группы.

Вуз 2	Вуз 3	Вуз 4	$n = 5$	$n = 14$
$(i + 2,1)$	$\left(i+2, \frac{n-1}{n}\right)$	$(i + 1,1)$	$i \geq 47$	$i \geq 13$
$\left(i+3, \frac{n-1}{n}\right)$	$(i + 2,1)$	$(i + 1,1)$	$4 \leq i \leq 46$	$i = 12$
$\left(i+3, \frac{n-1}{n}\right)$	$(i + 2,1)$	$\left(i+2, \frac{n-1}{n}\right)$	$2 \leq i \leq 3$	$i \leq 11$

Заметим, что при $n = 5$ абитуриенты всегда выбирают лучший вуз из категории B_{i+3} и худший вуз из (гарантированной) категории B_i. Различия в выборе абитуриентов разных уровней подготовки заключаются только в выборе трех средних вузов.

Назовем три группы абитуриентов условно сильными, средними и слабыми. Граница уровня подготовки, отсекающая слабых от средних, сдвигается вверх с увеличением n, а граница, отделяющая средних от сильных, сдвигается вниз с увеличением n. Таким образом, средняя группа сжимается и в конце концов исчезает.

Абитуриент из слабой группы выбирает два вуза из категории B_{i+3}, два вуза из категории B_{i+2} и гарантированный вуз своей категории B_i. Абитуриенты среднего уровня подготовки вместо второго вуза из категории B_{i+2} выбирают вуз из категории B_{i+1} и рискуют меньше, чем слабые абитуриенты. Наконец, сильные абитуриенты выбирают вместо второго вуза из категории B_{i+3} второй вуз из категории B_{i+2}. Таким образом, сильные абитуриенты рискуют в наименьшей степени.

Результаты для следующих групп абитуриентов были также рассчитаны, с ними можно ознакомиться в работе [Кисельгоф, 2011]. Однако здесь мы не будем останавливаться на них и перейдем к результатам моделирования приемной кампании.

3. Моделирование приемной кампании

Теперь, когда мы определили ожидаемое поведение абитуриентов, можно предсказать развитие приемной кампании. Для этого придется задать следующие параметры:

- параметр M, определяющий число групп качества вузов и абитуриентов;
- параметр n, определяющий число вузов в каждой группе одинакового качества и характеризующий уровень неопределенности;
- количество абитуриентов разного уровня подготовленности.

Поскольку в нашей модели считается одинаковым число вузов в группе и число мест в вузе, то места будут распределены равномерно по качеству.

Относительно предпочтений абитуриентов будем предполагать, что каждый вуз встречается на первом, втором и т.д. месте в ранжировке абитуриентов одинаковое число раз. Таким образом, мы пока не будем вводить случайную составляющую, связанную с предпочтениями абитуриентов.

3.1. Равномерное распределение абитуриентов

Рассмотрим следующий пример.

- $M = 10$, т.е. 10 категорий абитуриентов и 11 категорий вузов;
- $n = 20$, число вузов в категории;
- число абитуриентов совпадает с числом мест в вузах;
- абитуриенты распределены по категориям равномерно.

Общее число вузов будет равно $20 \cdot 11 = 220$, а мест – $220 \cdot L$. Абитуриенты распределены по категориям равномерно, поэтому

$$\forall i \;\; A_i = \frac{220L}{10} = 22L.$$

На первом этапе каждый абитуриент категорий $2 \le i \le 8$ подаст заявление в вуз своей категории, в два вуза категории B_{i+2} и два вуза категории B_{i+3}. Абитуриенты из двух верхних категорий, в принципе не имеющие категории вузов B_{i+3}, также сделают выбор, максимизирующий их ожидаемую полезность. Для абитуриента из лучшей категории A_{10} набор будет включать вуз своей категории и четыре вуза категории B_{11}. Для абитуриента из категории A_9 набор будет включать один вуз своей категории, один вуз категории B_{10} и три вуза категории B_{11}. Для абитуриента из группы A_1 наилучшим будет выбор, не включающий гарантированный вуз: один вуз из B_2, два вуза из B_3 и два вуза из B_4.

Иллюстрация демонстрирует ожидаемое число заявлений в одном вузе каждой категории. После того как спрогнозировано поведение абитуриентов, можно легко смоделировать проведение приемной кампании в соответствии с принятым в настоящее время механизмом. Ожидаемые результаты приемной кампании представлены на второй иллюстрации. В соответствии с этим прогнозом каждый вуз, кроме вузов из категории B_8, получит по $1{,}1L$ абитуриентов, а вузы восьмой категории останутся без абитуриентов вовсе.

