

**Я. О. Дунаева**

*Доцент кафедры информационных технологий в управлении  
Байкальская международная бизнес-школа  
Иркутского государственного университета*

## **АНАЛИЗ ДАННЫХ О РЕЗУЛЬТАТАХ СДАЧИ ТЕСТА TOEFL СТУДЕНТАМИ СИБИРСКО-АМЕРИКАНСКОГО ФАКУЛЬТЕТА МЕНЕДЖМЕНТА БАЙКАЛЬСКОЙ МЕЖДУНАРОДНОЙ БИЗНЕС-ШКОЛЫ ИРКУТСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА**

**Аннотация.** Рассматриваются результаты сдачи теста TOEFL студентами Сибирско-американского факультета менеджмента БМБШ ИГУ (САФ), статистическими методами исследуются зависимости результатов TOEFL от базовой подготовки и количества попыток сдать тест, а также выявляются отклонения результатов от ожидаемых значений.

**Ключевые слова:** анализ данных, языковая подготовка, результаты TOEFL, статистические методы исследования, однофакторный дисперсионный анализ.

Современное высшее образование является высоко конкурентной средой, поэтому мониторинг и глубокое понимание сути происходящих в нем внутренних процессов чрезвычайно важно для эффективного управления учебным процессом. Пониманию этому способствует анализ данных, накопленных в системе дифференцированного Интернет-обучения «Гекадем» и предоставляемых пользователям системы, обладающим соответствующими полномочиями. Основная часть отчетов, предоставляющих такую информацию, сосредоточена в подсистемах «Деканат» и «Директор».

В качестве исходные данных для выполненного в настоящей работе исследования использовались результаты SQL-запросов к базе данных системы «Гекадем», разработанных сотрудниками кафедры информационных технологий в управлении и лаборатории информационных технологий БМБШ для подсистем «Деканат» и «Директор», а также результаты дополнительных запросов, выполнявшихся специально для настоящего исследования. В целом преобразование форматов данных пока не автоматизировано и выполнялось вручную или с применением специально разработанных для этой цели небольших программ. В качестве инструментальных средств анализа данных использовались MS Excel и SAS.

Показателями, традиционно интересующими администрацию и преподавателей БМБШ, являются результаты сдачи студентами Сибирско-американского факультета (САФ) международного теста TOEFL (Test of English as a Foreign Language). Интенсивная подготовка студентов к сдаче TOEFL является одним из основных направлений деятельности кафедры английского языка и результаты, демонстрируемые студентами во время тестирования, могут служить показателями эффек-

тивности работы преподавателей кафедры, индикаторами интереса студентов к изучению английского языка, а также характеристикой базовой языковой подготовки студентов в средних учебных заведениях. Кроме того, от этих результатов зависит количество студентов САФ, поступающих на международные программы: бакалавриата – российско-американскую и российско-австралийскую.

Исходные данные представляют собой сведения о сдаче TOEFL студентами программ САФ-2006, САФ-2007, САФ-2008, САФ-2009, САФ-2010: о количестве попыток, баллах TOEFL и средних учебных заведениях, выпускниками которых являются студенты перечисленных программ. В рассматриваемые 5 лет студенты, сдававшие TOEFL с первой попытки, были выпускниками нескольких школ, лицеев и гимназий, преимущественно Иркутских (рис. 1). Полная картина (с учетом средних учебных заведений, только один выпускник которых сдал TOEFL с первой попытки) представлена на рис. 2. Представленная на рис. 2 диаграмма может привести к выводу о лучшей языковой подготовке в школах, гимназиях и лицеях Иркутска, но, учитывая, что суммарное количество иркутян на программах в рассматриваемом периоде составляло 63,75 % от общего количества студентов, такой вывод не обоснован.

Учитывая соответствующие данные по студентам из других городов (табл. 1), можно сказать, что в рассматриваемый период лучших результатов TOEFL добились выпускники учебных заведений Заларей, Улан-Удэ, Усть-Кута и Усолья-Сибирского.

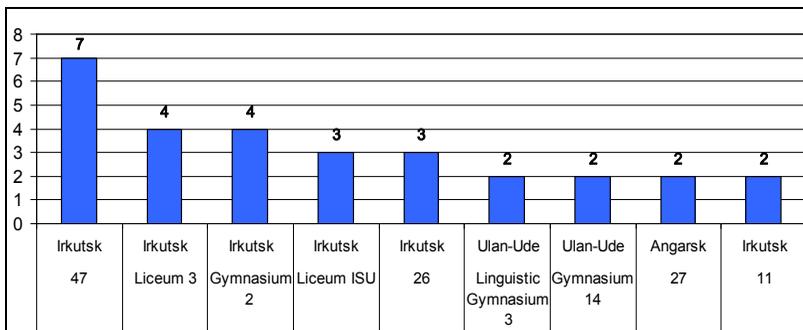


Рис. 1. Количество студентов, сдавших TOEFL с первой попытки за период с декабря 2005 по декабрь 2010 г. (в разрезе средних учебных заведений)

