С.М. ВЕРТЕШЕВ, Д.О. ПРОКОФЬЕВ, А.А. ХВАТЦЕВ

МЕТОДИКА УЧЁТА ВЛИЯНИЯ РЕЗУЛЬТАТОВ ЕГЭ НА УСПЕВАЕМОСТЬ СТУДЕНТОВ В ПЕРВОМ СЕМЕСТРЕ

Излагается методика применения дисперсионного и регрессионного анализов для определения степени влияния ЕГЭ по математике на результаты первой экзаменационной сессии по математическим дисциплинам в ППИ.

После введения в 2009 году в средней школе обязательных единых государственных экзаменов (ЕГЭ) по математике и русскому языку в обществе, а главным образом среди специалистов так или иначе связанных с подготовкой школьников и обучением студентов, развернулась острейшая дискуссия. Предметом этой дискуссии являются два вопроса: 1) действительно ли результаты ЕГЭ соответствуют уровню знаний современных выпускников школ; 2) позволяют ли ЕГЭ отобрать в вузы самых достойных студентов из числа абитуриентов.

Многолетний опыт работы одного из авторов данной статьи в качестве эксперта по проверке задач части С ЕГЭ по математике позволяет однозначно сделать следующее заключение. Несомненно, уровень задач части С вполне соответствует уровню задач традиционного вступительного экзамена по математике, а часто даже и превосходит их по трудности.

Однако, большой стаж (свыше 30 лет) преподавательской работы в ППИ позволяет также заключить: уровень базовых знаний по математике у студентов первого курса из года в год неуклонно снижается. Нынешние первокурсники значительно меньше владеют фактическим материалом, чем выпускники школ пятнадцати- двадцати летней давности. Многие из поступивших на первый курс института с трудом могут вспомнить хотя бы одну из тригонометрических формул, изучаемых в школе, у большинства отсутствуют навыки построения геометрических иллюстраций, выполнения алгебраических преобразований и т. п.

В [1] на основе методики, предложенной П.В. Герасименко [2], одним из авторов данной работы проведено исследование влияния оценок полученных на вступительном экзамене по математике, на результат экзамена по математическому анализу в первом семестре на факультете автоматизации и вычислительной техники (ФАВТ). Результат проведённого исследования следующий: на 20% разброс оценок на экзамене по математическому анализу в первом семестре можно объяснить за счёт различий в подготовке по элементарной математике поступивших на ФАВТ в 1998 году.

Несомненно, представляет большой интерес оценить степень влияния уже результатов ЕГЭ на успеваемость студентов в первом семестре и сравнить это влияние с результатами работы [1].

Успехи студента определяются не только базовыми знаниями, полученными им в школе, но также и его подготовленностью обучаться в вузе, умением самостоятельно изучать обширный фактический материал, организацией учебного процесса в вузе, мотивацией в обучении, «климатом» в академической группе, бытовыми условиями студента и многими другими причинами.

Количественный показатель знаний студента – его оценка на экзамене, как и оценка на вступительном экзамене, по своей сути является случайной величиной, так как её выставляют, в основном, как результат проверки небольшой выборки из большого объёма проверяемого материала. Пусть случайные величины η и ξ – оценки семестрового и вступительного экзаменов, которые соответственно могут принимать значения $\boldsymbol{y}_0 = 2, \boldsymbol{y}_1 = 3, \boldsymbol{y}_2 = 4, \boldsymbol{y}_3 = 5$ и $\boldsymbol{x}_1 = 3,\ \boldsymbol{x}_2 = 4,\ \boldsymbol{x}_3 = 5$.

В работе на основе дисперсионного анализа результатов экзаменов предлагается методика, которая устанавливает степень влияния результатов ЕГЭ на экзамен по математике в первом семестре, определяющем успешную математическую подготовку в последующих семестрах

В качестве результативного признака примем случайную величину η , которая принимает значения $y_j = 2 + j,\ (j = 0,1,2,3)$, а за факторный признак примем случайную величину ξ , которая принимает значения $x_i = 2 + i,\ (i = 1,2,3)$. Основные показатели, используемые в методике, приведены в таблице 1.

