СТАТИСТИЧЕСКИЙ АНАЛИЗ РЕЗУЛЬТАТОВ КОМПЬЮТЕРНОГО ТЕСТИРОВАНИЯ В КРЕДИТНОЙ СИСТЕМЕ ОБУЧЕНИЯ

И.Н. Куринин, В.И. Нардюжев, И.В. Нардюжев

Кафедра компьютерных технологий Российский университет дружбы народов ул. Миклухо-Маклая, 6, Москва, Россия, 117198

В статье описана программа для статистического анализа результатов компьютерного тестирования. Продемонстрированы примеры ее использования для оценки учебных достижений студентов в кредитной системе обучения. Рассмотрена оценка качества тестов.

Ключевые слова: студент, достижения, оценка, компьютерное тестирование, статистический анализ, оценка качества.

Результаты компьютерного тестирования определяют учебные достижения студентов и позволяют судить о качестве учебного процесса. В такой ситуации для преподавателей принципиально важными становятся оперативный анализ и правильная интерпретация результатов тестирования, как и оценка качества тестовых материалов для корректного использования их в учебной работе. Решить стоящие перед преподавателем задачи помогает разработанная авторами технология оперативного статистического анализа результатов компьютерного тестирования [1]. Она базируется на функциональных возможностях программы «СтатИнфо» [2]. Эффективность и надежность технологии проверена многолетним опытом ее применения для оценки учебных достижений студентов в кредитной системе обучения [3].

Программа «СтатИнфо» работает под управлением операционной системы MS Windows. Информационной основой этой программы является табличная (реляционная) база данных. База компактная. Для хранения результатов 1000 тестируемых требуются всего 0,5 Мбайта. Накопленный в базе материал можно использовать для научной и практической работы в области компьютерного тестирования и мониторинга качества образования. Допускается локальная или сетевая установка программы.

Программа «СтатИнфо» позволяет преподавателю выбрать объект статистической обработки (группа, факультет и т.п.) и получить для этого объекта количество тестируемых, их распределение по оценкам с учетом шкалы перевода тестовых баллов в оценки, распределение тестируемых по количеству набранных баллов, процент правильных ответов на вопрос теста. По полученным цифровым данным можно построить графики, провести сравнительный анализ результатов тестирования различных объектов статобработки, посмотреть списки тестируемых с возможностью произвольной сортировки и быстрого поиска, распечатать статистические формы и списки (или их части) на принтере. При необходимости любую статистическую форму или список можно сохранить в отдельном текстовом файле или загрузить в электронные таблицы для построения необходимых диаграмм и таблиц.

Рабочий экран программы «СтатИнфо» (рис. 1) содержит заголовок, меню, панель инструментов, строку состояния. По заголовку можно определить, какая база используется в данный момент. При работе с меню в строке состояния отображается назначение выбранного в данный момент пункта. Пункты меню разбиты на четыре функциональные группы: сводка — получение, открытие, сохранение сводок, загрузка результатов; режим — установка параметров текущей сводки; окно — размещение окон на экране и переход к указанному окну; ? — получение справочных сведений о программе «СтатИнфо».



Рис. 1. Рабочий экран программы «СтатИнфо»

Кнопки панели инструментов повторяют соответствующие пункты меню и используются для быстрого выполнения следующих типовых операций: получение новой сводки; открытие сводки, сохраненной в файле статистики; сохранение текущей сводки в файле статистики; переход в режим формы; выбор интересующей статистики; настройка шкалы оценок; сохранение выбранной формы в текстовом файле (для передачи в Excel); переход в режим списка; выбор интересующих столбцов (реквизитов тестируемых); сортировка списка; поиск информации в списке; переход в режим отчета; размещение окон сверху вниз без перекрытия; размещение окон слева направо без перекрытия; загрузка протоколов компьютерного тестирования в базу данных; выход из программы «СтатИнфо».