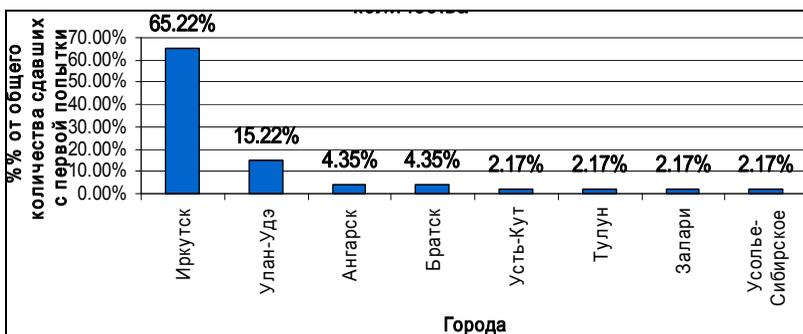


Рис. 2. Процент выпускников средних учебных заведений, сдавших TOEFL с первой попытки за период с декабря 2005 по декабрь 2010 г., относительно общего количества сдавших TOEFL в этот период с первой попытки

Таблица 1

СРАВНИТЕЛЬНЫЕ ДАННЫЕ О СДАВШИХ TOEFL СТУДЕНТАХ С ПЕРВОЙ ПОПЫТКИ И ПРОЦЕНТОМ СОСТАВЕ СТУДЕНТОВ ИЗ НЕСКОЛЬКИХ ГОРОДОВ

	% в общем кол-ве студ. на программах	% сдавших с 1-й попытки	% сдавших с 1-й попытки к общему кол-ву студ. на программах
Иркутск	63,75	65,22	1,02
Улан-Удэ	6,25	15,22	2,44
Ангарск	7,50	4,35	0,58
Братск	4,17	4,35	1,04
Усть-Кут	1,25	2,17	1,74
Тулун	2,50	2,17	0,87
Залари	0,42	2,17	5,21
Усолье-Сибирское	1,25	2,17	1,74

Анализ данных о количестве попыток, предпринимавшихся студентами, сдавшими TOEFL в период с декабря 2005 по декабрь 2010 г., показал, что практически все студенты, сдавшие тест в рассматриваемом периоде, сделали это с одной, двух, трех или четырех попыток. В процентном отношении больше всего студентов, сдавших TOEFL со второй попытки (46,49 %) и первой попытки (40,35 %).

Рассмотрим ситуацию со сдачей TOEFL в период с декабря 2005 г. по декабрь 2010 г. в целом. Согласно данным, полученным из базы данных системы «Гекадем», общее количество студентов, сдававших тест в этот период (с любым количеством попыток) составило 919 человек. Из них сдавших (получивших не менее 550 баллов) было 217 человек (23,61 %), чуть менее одной четвертой, а потерпевших неудачу (на-

бравших в каждой из попыток менее 550 баллов) – 702 человека (76,39 %), чуть более трех четвертей.

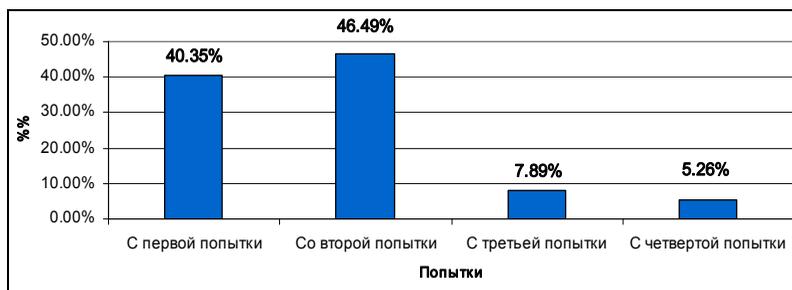


Рис. 3. Процент студентов САФ, сдавших TOEFL с одной, двух, трех и четырех попыток за период с декабря 2005 по декабрь 2010 г., относительно общего количества сдавших TOEFL в этот период

И, наконец, рассмотрим прогресс, который продемонстрировали студенты нескольких программ САФ при нескольких попытках сдачи TOEFL в период декабрь 2005 – декабрь 2010 г. (размерность случайной выборки равна 100).

Таблица 2  
СРЕДНЕЕ УВЕЛИЧЕНИЕ РЕЗУЛЬТАТОВ СДАЧИ TOEFL  
В ПЕРИОД С ДЕКАБРЯ 2005 Г. ПО ДЕКАБРЬ 2010 Г.

