

УДК 004.074

Б.В. Илюхин, А.В. Лепустин, Ю.Я. Кацман

Математическое моделирование влияния контекстных факторов на уровень подготовленности абитуриентов учреждений высшего профессионального образования Российской Федерации

Анализ результатов ЕГЭ выпускников является важной задачей для формирования единой системы обученности школьников. Предлагается статистический анализ результатов ЕГЭ для формирования дополнительных занятий по различным направлениям.

Ключевые слова: моделирование, тестирование, ЕГЭ, регрессия, вуз.

Существующий в Российской Федерации конкурсный отбор в вузы предполагает, что все поставлены в равные условия, но существуют территориальные, институциональные и информационные барьеры для получения высшего профессионального образования. Авторы не подвергают сомнению необходимость использования результатов объективных процедур оценки уровня обученности школьников для приема в вузы и оценки эффективности управления в сфере образования (в том числе и оценки результатов приема в вузы – средний проходной балл ЕГЭ и пр.). Тем не менее хотелось бы обратить внимание на некорректность прямой интерпретации результатов упомянутой процедуры оценки.

Показано, что образование не является замкнутой и закрытой системой, и на его результаты (индивидуальные достижения школьников) оказывают значительное влияние так называемые «контекстные факторы» [1, 2] – уровень социально-экономического развития территории, образовательный ценз родителей и прочие исследования [3, 4, 7] показали существенную зависимость результатов школьников от отношения к ЕГЭ администрации школ и учителей, а также от образовательного ценза родителей школьников. В работе [5] показано влияние на качество образования экономических факторов, а результаты исследования, доказывающего, что отдельные школы могут быть эффективны даже в сложных социальных контекстах, представлены в работе [6]. Подобные выводы содержатся и в работах, описывающих российский и международный опыт [6, 8–10].

Методика, описанная в [6], была взята как основа для проведения исследования результатов ЕГЭ выпускников школ Томской области в 2011–2012 гг. Целью исследования было выявление отдельных факторов, наиболее существенно влияющих на образовательные результаты, а также совершенствование методов корректного сопоставления различных школ [3].

Ресурсные, финансовые, социальные и иные показатели школ и контингента обучающихся были собраны в мае 2012 г. в рамках формирования социальных паспортов образовательных организаций как элемента построения региональной системы оценки качества образования.

В работе предполагалась проверка следующих гипотез:

- несопоставимость образовательных результатов отдельных школ, расположенных в различных типах населенных пунктов (крупных городов, сел, малых сел и пр.);
- влияние групп факторов [1–3] на образовательные результаты школьников;
- эффективность методики [6] оценки влияния различных факторов на образовательные результаты школьников;
- значимость различий отдельных параметров образовательных результатов в зависимости от исследуемых групп факторов;
- оценка качества линейной регрессионной модели, описывающей образовательные результаты, в зависимости от ее размерности.

Все исходные данные были сведены в единую таблицу, содержащую 229 строк (школы) и 183 столбца (переменные). Обработка данных проводилась в пакете STATISTICA.

В исследовании предполагалось, что для выпускников различных образовательных организаций (ОО) – сельских, городских, малокомплектных и пр. могут наблюдаться значимые различия в оценках эффективности.

Проведенный анализ результатов свидетельствует, что оценки эффективности для школ различных местностей значимо различны, о чем свидетельствуют не только средние значения, но и не перекрывающиеся (частично перекрывающиеся) доверительные интервалы.

Для проверки вида распределения (непараметрическая гипотеза) были использованы критерий Пирсона и критерий Колмогорова–Смирнова. Полученные исследования [11] показали, что при уровне значимости $\alpha = 0,05$ верна нулевая гипотеза – оценки ЕГЭ по двум предметам и для школ различных местностей имеют гауссово распределение.

Далее в работе исследовался вопрос, является ли различие эффективных оценок для разных ОО значимым или его можно объяснить случайными флуктуациями выборочных значений. Для ответа на этот вопрос использовался критерий Стьюдента, при этом исходная выборка была разбита на 4 непересекающиеся группы:

- 1 – школы города с населением более 500 000 человек;
- 2 – школы городов с населением не более 50 000 человек;
- 3 – малокомплектные школы;
- 4 – сельские и поселковые школы за исключением малокомплектных.

