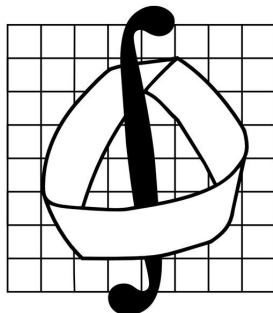


МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ имени М. В. ЛОМОНОСОВА



Механико-математический факультет
Институт "Вега"

Эконометрическое исследование на тему
**«Детерминанты проходного балла при
поступлении в вуз (на примере
факультетов МГУ с 2016 по 2019 года)»**

Выполнили:

студенты 309 группы
Самойлов Роман Александрович
Аксарина Евгения Денисовна
Дугин Максим Денисович

Научный консультант:

кандидат экономических наук
Станкевич Иван Павлович

Москва
2024

Содержание

1. Введение	3
2. Основные результаты	4
2.1. Описание данных	4
2.2. Основная регрессия	7
2.3. RE-, FE- и Pooled-модель	7
2.4. Проверка на подвыборках	9
3. Вывод	12

1. Введение

Одним из ключевых показателей, определяющих качество образования и конкурентоспособность выпускников, является проходной балл при поступлении в высшее учебное заведение. Данный показатель не только отражает академический уровень абитуриентов, но и формирует ожидания относительно образовательного процесса и перспективы трудоустройства. В контексте современной России, Московский государственный университет имени М.В. Ломоносова (МГУ) выступает в качестве одного из наиболее престижных и востребованных вузов, а факультеты этого университета привлекают студентов со всей страны и за ее пределами.

Данный исследовательский проект направлен на изучение детерминант проходного балла при поступлении на различные факультеты МГУ в период с 2016 по 2019 года. Мы рассматриваем широкий спектр факторов, которые могут оказывать влияние на проходной балл, включая общий уровень подготовки (средний балл ЕГЭ), востребованность специальностей среди компаний (количество абитуриентов, поступающих по целевой квоте), а также востребованность факультета у олимпиадников (количество абитуриентов, поступающих без вступительных испытаний).

Изучение этих детерминант имеет важное значение для понимания динамики поступления в вузы, выявления закономерностей в образовательных трендах и разработки стратегий повышения доступности высшего образования. Кроме того, это исследование поможет определить возможности для корректировки образовательной политики и даст вузам информацию для оптимизации приемной кампании в соответствии с текущими требованиями и ожиданиями общества.

В последнее время поступает много критики системы ЕГЭ, использующейся в ВУЗах России с 2001 года. Профессора утверждают о повышении стресса абитуриентов [3] после введения единого экзамена из-за наличия только одной попытки. Причем это не конец испытаний: бывшим школьникам приходится выступать в роли прогнозистов проходного балла, следить за списками поступающих, из-за нехватки баллов приходится менять стратегию поступления. Это исследование попытается ответить на вопрос о зависимости между параметрами в ранжированных списках поступающих, которые может увидеть любой абитуриент, и проходным баллом.

Выбор оптимальной стратегии при подготовке к поступлению изучался ранее. В [5] было установлено про конкурс 2010-2012 годов, что в среднем ученики выпускного класса тратили 12-14 тыс. в месяц на дополнительные занятия к ЕГЭ. Подготовка коррелировала положительно на результат экзамена, он увеличивался примерно на 3 балла(из 100) при прочих равных.

2. Основные результаты

2.1. Описание данных

Данные получены с сайтов факультетов, а также с сайта центральной приёмной комиссии МГУ [1]. После очистки они совмещены в одну таблицу. В таблице 2.1 указаны некоторые характеристики и описание участвующих в регрессии переменных.