Среднее увеличение результата в баллах (прогресс) между попытками		
1–2	2–3	3–4
34,63	43,95	46

В рассматриваемый период имели место два случая, в которых прогресс студентов составил более 100 баллов. Причиной столь существенного увеличения результатов, с одной стороны, может быть выдающееся усердие студента, с другой стороны – скрытое применение технических средств во время теста, что, к сожалению, весьма вероятно в некоторых случаях. Принимая во внимание дополнительные факторы, в частности, имевший место ранее персональный прогресс конкретного студента (из числа продемонстрировавших выдающийся прогресс в одной из попыток), увидим, что сомнения в объективности результатов обоснованы, и недобросовестность студента была весьма вероятна. Конкретный пример: 443, 487, 613, т. е. прогресс второй попытки в сравнении с первой был равен 44, а третьей в сравнении со второй – 126, почти в три раза больше. Учитывая приводимые ниже (табл. 3) характе-

ристики выборки, это более чем в 2 раза превышает стандартное отклонение и вызывает подозрения.

Рассмотрим конечные результаты TOEFL для рассматриваемой выборки. Характеристики выборки представлены в таблице 3. Частота попадания результатов в классы, количество которых равно 7, визуально представлена на рис. 4, согласно которому распределение частот – левостороннее, несимметричное. Несимметричность распределения частот означает, что распределение генеральной совокупности отличается от нормального и для выборки более представительно не среднее, а медиана, которая равна 557. Число классов вычислено по формуле Стерджеса [2].

Таблица 3

ХАРАКТЕРИСТИКИ СЛУЧАЙНОЙ ВЫБОРКИ КОНЕЧНЫХ РЕЗУЛЬТАТОВ TOEFL ЗА 5 ЛЕТ (2005–2010)

Среднее	541,84
Стандартное отклонение	50,98
Медиана	557
Xmin	423
Xmax	623
R	200
Число классов	7
Ширина класса	28,57

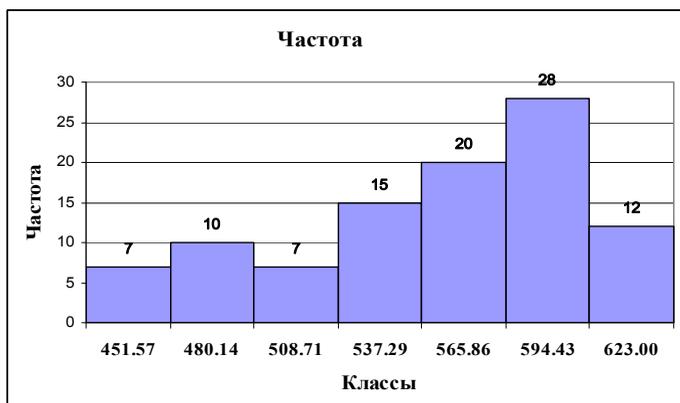


Рис. 4. Распределение частот попадания результатов TOEFL в 7 классов

Для поиска ответа на этот вопрос, имели ли место существенные отличия между студентами разных курсов САФ (наборов) с точки зрения результатов TOEFL, применим однофакторный дисперсионный анализ. Проведем анализ результатов прохождения первого теста для разных периодов времени: для программ САФ–1992 – САФ–1996, т. е. для пяти групп.

Сформулированы гипотезы:

$H_0$ : фактор года поступления не является существенным (наборы не отличаются друг от друга с точки зрения результатов TOEFL);

$H_a$ : фактор года поступления является существенным (наборы отличаются друг от друга с точки зрения результатов TOEFL).

Результаты дисперсионного анализа результатов сдачи TOEFL студентами программ САФ–1992 – САФ–1996 (табл. 4).

Таблица 4

РЕЗУЛЬТАТЫ ДИСПЕРСИОННОГО АНАЛИЗА (ANOVA) ПРОВЕРКИ  
НАЛИЧИЯ ОТЛИЧИЙ МЕЖДУ ПЯТЬЮ НАБОРАМИ НА САФ  
(1992–1996 ГГ.) С ТОЧКИ ЗРЕНИЯ ПОДГОТОВКИ СТУДЕНТОВ  
ПО АНГЛИЙСКОМУ ЯЗЫКУ

Источник вариации	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>P</i> -значение	<i>F</i> критическое
Между группами	39 217,7	4	9804,4	3,904	0,0051	2,447
Внутри групп	301 383,2	120	2511,5			
Итого	340 600,9	124				

Для второго примера  $F$  равно 3,9 и это значение больше  $F$  критического, равного 2,45, следовательно, нулевая гипотеза не отвергается, т. е. анализ не выявил разницы между группами с точки зрения первичных результатов TOEFL.

Данное исследование показало, что статистические методы анализа данных позволяют выявить основные закономерности, тенденции и отклонения в исследуемых данных. Целесообразно периодически анализировать данные о сдаче TOEFL, во-первых, с целью оценки результативности языковой подготовки студентов в целом и прогресса в сдаче TOEFL в случае неуспеха в первой попытке, во-вторых, выявления специфических отклонений, свидетельствующих о влиянии на результаты тестов неких сторонних факторов.

### Список литературы

1. URL: <http://www.toefl.ru/>.
2. URL: <http://math.immf.ru/lections/203.html>.