Чтобы загрузить в базу данных протоколы компьютерного тестирования, сформированные программой «Тестер» [3], необходимо воспользоваться кнопкой «Загрузка» (предварительно закрыв все сводки). На экране появится диалог «Загрузка протоколов». Для начала загрузки нажмите кнопку «Начать». В появившемся на экране стандартном диалоге «Загрузка протоколов тестирования в базу данных» выберите нужные файлы с расширением «.р». По умолчанию они хранятся в папке «Info \ Res». Выбрав файлы, нажмите кнопку «Открыть». Начнется загрузка, по мере продвижения которой будут расти полоски, отображающие процент выполненной работы, меняться имена загружаемых файлов, увеличиваться счетчик добавленных записей.

В итоге в заголовке диалога появится одна из надписей: «Загрузка завершена» — успешное завершение всей операции; «Загрузка отменена» — пользователь отменил загрузку до ее завершения; «Загрузка прервана» — возникла ошибка, на-

пример, при попытке пользователя повторно загрузить подгруженный ранее файл. В любом случае программа покажет количество успешно загруженных файлов и добавленных записей. Нажмите кнопку «Закрыть», чтобы закрыть диалог.

Перечень полей в табличной базе данных с результатами компьютерного тестирования доступен для формирования запросов (рис. 2). Существенной особенностью программы является «гибкость» возможного запроса. Диалог «Запрос» содержит таблицу фильтров, в которой для каждого столбца (поля) информационной базы данных можно указать диапазоны допустимых в запросе значений.

Запрос		×
Столбцы :	Фильтры :	Ключ:
Пункт		Пчнкт
<u>Участник</u>		
Факультет		
Группа		2
Ф.И.О.		Значения:
Предмет	-03 09-11 15 51-	один итог 👤
Тест		
Дата		
Балл	15-	П
Процент		Предметы:
Ответ		без итога 💌
Проверка		
Статус		
Протокол		
Штамп		
		<u>П</u> ринять
		<u>О</u> тменить

Рис. 2. Формирование запроса в программе «СтатИнфо»

При формировании запросов каждому полю таблицы можно назначить фильтр, указав диапазон допустимых значений. Например, установив для поля «Предмет» фильтр «-03 09-11 15 51-», из таблицы можно отобрать все записи, содержащие такой код предмета Π , что (Π <=03) или (09<= Π <=11) или (Π =15) или (Π >=51). Используя дополнительный фильтр «15-» для поля «Балл», можно отобрать всех участников тестирования по указанным предметам, получивших не менее 15 баллов. Или, задав дополнительный фильтр «90» для поля «Балл», можно отобрать всех учащихся, прошедших тестирование по указанным предметам и получивших 90 баллов.

В выпадающем списке «Ключ» выбирается поле таблицы, по значениям которого будут группироваться записи о тестируемых.

Выпадающие списки «Значения» и «Предметы» задают режимы подведения итогов по ключевому полю и по предметам: «один итог» (показывать без распределения), «с итогом» (показывать как итог, так и распределение), «без итога» (показывать только распределение, без итога).

По кнопке «Принять» программа «СтатИнфо» принимает сформированный запрос и выдает на экран сводку с результатами (рис. 3).

	Полученные баллы (707 чел.)												
												7	
	1												
1	Итог		0	1	2	3	4	5	6	7	8	9	
	105 ИТ (информат	0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
		1	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,4	0,0	
		2	0,0	0,3	0,3	0,0	0,1	0,1	0,7	0,1	1,1	0,4	
		3	0,8	1,1	1,1	0,1	1,8	1,0	0,8	1,4	2,1	1,1	
		4	2,0	0,7	3,3	1,1	2,0	1,3	2,4	1,4	4,0	1,8	
		5	3,1	1,8	3,4	2,5	2,5	3,3	4,1	2,5	4,1	2,0	
		6	2,8	2,3	3,7	2,3	1,7	3,1	2,3	2,1	1,7	2,1	
		7	2,0	1,0	1,1	1,1	1,3	0,8	0,7	0,4	1,4	0,3	
		8	0,3	0,0	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0	
		9	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	
		10	0,0										
						.ол-во		Наим.		Сред.		Наиб.	
						707		16,0		53,8		92,0	4

Рис. 3. Сводка с результатами в программе «СтатИнфо»

Для полученной сводки по кнопке «Статистика» открывается диалог (рис. 4), в котором можно выбрать интересующую статистическую форму. Возможны следующие варианты: 1) количество тестируемых; 2) поставленные оценки; 3) полученные баллы; 4) правильные ответы.