Название	Среднее	Стандартное отклонение	Минимум	25%	50%	75%	Максимум	Описание
year	2017.50	1.12	2016.00	2016.75	2017.50	2018.25	2019.00	год
num_bvi	9.95	17.45	0.00	0.00	3.00	12.00	154.00	кол-во поступивших без вступительных экзаменов
num_target_students	1.85	4.64	0.00	0.00	0.00	2.00	38.00	кол-во поступивших по целевой квоте
num_quota_students	3.03	5.45	0.00	0.00	1.00	4.00	38.00	кол-во поступивших по особой квоте
num_student	399.99	350.04	21.00	137.75	276.50	545.25	1526.00	кол-во участников конкурса
passing_score	353.84	48.21	202.00	324.00	346.00	379.00	469.00	проходной балл первой волны
average_score	325.99	39.53	263.38	298.43	312.45	333.53	419.22	среднее кол-во баллов по всем предметам
faculty_humanities	0.50	0.50	0.00	0.00	0.50	1.00	1.00	гуманитарность факультета
abbreviation_length	20.20	12.10	6.00	10.00	15.00	29.00	55.00	длина названия факультета
num_budget_places	47.04	59.27	0.00	9.00	18.00	53.75	269.00	кол-во бюджетных мест
gender	0.35	0.18	0.05	0.19	0.34	0.44	0.76	пол (0 - жен., 1 - муж.)
specialist_or_bachelor	0.20	0.40	0.00	0.00	0.00	0.00	1.00	специалитет или бакалавриат
non_moscow_percentage	0.43	0.10	0.10	0.35	0.43	0.50	0.64	процент людей, имеющих право на общежитие
avg_additional_points	3.82	0.54	2.55	3.44	3.81	4.21	5.24	среднее кол-во дополнительных баллов
num_bronze	1.69	5.75	0.00	0.00	0.00	0.00	33.00	число результатов первокурсников на бронзу кто бег + лыжники

На рисунке 2.1 можно заметить, что диаграммы среднего и проходного балла имеют два пика. Это вызвано тем, что на некоторых специальностях требуется 4 предмета, а на некоторых 5.

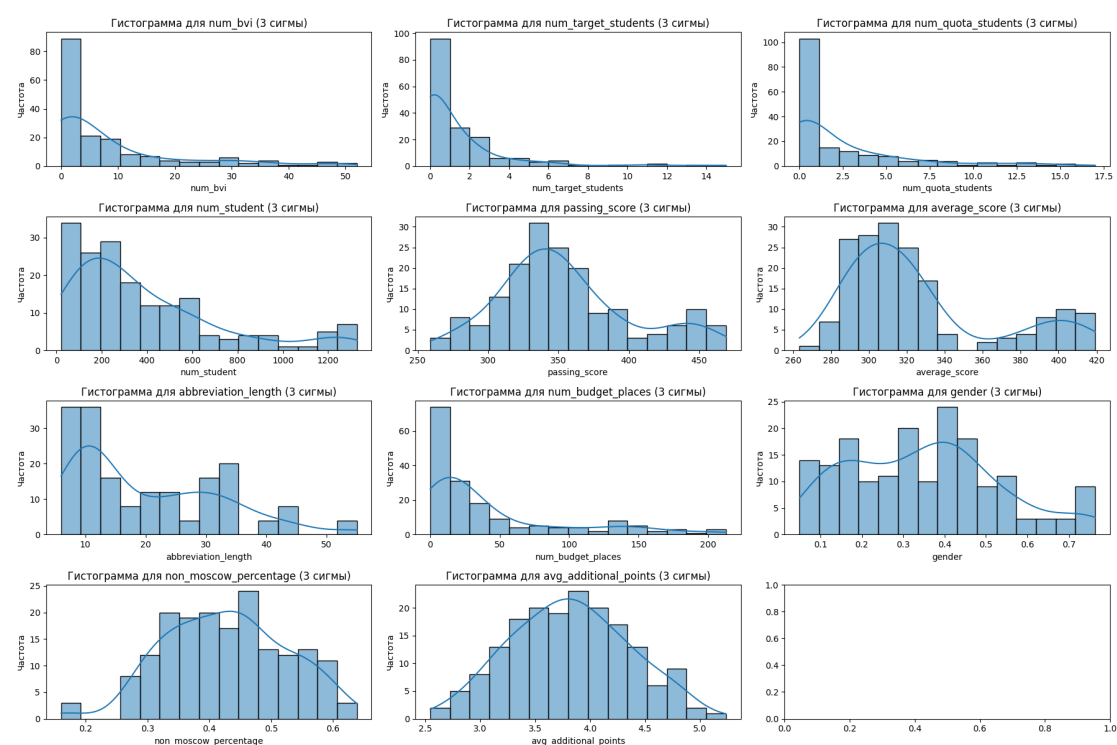
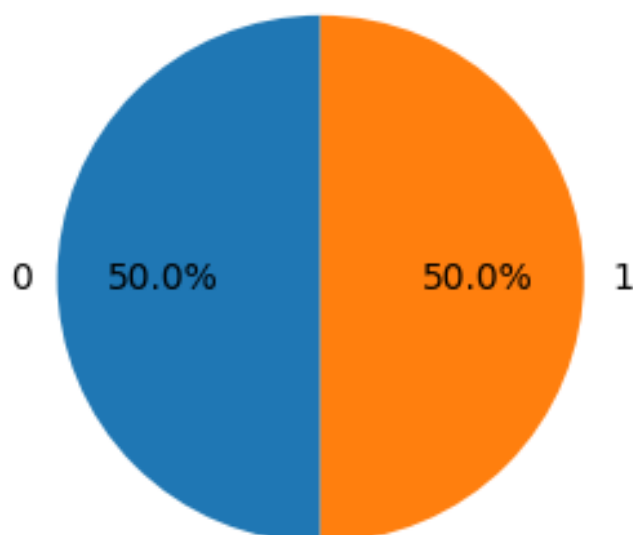