Таблица 1

$\xi \backslash \eta$	$y_0 = 2$	$y_1 = 3$	$y_2 = 4$	$y_3 = 5$	$n_i = \sum_{j=0}^{3} m_{ij}$	групповое среднее $\overline{y}^{(i)}$	групповая дисперсия $\widetilde{\sigma}_{iy}{}^2$
$x_1 = 3$	m_{10}	m_{11}	m_{12}	m_{13}	n_1	$\overline{y}^{(1)}$	$\widetilde{\sigma}_{1y}^2$
$x_2 = 4$	m_{20}	m_{21}	m_{22}	m_{23}	n_2	$\overline{y}^{(2)}$	$\widetilde{\sigma}_{2y}^2$
$x_3 = 5$	m_{30}	m_{31}	m_{32}	m_{33}	n_3	$\overline{y}^{(3)}$	$\widetilde{\sigma}_{3y}^2$
$m_j = \sum_{i=1}^{3} m_{ij}$	m_0	m_1	m_2	m_3			
групповое среднее $\overline{x}^{(j)}$	$\overline{x}^{(0)}$	$\overline{x}^{(1)}$	$\overline{x}^{(2)}$	$\overline{x}^{(3)}$	$N = \sum_{i=1}^{3} n_i = \sum_{j=0}^{3} m_j$		
групповая дисперсия $\widetilde{\sigma}_{jx}{}^2$	$\widetilde{\sigma}_{0x}^2$	$\widetilde{\sigma}_{1x}^2$	$\widetilde{\sigma}_{2x}^2$	$\widetilde{\sigma}_{3x}^2$			

В этой таблице число, стоящее на пересечении строки x_i и столбца y_j, – частота m_{ij} – количество студентов, получивших в семестре на экзамене по математике оценку y_j из группы абитуриентов, поступивших в вуз с оценкой вступительного экзамена по математике x_i.

$$n_i = \sum_{j=0}^{3} m_{ij},\ (i = 1,2,3); \tag{1}$$

$$m_j = \sum_{i=1}^{3} m_{ij},\ (j = 0,1,2,3); \tag{2}$$

$$\overline{x}^{(j)} = \frac{1}{m_j} \sum_{i=1}^{3} m_{ij} x_i,\ (j = 0,1,2,3) \tag{3}$$

– групповые средние значения оценок вступительного экзамена тех студентов, которые получили на семестровом экзамене оценку y_j.

$$\widetilde{\sigma}_{jx}{}^2 = \frac{1}{m_j} \sum_{i=1}^{3} \left(x_i - \overline{x}^{(j)}\right)^2 m_{ij},\ (j = 0,1,2,3) \tag{4}$$

групповые выборочные дисперсии

$$\overline{y}^{(i)} = \frac{1}{n_i} \sum_{j=0}^{3} m_{ij} y_j,\ (i = 1,2,3) \tag{5}$$

– групповые средние значения оценок семестрового экзамена студентов, которые получили на вступительном экзамене по математике оценку x_i.

$$\tilde{\sigma}_{iy}{}^2 = \frac{1}{n_i}\sum_{j=0}^{3}\left(y_j - \bar{y}^{(i)}\right)^2 m_{ij}, \ \ (i = 1,2,3) \tag{6}$$

– групповые выборочные дисперсии.

Между ξ и η существует стохастическая зависимость, так как каждому фиксированному значению x_i соответствует не одно, а множество значений y_i, причем сказать до семестрового экзамена, какое именно значение примет η нельзя. Как известно $[3]$, функцией, которая наилучшим образом (в смысле среднего квадратического отклонения) описывает зависимость между величинами η и ξ, является регрессия η на ξ, т. е. условное математическое ожидание $M(\eta|\xi = x) = \varphi(x)$. В классе линейных функций корреляционную зависимость между оценками на семестровом экзамене (y) и оценками на ЕГЭ (x) наилучшим образом представляет уравнение линейной регрессии. Степень стохастической зависимости η от ξ измеряется коэффициентом детерминации, который показывает, какая доля дисперсии η объясняется корреляционной зависимостью η от ξ. Остальная часть дисперсии η обусловлена зависимостью η от случайных (остаточных) факторов, которые влияют только на η и не передают свое влияние на η через ξ.

