

В.Е. Бахрушин, С.В. Журавель, М.А. Игнахина

Эмпирические функции распределения результатов тестирования выпускников школ

Показано, что результаты единого государственного экзамена выпускников школ Российской Федерации по русскому языку и математике могут быть описаны как смесь трех нормально распределенных выборок. Рекомендовано использование непараметрических статистических методов для обработки результатов тестирования и построения шкал перевода тестовых баллов в оценки.

It is shown that the results of the united state examination of Russian Federation secondary school graduates in Russian language and mathematics may be described as a mixture of three normally distributed components. Thus non-parametric statistical methods are recommended for test results processing and building of scales for transformation of testing results into marks.

Показано, що результати єдиного державного екзамену випускників шкіл Російської Федерації з російської мови та математики можуть бути подані як суміші трьох нормально розподілених вибірок. У зв'язку з цим рекомендується застосування непараметричних статистичних методів для обробки результатів тестування й побудови шкал переведення тестових балів в оцінки.

Введение. Вопрос усовершенствования методов и алгоритмов обработки результатов тестирования актуален в связи с переходом Украины к проведению единого независимого тестирования выпускников общеобразовательных школ, внедрением кредитно-модульной системы организации учебного процесса в высших учебных заведениях, а также активной разработкой средств дистанционного и электронного обучения.

Теоретические основы обработки результатов тестирования рассмотрены в работах В.С. Аванесова и других авторов [1, 2]. Однако вопрос об устойчивости используемых методик к отклонениям результатов от нормального закона распределения исследован недостаточно.

Известно [3, 4], что отклонения данных от нормального закона распределения во многих случаях ведут к неправомерности использования традиционных методов обработки и анализа данных, а также некоторых параметров, применяемых для их описания. В частности, для одномодальных распределений с коэффициентом эксцесса ниже $-0,6$ оценка центра распределения – центр размаха, а для распределений с коэффициентом эксцесса выше $0,8$ – медиана [4]. Согласно тому же источнику, объем выборки, необходимый для определения показателей разброса относительно центра, резко возрастает при увеличении коэффициента эксцесса. Если для нормального распределения ($\varepsilon = 0$)

по обеспечению относительного значения стандартного отклонения не более $0,05$ требуется объем выборки не менее 200 элементов, то при $\varepsilon = 5$ необходимый объем выборки возрастает до 700, а при $\varepsilon = 10$ – до 1200. Более эффективными оценками разброса относительно центра для распределений с положительными значениями коэффициента эксцесса являются энтропийный коэффициент, а при ε больше 17–18 – квантильные оценки.

Выбор конкретных методов проверки статистических гипотез различного типа также в значительной мере определяется видом функции распределения. Если закон распределения неизвестен, то используют более устойчивые к различиям в законах распределения непараметрические методы. Однако и в этом случае эффективность оценок может быть повышена, если закон распределения анализируемых данных известен. В связи с этим актуальна проблема исследования эмпирических функций распределения результатов тестирования. В работе [5] показано, что результаты внешнего независимого оценивания знаний выпускников украинских школ 2006 г. по украинскому языку и математике находятся в соответствии с гипотезой об однородном логнормальном распределении.

Цель статьи – идентификация закона распределения результатов тестирования на более представительной выборке результатов единого государственного экзамена 2008 г. для вы-

пускников средних общеобразовательных учреждений Российской Федерации.

Исходные данные и методика исследования

В качестве исходных данных с сайта [6] взяты результаты единого государственного экзамена по русскому языку и математике в Российской Федерации, проведенных в 2008 г. В тестировании по русскому языку принимало участие 1 055 842 выпускника, а по математике – 937 652.

Эмпирические функции распределения строили в соответствии с методикой [7]. Исходные данные приведены по 60-балльной шкале для результатов по русскому языку и по 37-балльной – по математике. Кроме того, использовали непрерывную аппроксимацию реального закона распределения, для которой значения модельной функции распределения совпадали с наблюдаемыми при целых значениях аргумента.