Рис. 2.1: Распределение данных

Также на 2.2 видно, что в МГУ одинаковое количество гуманитарных и технических факультетов (как минимум среди тех, что мы используем), а программ для подготовки бакалавров больше в 4 раза, чем специалитетов.

2.3 Показывает сильную связь между средним и проходным баллом, количеством абитуриентов и количеством бюджетных мест.

Соотношение значений в faculty_humanities



Соотношение значений в specialist_or_bachelor

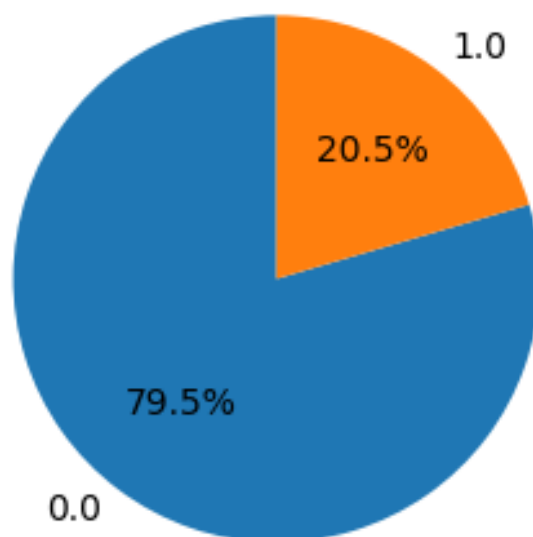


Рис. 2.2: Распределение данных

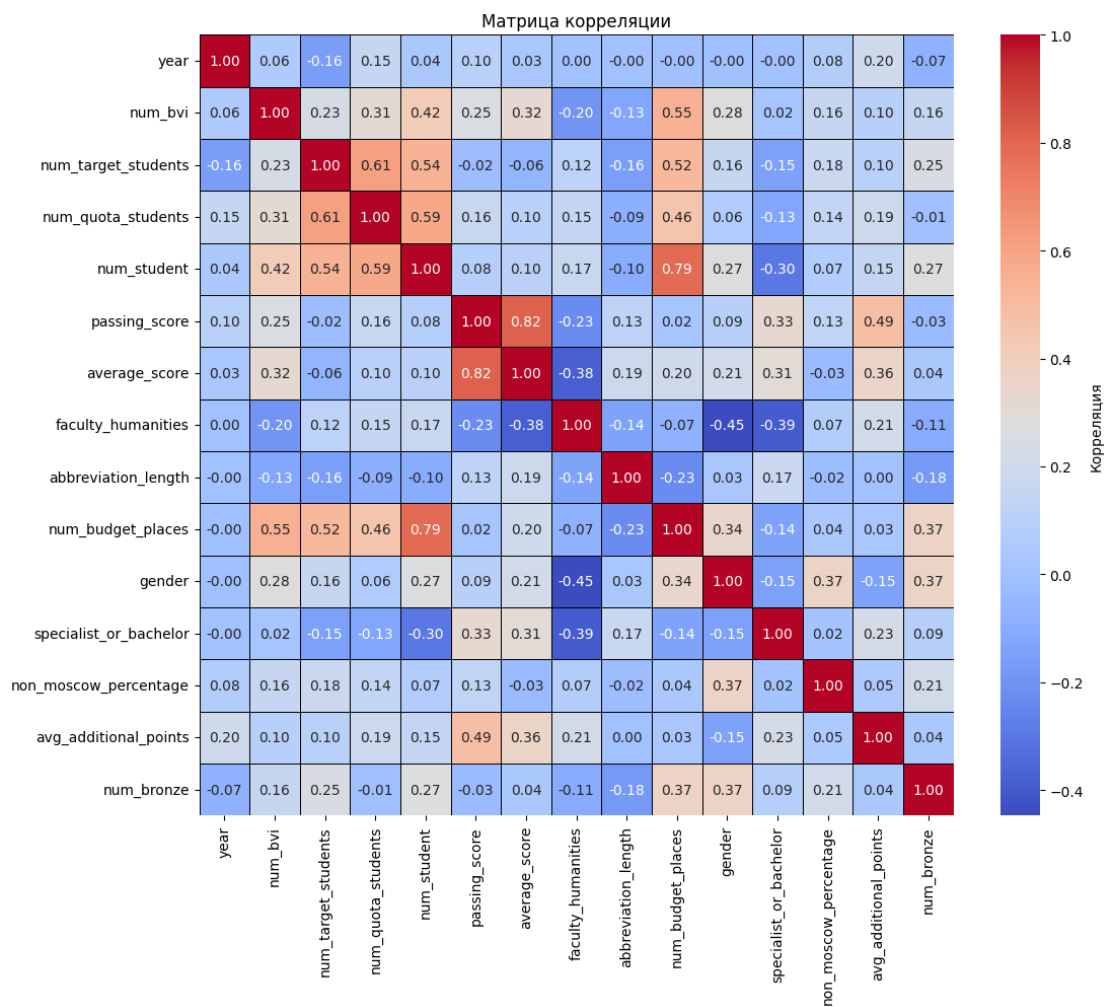


Рис. 2.3: Корреляция данных

2.2. Основная регрессия

Модель представляет собой линейную регрессию панельных данных:

$$\begin{aligned} passing_score \sim & num_bvi + num_target_students + num_quota_students + \\ & + num_student + num_student^2 + average_score + faculty_humanities + \\ & + abbreviation_length + num_budget_places + specialist_or_bachelor + \\ & + avg_additional_points + num_bronze \end{aligned}$$

Мы выбрали модель со случайными эффектами (RE) для учета неоднородности между факультетами, предполагая, что такие факторы, как местоположение, репутация, инфраструктура, распределены случайно и не коррелируют с регрессорами.

Результаты анализа представлены в таблице 2.2. На проходной балл оказали существенное влияние такие переменные, как количество абитуриентов в конкурсе, средний балл, количество бюджетных мест и формат обучения. Эти результаты можно объяснить следующим образом: чем больше абитуриентов с высокими баллами ЕГЭ участвуют в конкурсе и чем меньше количество бюджетных мест, тем сильнее конкуренция за каждое место, что приводит к повышению проходного балла. Положительное влияние формата обучения на проходной балл связано с тем, что специалитет обычно предлагается на более фундаментальных направлениях, где требуется более высокий уровень подготовки.