Зная распределение долей дисперсии, обусловленных различными факторами, можно оптимально распределить усилия по устранению недостатков в организации учебного процесса.

Разброс оценок семестрового экзамена вокруг их среднего значения

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{3}\bar{y}^{(i)} n_i \tag{7}$$

характеризует выборочная дисперсия

$$\tilde{\sigma}_y^2 = \sum_{j=0}^{3}\left(y_j - \bar{y}\right)^2 m_j \tag{8}$$

Для определения коэффициента детерминации $\tilde{\rho}_{\eta|\xi}{}^2$ представим выборочную дисперсию оценок семестрового экзамена $\tilde{\sigma}_y{}^2$ в виде суммы

$$\tilde{\sigma}_y{}^2 = \tilde{\sigma}_\phi{}^2 + \sigma_0{}^2, \tag{9}$$

где $\tilde{\sigma}_\phi{}^2$ – выборочная дисперсия, характеризующая разброс условных математических ожиданий $M(\eta|\xi = x_i)$ относительно $M(\eta)$, а $\tilde{\sigma}_0{}^2$ – дисперсия остаточных факторов.

$$\tilde{\sigma}_\phi{}^2 = \frac{1}{N}\sum_{i=1}^{3}\left(\bar{y}^{(i)} - \bar{y}\right)^2 n_i \tag{10}$$

$$\tilde{\sigma}_0{}^2 = \frac{1}{N}\sum_{i=1}^{3}\tilde{\sigma}_{iy}{}^2 n_i \tag{11}$$

Тогда

$$\tilde{\rho}_{\eta|\xi}{}^2 = \frac{\tilde{\sigma}_\phi{}^2}{\tilde{\sigma}_y{}^2} \tag{12}$$

Уравнение линейной регрессии, устанавливающее зависимость уровня средних оценок между семестровым экзаменом и оценок ЕГЭ может быть представлено в виде

$$y = \tilde{r}_{xy} \frac{\tilde{\sigma}_x}{\tilde{\sigma}_y}\left(x - \bar{x}\right) + \bar{y}, \tag{13}$$

где $\bar{x} = \frac{1}{N}\sum_{j=0}^{3}\bar{x}^{(j)}m_j$ – средний балл поступивших на первый курс, $\tilde{\sigma}_x^2 = \sum_{i=1}^{3}\left(x_i - \bar{x}\right)^2 n_i$ – дисперсия оценок ЕГЭ.

Выборочный коэффициент корреляции $\tilde{r}_{xy}$ вычисляется по формуле

$$\tilde{r}_{xy} = \frac{\tilde{m}_{11} - \bar{x}\cdot\bar{y}}{\tilde{\sigma}_x \cdot \tilde{\sigma}_y}, \tag{14}$$

где $\tilde{m}_{11} = \frac{1}{N}\sum_{i=1}^{3}\sum_{j=0}^{3}x_i y_j m_{ij}$ – выборочный начальный момент первого порядка.

Описанная методика реализована в виде компьютерной программы, написанной Прокофьевым Д.О. на языке Object Pascal / Delphi. Входными данными программы является массив чисел, содержащий тестовые баллы по 100 балльной шкале поступивших на первый курс студентов и их оценки по дисциплине в первом семестровом экзамене. Необходимо также задать шкалу перевода тестовых баллов в пятибалльную шкалу. Формирование таблицы 1 в программе осуществляется автоматически. Программа дополнена отдельной утилитой для набора оценок и обладает простым графическим интерфейсом, что позволяет легко отображать графики стохастической зависимости между исследуемыми величинами.