Сравнение оценок ЕГЭ по русскому языку (Р) и математике (М) городских выпускников и выпускников малокомплектных школ подтвердило значимость различия всех оценок (табл. 1).

Таблица 1

Сравнение результатов ЕГЭ для городских и малокомплектных школ (критерий Стьюдента)

Переменные	Т-тест; Группы: 1 – школы г. Томска, 2 – школы городов с населением менее 50 000 жителей, 3 – малокомплектные школы, 4 – сельские и поселковые школы (кроме малокомплектных) Группа 1: 1 Группа 2: 3										
	Среднее 1	Среднее 3	t-статистика	df – число степен. свободы	p	N 1	N 3	Стандарт. отклон. 1	Стандарт. отклон. 3	F-статистика	p
Баз_Р	74,64	67,47	4,37	96	0,000	58	40	7,783	8,272	1,13	0,67
Средн_Р	64,22	58,28	4,20	96	0,000	58	40	6,936	6,812	1,04	0,92
Баз_М	64,59	58,64	2,55	96	0,012	58	40	11,31	11,43	1,02	0,93
Средн_М	42,98	38,50	2,46	95	0,016	57	40	9,424	7,884	1,43	0,24

Описание переменных в табл. 1: Баз_Р – средний процент выполнения заданий базового уровня сложности по русскому языку ЕГЭ в ОО, Средн_Р – средний тестовый балл по русскому языку ЕГЭ в ОО, Баз_М – средний процент выполнения заданий базового уровня сложности по математике ЕГЭ в ОО, Средн_М – средний тестовый балл по математике ЕГЭ в ОО.

Полученные результаты свидетельствуют, что только на уровне менее 1,6% можно считать, что результаты ЕГЭ равны. Следовательно, верна альтернативная гипотеза – результаты ЕГЭ значимо различны. Так как применение критерия Стьюдента корректно только при нормальном распределении сравниваемых данных и равенстве соответствующих дисперсий, то в последнем и предпоследнем столбце (см. табл. 1) приведены результаты проверки равенства дисперсий двух выборок. Критерий Фишера на уровне значимости подтвердил справедливость нулевой гипотезы, а именно – применение *t*-критерия корректно. Особенно наглядно различие эффективности работы школ видно на диаграммах размаха рис. 1.

Действительно, при проецировании отклонения значения оценок средних баллов по русскому языку на величину одного среднеквадратического отклонения (т.н. «ящичков с усами») на ось ординат видно, что не только средние баллы по русскому языку у данных групп различны, но и соответствующие им 95% доверительные интервалы не перекрываются.

Проведенный анализ результатов исследований показал:

- оценки выпускников малокомплектных школ значимо отличаются от соответствующих оценок выпускников обычных школ;
- результаты ЕГЭ по русскому языку для выпускников различных групп различны, в то же время оценки по математике можно считать равными на заданном уровне значимости.

На следующем этапе было оценено влияние контекстных факторов (переменных) на оценку эффективности работы школ (оценки ЕГЭ по русскому языку и математике). Контекстные факторы, сгруппированные по различным основаниям (социально-экономические характеристики семей

школьников, материально-техническое обеспечение школ, характеристики педагогических кадров, общая информация об образовательной организации и пр.) Для оценки тесноты связи между случайными переменными X и Y в рамках линейного корреляционного анализа рассчитывался эмпирический коэффициент парных корреляций ρ_{xy} . Кроме того, были построены одномерная и многомерная регрессионные модели. Очевидно, что многомерная модель значительно корректнее описывает оценки ЕГЭ по русскому языку и математике, чем одномерная. Исключая последовательно менее значимые факторы, авторы получили модели, зависящие от трех-пяти переменных (для образовательных организаций различных типов населенных пунктов).