Как модели рассматривались однородные функции нормального распределения и смеси нормально распределенных выборок. Для уточнения параметров моделей минимизировали суммы квадратов их отклонений от эмпирических функций распределения, а также расчетные значения критериев χ^2 и Колмогорова-Смирнова, используя процедуру «Поиск решения» электронных таблиц *MS Excel*. При этом выполняли предварительную проверку устойчивости получаемого результата к выбору начального приближения. Для проверки гипотезы о соответствии используемой модели эмпирической функции распределения использовали критерии Колмогорова-Смирнова и χ^2 . При этом, следуя [3], принимали, что при определении параметров эмпирического распределения по выборке критическое значение критерия Колмогорова-Смирнова для уровня значимости 0,05 равно 0,895.

Эмпирические функции распределения результатов тестирования

Эмпирические и модельные функции распределения результатов тестирования по русскому языку и математике представлены на

рис. 1 и 2. Модели представляли собой однородные функции нормального распределения и смеси нормально распределенных выборок с параметрами, приведенными ниже.

Как видно из представленных данных, эмпирические функции распределения соответствуют гипотезе об однородном нормальном распределении. Расчетные значения критерия Колмогорова-Смирнова для нормального закона распределения рассмотренных выборок составляют соответственно 0,189 и 0,136. В обоих случаях они значительно меньше критического значения.

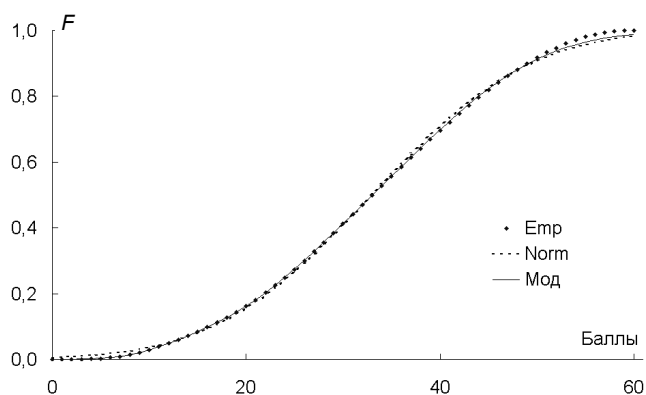


Рис. 1. Эмпирическая и модельные функции распределения результатов единого государственного экзамена по русскому языку

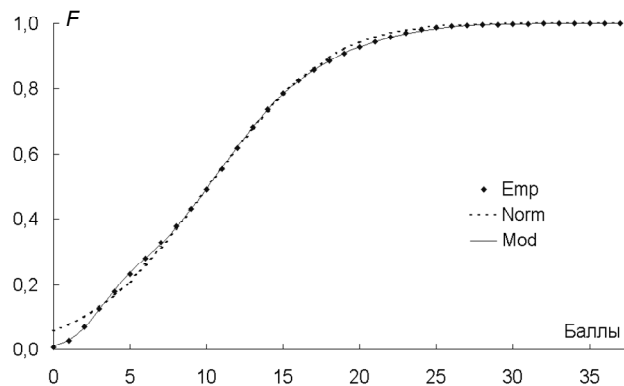


Рис. 2. Эмпирическая и модельные функции распределения результатов единого государственного экзамена по математике

Таким образом, результаты проверки по критерию Колмогорова-Смирнова не дают основания для отклонения гипотезы об однородном нормальном распределении рассматриваемых выборок.

Гистограммы распределения

Для дополнительного анализа рассматриваемых выборок построены их гистограммы (рис. 3 и 4) и проведена проверка гипотез о законе распределения на основании критерия χ^2 . Этот критерий считается менее мощным, чем критерий Колмогорова-Смирнова, и поэтому не рекомендуется для выборок числовых данных [3]. Вместе с тем приведенные гистограммы наглядно свидетельствуют о том, что рассматриваемые выборки не однородны и не могут быть описаны нормальным законом распределения. Расчетные значения критерия χ^2 для такой гипотезы значительно превышают критические, а распределение результатов по математике вообще не является унимодальным.