В нашей модели эндогенными могут быть переменные, отвечающие за средний балл и средние дополнительные баллы абитуриентов данного факультета. Действительно, возможно влияние регрессанта на эти регрессоры. Высокий проходной не дает шансов людям с низкими суммарными баллами или дополнительными достижениями и они не вступают в конкурс. И наоборот, если есть большая вариативность в баллах зачисленных, то скорее всего, и средняя сумма не близка к максимальной и можно не тревожиться из-за отсутствия золотой медали и т.д.

Для устранения эндогенности используется инструментальная переменная, обозначающая результаты студентов рассматриваемого года набора на студенческих спортивных соревнованиях. Этот показатель явным образом коррелирует со средним гто набора, а значит дополнительными баллами и общим средним в какой-то степени, и не коррелирует ни с чем другим. Показатель $numb_bronze_{it}$ отражает долю студентов в данном наборе рассматриваемого факультета, выполнивших требования на бронзовый значок ГТО (по нормам 16-17 лет) в "Кросс-тесте МГУ: Малый газон" в беге и матче "БЕГУНЫ vs. ЛЫЖНИКИ" на лыжах[2]. Из-за специфики данных объединяются непересекающиеся множества.

Также используется HC SE для устойчивости к гетероскедастичности.

2.3. RE-, FE- и Pooled-модель

В таблице 2.3 приведены результаты регрессии RE-, FE- и Pooled-модель (по столбцам). Можно заметить, что значимым во всех моделях является средний балл. Значит он устойчив.