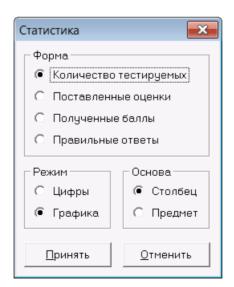


Рис. 4. Рабочий экран «Статистика» программы «СтатИнфо»

Любую форму можно посмотреть на экране (а потом напечатать на принтере) в цифрах или в графике, меняя положение переключателя «Режим». Переключатель «Основа» определяет, значение какого из двух параметров (ключевого столбца или предмета) остается в форме неизменным по горизонтали. Выбор в каждом конкретном случае определяется удобством представления информации на экране или в распечатке на принтере.

Для полученной сводки по кнопке «Шкала оценок» есть возможность настроить шкалу перевода тестовых баллов в оценки. Подстройка шкалы повлияет на вид статистической формы «Поставленные оценки» и поможет учесть специфику или потребности конкретного вуза. При необходимости сформируйте новый запрос в базу данных с результатами компьютерного тестирования и нажмите кнопку «Шкала оценок», чтобы открыть одноименный диалог и настроить параметры предмета, указав напротив кода его название, количество заданий, верхние границы для оценок по 100-балльной шкале. Важное место в системе занимает работа со списками тестируемых. Для переключения на работу со списками тестируемых используется кнопка «Список».

В этом режиме можно выполнить следующие операции:

- 1) выбрать по кнопке «Столбцы» (рис. 5) необходимые для просмотра на экране (печати на принтере) реквизиты тестируемых и определить порядок их следования друг за другом;
- 2) по кнопке «Сортировка» (рис. 6) отсортировать список тестируемых по любой совокупности реквизитов. По умолчанию записи в списке отсортированы по фамилиям;
- 3) по кнопке «Поиск» найти в базе данных участника тестирования по любой совокупности реквизитов (по ФИО, по коду предмета, по числу набранных баллов и т.д.). Перед началом поиска список должен быть отсортирован по соответствующим реквизитам.

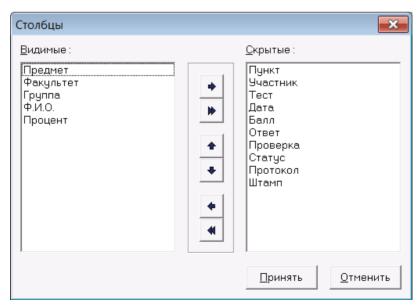


Рис. 5. Выбор столбцов (полей из базы данных) для списков

Список тестируемых (707 чел.)								
	1_		1 +40	In 1				
Предмет	Факультет		Ф.И.О.	Процент				
105 ИТ (информатика)	Гумсоц	ГМБ-15	Гадаева Ирина	69				
105 ИТ (информатика)	Гумсоц	ГНБ-21	Кирсанова Юлия	69				
105 ИТ (информатика)	Гумсоц	ГМБ-11	Осадчук Екатерина	69				
105 ИТ (информатика)	Гумсоц	ГМУ-32	Агафонов-Максимов	69				
105 ИТ (информатика)	Гумсоц	ГФБ-21	Рубаха Елизавета	69				
105 ИТ (информатика)	Гумсоц	ГФБ-21	Терещенко Диана	69				
105 ИТ (информатика)	Гумсоц	ГПБ-42	Русакова Екатерина	69				
105 ИТ (информатика)	Гумсоц	ГМБ-16	Джавлах Кирилл	69				
105 ИТ (информатика)	Гумсоц	ГМУ-34	Полухина Анна	70				
105 ИТ (информатика)	Гумсоц	ГМУ-34	Хасин Роман	70				
105 ИТ (информатика)	Гумсоц	ГМБ-11	Осадчук Екатерина	70				
105 ИТ (информатика)	Гумсоц	ГНБ-21	Потапова Анна	70				