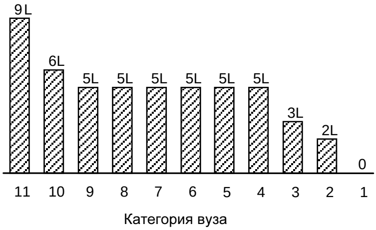

Рис. 1. Число заявлений в расчете на вуз

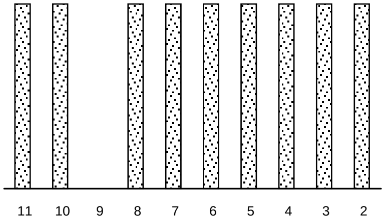

Рис. 2. Окончательное зачисление

3.2. Неполная информация о распределении абитуриентов

Рассмотрим следующий пример.

- $M = 10$, т.е. 10 категорий абитуриентов и 11 категорий вузов, однако вузы могут более четко оценить уровень подготовки абитуриента, чем сами абитуриенты. Поэтому с точки зрения вуза каждая из категорий абитуриентов разделяется на три разноуровневые группы,
- $n = 20$, число вузов в категории.
- Абитуриенты распределены по категориям равномерно.
- Число абитуриентов меньше числа мест в вузах, а именно равно $0{,}75 \cdot 220L$.

Для данного примера на рис. 3 графически представлен результат приемной кампании.

Здесь разными штриховками представлены абитуриенты из групп разного уровня подготовки (для упрощения иллюстрации абитуриенты из подгрупп внутри одной группы выделены одинаковой штриховкой. Отметим, что в такой ситуации будет наблюдаться «несправедливое» зачисление, например, абитуриенты уровня 6 зачислены в вузы группы 8, в то время как некоторые абитуриенты уровня 7 (более сильные, чем уровень 6) попали в вузы более слабой группы 7.

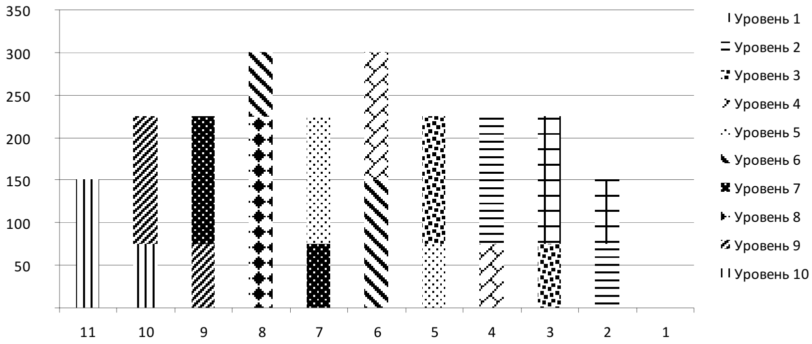

Рис. 3. Окончательное зачисление
(неполная информация)

Заключение

В данном исследовании предложен способ моделирования поведения абитуриента при выборе набора вузов для подачи заявлений, а также смоделирован ход приемной кампании для различных ситуаций – при разной имеющейся у вузов и абитуриентов информации и разном соотношении числа мест и числа абитуриентов. Данная модель позволяет высветить слабые места существующего в России механизма распределения абитуриентов по вузам. Дальнейшее развитие данного исследования предполагает два направления. Во-первых, это моделирование зачисления в течение нескольких лет, при котором абитуриенты второго и последующих лет получают дополнительную историческую информацию. Второе направление – моделирование поведения вузов как активных игроков, которое позволило бы описать и предсказать случаи и характер манипулирования механизмом зачисления.

Литература

Алескеров Ф.Т., Хабина Э.Л., Шварц Д.А. Бинарные отношения, графы и коллективные решения. М.: Изд. дом ГУ ВШЭ, 2006.

Кисельгоф С.Г. Выбор вузов абитуриентами с квадратичной функцией полезности: Препринт WP7/2011/01. М.: Изд. дом НИУ ВШЭ, 2011.

Abdulkadiroglu A., Sonmez T. School Choice: A Mechanism Design Approach // American Economic Review. 2000. Vol. 93. P. 729–747.

Balinski M., Sonmez T. A Tale of Two Mechanisms: Student Placement // Journal of Economic Theory. 1999. Vol. 84(1). P. 73–94.

Gale D., Shapley L.S. College Admissions and the Stability of Marriage // American Mathematical Monthly. 1962. Vol. 69. P. 9–14.