Для анализа степени влияния ЕГЭ на оценки по математике в первом семестре были выбраны результаты 300 студентов первого курса очной формы обучения (по 50 человек с каждого из 6 факультетов), поступивших в ППИ в 2009 году и 50 студентов Псковского филиала Санкт-Петербургского университета статистики и экономики.

В 2009 году распоряжением Федеральной службы по надзору в сфере образования и науки минимальное количество баллов ЕГЭ по математике, подтверждающее освоение выпускником основных программ среднего общего образования, установлено на уровне 21 балла. Рекомендации по переводу тестовых баллов ЕГЭ в пятибалльную шкалу отсутствуют. Поэтому воспользуемся рекомендациями 2008 года и примем следующую шкалу перевода.

Таблица 2

Интервал тестовых баллов	21-46	47-64	65-100
Пятибалльная шкала	3	4	5

В таблице 3 представлены результаты этого анализа: уравнения линейной регрессии, средний балл на экзамене и коэффициенты детерминации по каждому факультету и по ППИ в целом.

Таблица 3

№	Факультет (вуз)	Уравнение регрессии	Средний балл в сессию	Коэффициент детерминации
1	ИСФ	$y = 1{,}054x - 1{,}213$	3,28	0,199
2	ММФ	$y = 0{,}273x + 0{,}227$	2,62	0,056
3	ФАИ	$y = 1{,}029x - 1{,}757$	3,36	0,1152
4	ФУЭ	$y = 1{,}179x - 1{,}477$	3,56	0,226
5	ФЭФ	$y = 1{,}235x - 1{,}535$	3,46	0,505
6	ЭлМФ	$y = 1{,}011x - 1{,}211$	3,16	0,148
7	ППИ	$y = 1{,}083x - 1{,}313$	3,24	0,228
8	СПбГУСЭ	$y = 1{,}290x - 1{,}604$	3,50	0,411

Как следует из таблицы 3 между тестовыми баллами на ЕГЭ и оценками по математике в первую экзаменационную сессию существует прямая пропорциональная зависимость: с ростом числа тестовых баллов ЕГЭ по математике растёт и средний балл в первую сессию. Коэффициент детерминации в исследованных потоках изменяется от 5,6% до 50%. Довольно сильное отличие в значениях коэффициента детерминации на технических и экономических факультетах ППИ, по-видимому, можно объяснить более высокими требованиями к математической подготовке студентов и сложностью программы на технических факультетах. Очевидно, что не последнюю роль здесь играет и более высокий проходной балл на экономических специальностях. В среднем в Псковском политехническом институте примерно на 23% разброс в оценках можно объяснить различием в подготовке по элементарной математике. Этот показатель за десятилетие, прошедшее со времени исследования [1], практически не изменился, что может свидетельствовать о том, что качество отбора поступающих в вузы с введением обязательных ЕГЭ, по крайней мере, не ухудшилось по сравнению с тем, когда этот отбор осуществлялся по итогам вступительных экзаменов.

В заключение следует отметить, что рассмотренная методика может быть достаточно легко распространена на исследование влияния группы ЕГЭ на успеваемость студентов, а также исследование процесса обучения по конкретной дисциплине в ходе нескольких семестров.

ЛИТЕРАТУРА

1. Хватцев А.А., Цветкова Т.С. Методика оценивания результатов обучения математики в вузе / А.А. Хватцев, Т.С. Цветкова // Труды Псковского политехнического института. – 2001. – № 5: Естествознание и математика. Гуманитарные науки. – С. 20-25.
2. Герасименко П.В. Анализ степени влияния основных факторов на результаты обучения высшей математике в современных условиях / П.В. Герасименко // Математика в вузе. Современные интеллектуальные технологии: Материалы международной научно-методической конференции. 21 – 25 июня 2000 г. / НовГУ им. Ярослава Мудрого. Великий Новгород, 2000. – С. 7-9.
3. Хватцев А.А. Математическая статистика : учеб. пособие/ А.А. Хватцев. – Псков : ППИ, 2005. – 55 с.