Рис. 6. Список тестируемых в программе «СтатИнфо»

Печать отчетов. Для ввода комментария к форме или списку служит расположенный над таблицей редактор. Его первая строка формируется программой автоматически в зависимости от выбранной статистической формы или списка, а вторая и третья строки доступны для редактирования. Введенный комментарий будет присутствовать в распечатках, поэтому в нем рекомендуется отразить содержимое запроса, результатом выполнения которого является сводка. Сводка «Правильные ответы на задания теста» показана на рис. 7 (как пример).

Правильные ответы (707 чел.)													
	1												
	Итог				Наим.		Сред.		Наиб.				Кол-во
	105 ИТ (информатика)				31,1		50,3		69,0				707
		10	20	30	40	50	60	70	80	90	100	%	чел.
	1											69,0	488
	2											54,6	386
	3											67,8	479
	4											43,8	310
	5											49,6	351
	6											47,2	334
	7											53,2	376
	8											57,0	403
	9											63,5	449
	10											54,2	383

Рис. 7. Сводка «Правильные ответы на задания теста»

Для печати на принтере полученных на экране форм и списков необходимо перейти в режим отчета (кнопка «Отчет»). Окно предварительного просмотра даст представление о будущей распечатке (рис. 8). При помощи дополнительной (появляющейся только в этом режиме) панели инструментов можно установить масштаб, позволяющий увидеть всю страницу или всю страницу по ширине; установить реальный масштаб (100%-е соответствие бумаге); выбрать для предварительного просмотра любую страницу будущего отчета; установить книжную или альбомную ориентацию страницы; изменить поля страницы; вывести на принтер «по умолчанию» все страницы / только текущую / диапазон номеров.

<mark>ш</mark> Сводка №1												•	×
	1/2		4			•				<u>=</u>			
				П	рав	ильнь	ie o	тветь	ı (7	07	чел.)	•
													=
TOT				Наим.		Сред.		Наиб.			- 1	Сол-во	u
05 ИТ (информатика)		n 4		31,1		50,3	10	69,0				707	
	10	20	30	40	50	60	70	80	90	00	%	чел.	
1	2000000								11		59,0	488	
2							10		- 8		54,6	386	
3	210110110										57,8	479	
4	-				***						43,8	310	
5	21001000						- 1				49,6	351	
6	-						- 1				47.2	334	
٠7	200000					****					53,2	376	
8	-										57,0	403	
9											53,5	449	
10	-										54,2	383	
4				П									

Рис. 8. Просмотр отчета перед печатью в программе «СтатИнфо»

Хранение всей необходимой информации (результаты компьютерного тестирования) в единой реляционной (табличной) базе данных позволяет значительно ускорить процесс расчета базовых величин и дает возможность автору или тестологу быстро перейти от полученной в отчете цифры к стоящему за ней тестовому заданию.

При необходимости любую статистическую форму или список можно сохранить по кнопке «Экспорт» в отдельном текстовом файле (с разделителем ";"). Полученный файл можно, например, загрузить в виде стандартной электронной таблицы в MS Excel и использовать весь богатый набор возможностей этого пакета для расчета параметров и построения любых необходимых диаграмм или таблиц.

Для проведения расчетов в MS Excel часто требуется матрица ответов. Она представляется прямоугольной таблицей, в каждой позиции которой указываются ответы участника тестирования. Обычно номер строки соответствует номеру испытуемого, а номер столбца соответствует номеру задания теста. Для подготовки матрицы ответов в программе «СтатИнфо» нужно выбрать столбец «Ответы» (рис. 9) или столбец «Проверка» (рис. 10) и в режиме «На экспорт» сделать текстовый файл (с разделителем ";").