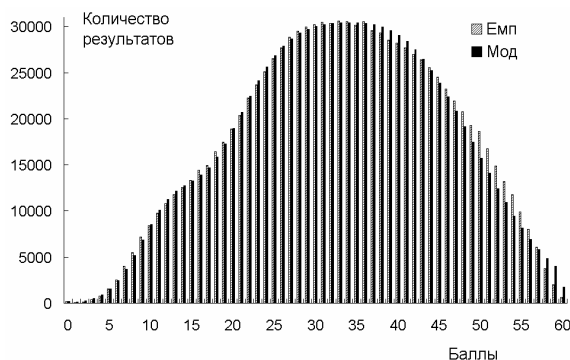


Рис. 3. Гистограмма распределения результатов единого государственного экзамена по русскому языку

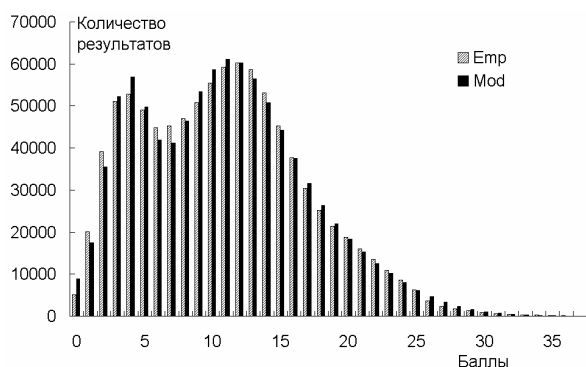


Рис. 4. Гистограмма распределения результатов единого государственного экзамена по математике

В связи с этим рассмотрена гипотеза о возможности представления имеющихся данных как смесей нормально распределенных выборок. Оптимальные с точки зрения сочетания количества компонент смеси и правильности

отображения функций и гистограмм распределения получены для случая трех компонент. Параметры компонент уточняли, минимизируя расчетные значения критериев χ^2 и Колмогорова-Смирнова. При этом получены модели функций распределения:

$$F_{\text{Ряз}} = 0,061N(11,92; 3,59) + 0,278N(24,23; 6,89) + 0,661N(39,32; 9,98);$$

$$F_{\text{Матем}} = 0,204N(3,49; 1,63) + 0,507N(10,60; 3,82) + 0,288N(16,54; 5,57).$$

Такие результаты ожидаемы, поскольку хорошо известны различия в уровне подготовки школьников между регионами, а также между городскими и сельскими школами, «рядовыми» и «элитными» образовательными учреждениями и т.д. Для результатов внешнего независимого тестирования выпускников украинских школ аналогичный разброс уровня подготовки проявился [5] в значительной асимметрии распределений результатов.

Таким образом, суммируя приведенные данные и результаты [5], можно увидеть, что реальные распределения результатов тестирования для выборок большого объема в большинстве случаев не описываются моделями однородного нормального распределения.

Закключение. Результаты исследования эмпирических функций распределения итогов тестирования выпускников средних школ в Украине и России позволяют сделать такие выводы.

Распределение результатов тестирования в общем случае может отклоняться от нормального закона и быть неоднородным, а получаемые выборки результатов могут характеризоваться высокими значениями коэффициентов асимметрии и эксцесса. В связи с этим при анализе результатов тестирования и построении шкал перевода их в оценки целесообразно использовать непараметрические статистические методы, устойчивые к типу распределения данных. Необходимо также более детально изучить вопрос о выборе стандартных показателей центра распределения результатов тестирования и разброса данных относительно центра.

Окончание на стр. 91

В настоящее время актуально дальнейшее накопление эмпирических данных о функциях распределения результатов тестирования и влиянии на них различных факторов.

1. *Аванесов В.С.* Научные основы тестового контроля знаний. – М.: Исследовательский центр, 1994. – 135 с.
2. *Аванесов В.С.* Композиция тестовых заданий. – М.: Адепт, 1998. – 217 с.
3. *Орлов А.И.* Прикладная статистика. – М.: Экзамен, 2006. – 671 с.
4. *Новицкий П.В., Зограф И.А.* Оценка погрешностей результатов измерений. – Л.: Энергоатомиздат, 1991. – 304 с.
5. *Бахрушин В.Е., Игнахина М.А., Шумада Р.Я.* Эмпирические функции распределения результатов тестирования // Зб. пр. III Міжнар. конф. «Нові інформаційні технології в освіті для всіх: система електронної освіти» / Ред. В. Гриценко. – К.: МННЦІТ та С, 2008. – С. 79–84.
6. <http://www.ege.ru>
7. *Бахрушин В.Є.* Аналіз даних. – Запоріжжя: ГУ «ЗІДМУ», 2006. – 128 с.

© В.Е. Бахрушин, С.В. Журавель, М.А. Игнахина, 2009