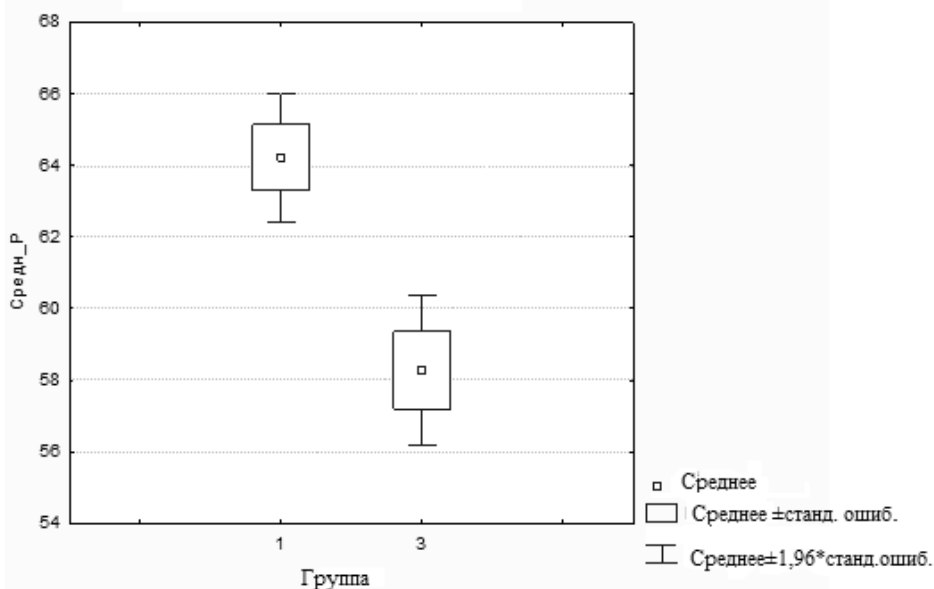


Рис. 1. Результаты ЕГЭ по русскому языку для городских и малокомплектных школ

Пример модели для школ города приведен в табл. 2.

Таблица 2

Характеристики многомерной регрессионной модели (школы города Томска области) для трех наиболее значимых переменных

Номер в таблице данных	N = 57	Результаты регрессионного анализа для зависимой переменной: Средн_Р $R = ,75859231$ $R^2 = ,57546229$ Adjusted $R^2 = ,55143185$ $F(3,53)=23,947$ $p < ,00000$ Std. Error of estimate: 4,6776					
		Beta	Std.Err. of Beta	B	Std.Err. of B	t(52)	p-level
	Intercept			59,712	2,22056	26,89041	0,000000
134	Доля детей, у которых оба родителя имеют высшее образование	0,485586	0,10447	15,068	3,24186	4,64788	0,000023
136	Доля детей, у которых один из родителей имеет высшее образование	0,183827	0,09097	8,622	4,26724	2,02055	0,048392
145	Доля школьников, состоящих на учете в КДН	-0,330754	0,10594	-111,27	35,6453	-3,1218	0,002909

Обозначения в табл. 2: N – объем выборки; R – коэффициент множественной корреляции; R^2 – коэффициент детерминации; adjusted R^2 – скорректированный коэффициент детерминации; F – критерий Фишера; p – вероятность нулевой гипотезы для F -критерия; *Standard error of estimate* – стандартная ошибка оценки (уравнения); *Intercept* – оценка свободного члена полученного уравнения регрессии; *Std. Err. of B* – стандартная ошибка оценки коэффициента; *Beta* – коэффициенты уравне-

ния регрессии; $t(52)$ – t -критерий для оценки коэффициента уравнения, где 52 – число степеней свободы в распределении Стьюдента; p -level – вероятность нулевой гипотезы для критерия Стьюдента.

Данные, приведенные в табл. 2, свидетельствуют не только о возможности описания связи образовательных результатов (баллов ЕГЭ) с приведенными в таблице факторами при помощи многомерной регрессионной модели, но и позволяют оценить качество этой модели. Кроме того, в таблице приведены полученные впервые для данной модели значения оценки качества коэффициентов уравнения регрессии. Таким образом, данные табл. 2 свидетельствуют о существовании линейной зависимости между результатами ЕГЭ и приведенными в таблице переменными, а также позволяют оценить: степень зависимости, качество модели в целом, качество каждого отдельного коэффициента уравнения регрессии в отдельности.