Таблица 2.1: Модель со случайными эффектами

	<i>Dependent variable:</i>
	passing_score
num_bvi	0.117 (0.179)
num_target_students	0.223 (0.277)
num_quota_students	0.311 (0.270)
num_student	0.084** (0.033)
I(num_student^2)	-0.00003 (0.00002)
average_score	0.901*** (0.087)
faculty_humanities	-1.265 (6.402)
abbreviation_length	-0.378 (0.263)
num_budget_places	-0.377*** (0.109)
specialist_or_bachelor	19.041** (9.130)
avg_additional_points	4.816 (4.265)
num_bronze	0.205 (0.534)
Constant	34.682 (38.886)
Observations	176
R ²	0.554
Adjusted R ²	0.521
F Statistic	202.209***
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

При проведении теста Хаусмана получаем $p\text{-value} = 0.166$, что означает, что гипотезу H_0 отвергнуть не можем и лучше RE.

При проведении теста Бройша-Пагана для сравнения Pooled с RE получаем $p\text{-value} < 2.2e-16$, что означает, что гипотеза H_0 отвергается и лучше RE.

При проведении f-теста для сравнения Pooled с FE получаем $p\text{-value} < 2.2e-16$, что означает, что гипотеза H_0 отвергается и лучше FE.

Итого, $RE > FE > Pooled$, что подтверждает слова в 2.2.

2.4. Проверка на подвыборках

Проведём регрессии только для женщин и мужчин, а также для москвичей (тех, кто не имеет право на общежитие) и иногородних (тех, кто имеет).2.4.

Пройдёмся по всем параметрам:

- **Число абитуриентов, поступающих по целевой квоте**

Положительно влияет для москвичей и отрицательно для регионов. Это объясняется тем, что для москвичей данное число ассоциируется с престижностью направления. А для региона данное число показывает, что часть людей не будет участвовать в общем конкурсе, а значит можно подавать согласие с более низким баллом. Следовательно, проходной балл будет ниже.

- **Число абитуриентов, участвующих в общем конкурсе**

Для всех групп, за исключением москвичей, переменная значима и оказывает положительное влияние. Логика влияния аналогична той, что описана в разделе 2.2. Для иногородних также значим квадрат этой переменной, где вершина находится около 1100. Это означает, что если число участников конкурса превышает 1100, то проходной балл начинает снижаться.

- **Средний балл по ЕГЭ**

Как и в 2.3, устойчив.

- **Число бюджетных мест**

Устойчиво значима. Влияет отрицательно, рассуждения такие же, как в 2.2.

- **Форма обучения**

Для всех выборок, кроме москвичей, значима. Влияет положительно, рассуждения такие же, как в 2.2. Также стоит заметить, что коэффициент у мужчин практически в 2 раза больше, чем у женщин. Это связано с отсрочкой от армии.

- **Средний балл по дополнительным показателям**

Для женщин и москвичей значима. Влияет положительно, связано с тем, что конкурс для этих групп плотный и каждый балл важен.

Таблица 2.2: Сравнение моделей

	<i>Dependent variable:</i>		
	passing_score		
	(1)	(2)	(3)
num_bvi	0.117 (0.179)	0.069 (0.284)	0.239 (0.215)
num_target_students	0.223 (0.277)	−0.002 (0.349)	0.531 (0.415)
num_quota_students	0.311 (0.270)	0.346* (0.207)	0.426 (0.435)
num_student	0.084** (0.033)	0.182*** (0.066)	0.028 (0.027)
I(num_student^2)	−0.00003 (0.00002)	−0.0001** (0.00003)	0.00001 (0.00002)
average_score	0.901*** (0.087)	0.552*** (0.192)	0.955*** (0.078)
faculty_humanities	−1.265 (6.402)		−1.547 (7.032)
abbreviation_length	−0.378 (0.263)		−0.418* (0.239)
num_budget_places	−0.377*** (0.109)	−0.418 (0.280)	−0.394*** (0.109)
specialist_or_bachelor	19.041** (9.130)		12.415* (7.143)
avg_additional_points	4.816 (4.265)	5.918* (3.539)	12.355* (6.309)
num_bronze	0.205 (0.534)	0.411 (0.476)	−0.261 (0.712)
Constant	34.682 (38.886)		1.997 (30.771)
Observations	176	176	176
R ²	0.554	0.313	0.783
Adjusted R ²	0.521	0.023	0.767
F Statistic	202.209***	6.231*** (df = 9; 123)	48.938*** (df = 12; 163)

Note:

*p<0.1; **p<0.05; ***p<0.01

Таблица 2.3: Сравнение подвыборок (общая, женщины, мужчины, москвичи, не москвичи)

	<i>Dependent variable:</i>				
	passing_score				
	(1)	(2)	(3)	(4)	(5)
num_bvi	0.117 (0.179)	0.528 (0.341)	0.026 (0.175)	0.616 (0.456)	−0.004 (0.129)
num_target_students	0.223 (0.277)	0.770 (0.509)	0.239 (0.443)	0.985* (0.569)	−0.583* (0.307)
num_quota_students	0.311 (0.270)	0.113 (0.333)	0.589 (0.360)	−0.099 (0.365)	0.179 (0.240)
num_student	0.084** (0.033)	0.073* (0.039)	0.116*** (0.037)	0.063 (0.046)	0.221*** (0.043)
I(num_student^2)	−0.00003 (0.00002)	−0.00001 (0.00002)	−0.00003 (0.00002)	−0.00000 (0.00003)	−0.0001*** (0.00005)
average_score	0.901*** (0.087)	0.986*** (0.078)	0.696*** (0.131)	0.923*** (0.079)	0.509*** (0.108)
faculty_humanities	−1.265 (6.402)	1.825 (5.843)	2.617 (7.755)	0.067 (7.309)	−3.778 (8.378)
abbreviation_length	−0.378 (0.263)	−0.391** (0.195)	−0.360 (0.310)	−0.232 (0.243)	−0.257 (0.373)
num_budget_places	−0.377*** (0.109)	−0.685*** (0.160)	−0.587*** (0.146)	−0.702*** (0.237)	−0.554*** (0.113)
specialist_or_bachelor	19.041** (9.130)	16.951** (7.015)	29.835** (12.533)	10.038 (7.061)	34.255** (15.233)
avg_additional_points	4.816 (4.265)	6.151** (2.849)	−3.083 (2.556)	5.834*** (2.240)	−3.328 (2.726)
num_bronze	0.205 (0.534)	−0.091 (0.280)	0.252 (0.261)	0.007 (0.472)	0.086 (0.228)
Constant	34.682 (38.886)	5.365 (24.126)	136.015*** (47.267)	38.089 (30.209)	187.855*** (39.231)
Observations	176	176	176	176	176
R ²	0.554	0.668	0.441	0.575	0.395
Adjusted R ²	0.521	0.644	0.400	0.544	0.351
F Statistic	202.209***	328.669***	128.835***	220.613***	106.464***

Note:

*p<0.1; **p<0.05; ***p<0.01

3. Вывод

Проведя все исследования, мы понимаем, что устойчиво и значимо влияет только средний балл по ЕГЭ. То есть проходной первой волны не только зависит от уверенности в себе абитуриентов с граничным результатом, но и отражает общий спрос на факультет в этом году. Также получены полезные промежуточные результаты: специалитет влияет положительно; для иногородних чем больше 1100 человек в конкурсе, тем меньше проходной балл, и для женщин число букв в названии влияет отрицательно.

Список литературы

- [1] 1. Ранжированные списки лиц, поступающих на бюджетные места в МГУ имени М.В.Ломоносова в 2016-2019 годах
2. Ранжированные списки лиц, поступающих на бюджетные места в МГУ имени М.В.Ломоносова в 2016-2019 годах (для отдельных категорий жителей Крыма и Севастополя)
3. Приказы о зачислении в МГУ имени М.В.Ломоносова в 2016-2019 годах
4. Контрольные цифры приема в МГУ имени М.В.Ломоносова для граждан Российской Федерации, поступающих на первый курс на обучение по программам бакалавриата и программам специалитета в 2016-2019 годах
- [2] 1. Результаты. Кросс-тесты МГУ: Малый газон 2017, 2018, 2020 годов
2. Результаты. Матч "БЕГУНЫ vs. ЛЫЖНИКИ 2017 - 2019
http://polyathlon-russia.com/msuathletics/kross_MSU.html
- [3] Б.П. Елисеев *К оценке ЕГЭ как одной из форм системы отбора абитуриентов*. Научный вестник МГТУ ГА, 166(4), 98-101
- [4] Ф.С. Картаев *Дружелюбная эконометрика*.
<https://books.econ.msu.ru/Introduction-to-Econometrics/>
- [5] И.А.Прахов *Динамика инвестиций и отдача от дополнительной подготовки в вуз*. Прикладная эконометрика - №37(1) - 2015