Промежуточные выводы:

1. Статистические исследования показали, что результаты городских и сельских школьников значимо различны и не подлежат прямому сопоставлению. Результаты ЕГЭ школьников из малых (до 70 тыс. населения) городов и крупных (более 500 тыс. населения) городов отличаются незначительно.

2. Исследования исходных данных в рамках линейного корреляционного анализа показали [11], что значимо отличны от нуля только 15–20 факторов (из ~ 180). Наиболее существенное влияние на образовательные результаты оказывают следующие факторы:

- Процент полных семей выпускников, обучающихся в данной ОО.
- Процент родителей, имеющих высшее образование среди обучающихся в данной ОО.
- Процент обучающихся, состоящих на различных формах учета (инспекции КДН и пр.).
- Процент участия обучающихся в олимпиадах (конкурсах) различного уровня.

3. Методами математической статистики подтверждена эффективность методики [3] по оценке влияния различных факторов на образовательные результаты школьников.

4. Проведенные исследования для четырех оценок качества обучения (по двум предметам базовые и средние оценки) показали несущественные различия между отдельными параметрами образовательных результатов. Это позволяет ограничиться исследованием только одних, например средних.

5. Для оценки качества регрессионной модели используется коэффициент детерминации. Линейная многомерная регрессионная модель оказалась более качественной, на что указывает более высокий коэффициент детерминации (около 58%), нежели одномерная (43, 5%). Но, тем не менее, очевидна необходимость дальнейшего исследования степени влияния отдельных факторов внутри различных групп школ (городских, сельских и пр.).

6. Неудовлетворенность качеством многомерной регрессионной модели (желательно, чтобы коэффициент детерминации был более 75%) требует дальнейших исследований. Одним из возможных путей исследований представляется применение методов многомерного факторного и кластерного анализа, построение нелинейных регрессионных моделей для повышения качества стохастической модели.

Для совершенствования образовательного процесса и повышения эффективности работы учреждений высшего профессионального образования (особенно реализующих программы бакалавриата, магистратуры и специалитета технической направленности) целесообразно учитывать результаты настоящего исследования.

При приеме абитуриентов на первый курс учреждений высшего профессионального образования необходимо учитывать территорию расположения образовательной организации, которую окончил абитуриент. Несмотря на то, что действующее законодательство не предполагает предпочтений выпускникам малокомплектных сельских школ при поступлении в вузы, представляется целесообразным внедрение моделей как дистанционной подготовки таких абитуриентов, так и разработка и внедрение системы помощи им в течение первых лет обучения.

В качестве таких мер целесообразно рекомендовать следующие:

– проведение дополнительных занятий с абитуриентами, выпускниками городских школ, набравшими по сумме трех вступительных испытаний (русский язык, математика, физика или информатика) не более 170 баллов, и зачисленными по конкурсу на первый курс направлений и специальностей технической направленности. Возможным и необходимым является проведение подобного рода занятий в период с 10 августа по 1 сентября. Данные занятия необходимы (на примере изучения авторами учебных планов технических специальностей радиотехнического факультета ТУСУРа) по следующим разделам математики – логарифмы, производные, тригонометрические уравнения и неравенства, геометрия (стереометрия), теория вероятностей; физики – динамика, кинематика,

электростатика, электричество и магнетизм; информатики – формирование алгоритмического мышления, представление данных в памяти ЭВМ. Для выпускников школ сельской местности, зачисленных на первый курс, подобные занятия необходимы при условии набора ими при поступлении не более 150 баллов.

Таким образом, с целью выравнивания знаний студентов первого курса, независимо от типа ОО предлагается следующее:

– смещение курса «Введение в специальность» на первые четыре-шесть недель первого семестра, выделение для данного курса в расписании отдельных дней занятий и проведение в такие дни не только лекционных занятий, но и экскурсий на действующие предприятия отрасли;

– формирование для студентов первого курса индивидуального плана выравнивания знаний и умений, используя результаты сдачи единого государственного экзамена. Использование детализированных результатов абитуриента (разложенных в разрезе кодификаторов элементов содержания и требований – сайт www.fipi.ru) позволит предельно точно сформулировать элементы содержания дисциплин, проблемных для каждого отдельного студента;

– прикрепление к каждому студенту-первокурснику куратора (отлично успевающего студента старшего курса, магистранта), в обязанности которого будет входить знакомство данного студента со студенческими конструкторскими бюро, группами проектного обучения, различными объединениями студентов, будет способствовать формированию представления об инженерной деятельности, будущей специальности и формировать положительную мотивацию на обучение и восходящую мобильность данного студента.

Литература

1. Болотов В.А. Информирование различных целевых групп как условие эффективного использования результатов оценки учебных достижения школьников [Электронный ресурс] / В.А. Болотов, И.А. Вальдман // Проблемы современного образования. – 2012. – № 6. – С. 187–202. – Режим доступа: http://www.pmedu.ru/res/2012_6_13.pdf

2. Болотов В.А. Условия эффективного использования результатов оценки образовательных достижений школьников / В.А. Болотов, И.А. Вальдман // Педагогика. – 2012. – №6. – С. 67–73.

3. Илюхин Б.В., Фоновые факторы, влияющие на результаты ЕГЭ / Б.В. Илюхин, В.В. Кашпур, М.В. Рачилина. – Томск: Дельтаплан, 2008. – 100 с.

4. Илюхин Б.В., Общественное участие в процедурах оценки качества образования как элемент открытости системы образования в целом // Открытость образования: разные взгляды – общие ценности: сб. материалов / Обществ. палата Рос. Федерации, Комиссия по развитию образования; Нац. исслед. ун-т «Высшая школа экономики», Ин-т образования. – М.: Изд. дом Высшей школы экономики, 2013. – С. 193–200.

5. Индикаторы оценки качества образования / М.Л. Агранович, П. Кондрашов // Директор школы. – 2007. – № 5. – С. 5–16. – (Образовательная политика).

6. Школы, эффективно работающие в сложных социальных контекстах / М.А. Пинская, С.Г. Косарецкий, И.Д. Фрумин // Вопросы образования. – 2011. – № 4. – С. 148–177.

7. Давыдова Е.А. Проектное обучение – парадигма элитного инженерного образования в России в условиях стратегии инновационного развития / Е.А. Давыдова, Р.В. Мещеряков, А.А. Шелупанов // Высшее образование сегодня. – 2006. – №8. – С. 9–15.

8. Илюхин Б.В. Оценка качества образования и принцип разумной достаточности // Народное образование. – 2012. – № 6. – С. 118–126.

9. Илюхин Б.В. Как построить систему оценки качества образования? / Б.В. Илюхин, П.И. Горлов // Журнал руководителя управления образованием. – 2012. – № 6. – С. 41–46.

10. Илюхин Б.В. Возможность использования ресурсов Центра оценки качества образования Томской области для создания элементов региональной системы оценки качества образования / Б.В. Илюхин, П.И. Горлов // Качество образования в Евразии. – 2013. – №1. – С. 123–135.

11. Илюхин Б.В. Влияние контекстных факторов на оценку эффективности работы школ Томской области / Б.В. Илюхин, Ю.Я. Кацман, А.В. Лепустин // Современные проблемы науки и образования. – 2014. – № 6. – URL: <http://www.science-education.ru/120-16117>.

Илюхин Борис Валентинович

Доцент каф. радиотехнических систем ТУСУРа

Тел.: 8 (382-2) 42-63-27

Эл. почта: bvi@tusur.ru

Лепустин Алексей Владимирович

Ст. преподаватель каф. вычислительной техники (ВТ)

Национального исследовательского Томского политехнического университета (НИТПУ)

Тел.: 8 (382-2) 42-63-27

Эл. почта: kim@ege.tomsk.ru

Кацман Юлий Янович

Доцент каф. ВТ НИТПУ

Тел.: 8 (382-2) 42-63-27

Эл. почта: katsman@tpu.ru

Iluchin B.V., Lepustin A.V., Kazman Yu.Ya.

The mathematical modeling of the impact of context factors on the educational level of high school graduates in the Russian Federation

The analysis of the unified state exam results of school graduates is an important issue in the view of the creation of the uniform educational system. The statistical analysis of the unified state exam results is proposed to form the additional studies in different education ways.

Keywords: modeling, testing, unified state exam, regression, institute of higher education.