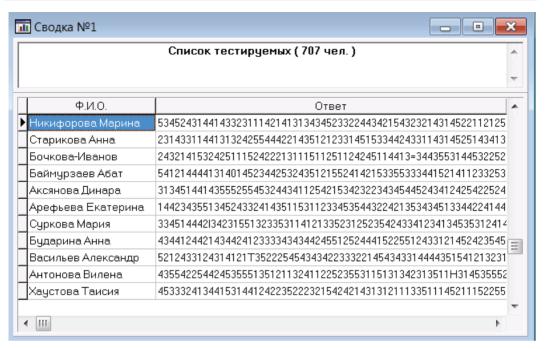


Рис. 9. Ответы тестируемых на задания теста

Список тестируемых (707 чел.)							
Ф.И.О.	Проверка						
Никифорова Марина	111011111110110100110000000111011111011010						
Старикова Анна	1111101111011001001111111000111111110101						
Бочкова-Иванов	1100100111111010010111111100000011100101						
Баймурзаев Абат	10111110110110011111110111110101011110010011010						
Аксянова Динара	1111101100101111110111100001001001011111						
Арефьева Екатерина	111001110011001111011110000010101111011100100100100100101						
Суркова Мария	1011011110111101111001111101111010110000						
Бударина Анна	111101110101010000101111110100001111111						
Васильев Александр	100100111110111111110001111101111111111						
Антонова Вилена	11101010101011011011011011001110010010110111010						
Хаустова Таисия	1111011011110101101111010101100001110000						
Турлыбеков Еркебулан	1100100001100110001100001101000110100000						

Рис. 10. Результаты проверки ответов тестируемых на задания теста

Для статистического анализа качества тестовых заданий обычно используются следующие параметры:

— средний балл, который определяется средним значением выборки с учетом баллов всех тестируемых и общего числа участников тестирования;

- стандарт, определяемый через отклонение, показывающее насколько далеко данное значение результата отстоит от среднего значения;
- коэффициент вариации, определяемый через отношение отклонения к среднему значению балла;
- трудность задания, вычисляемая как процент тестируемых выполнивших верно данное задание. На практике, когда нужно разделить тестируемых по уровню их подготовки, для теста отбираются задания с трудностью в интервале от 40% ло 60%.
- точечно-биссериальная корреляция, которая позволяет оценить как связано выполнение данного задания с выполнением всего теста. Поэтому мера соответствия между успешностью решения одного задания и всего теста по выборке тестируемых является показателем пригодности задания для теста.
- дифференцирующая способность задания. Рассчитывается как разность двух значений: трудности задания для сильной группы тестируемых и трудности задания для слабой группы тестируемых. Чем выше дифференцирующая способность задания, тем лучше оно разделяет тестируемых по уровню подготовки.
- статистическая независимость заданий в тесте, определяемая выборочным коэффициентом корреляции по двумерной выборке.

Большое внимание в проводимом статистическом анализе качества тестовых заданий уделяется дистракторному анализу. Анализ работы дистракторов в задании используется для оценки как «работает» каждый из предложенных вариантов ответов. Для этого рассчитываются следующие показатели: количество тестиру, решавших конкретное задание, частота выбора правильного ответа на задание, частота выбора ложных ответов (дистракторов), частота выбора ответа группами слабейших и сильнейших тестируемых, дифференцирующая способность ответа, отклонение частоты выбора неправильного ответа от среднего значения частот выбора неправильных ответов, сигналы о наличии отклонений от нормы и т.д.

В табл. 1 и 2 в качестве примера показаны реальные результаты итогового компьютерного тестирования студентов филологического факультета и факультета гуманитарных и социальных наук по курсу «Информатика» за 15 семестров 2005—2012 гг.

Таблица 1
Полученные баллы (% правильных ответов на все задания теста)

Период	Колич.	Полученные баллы (% правильных ответов на все задания теста)						
	тестов		средн.	наиб.				
2005 г. — весна	190	20,0	60,2	89,0				
— осень	171	12,0	58,7	87,0				
2006 г. — весна	79	16,0	50,2	82,0				
— осень	280	11,0	54,6	87,0				
2007 г. — весна	258	15,0	47,3	80,0				
— осень	325	13,0	52,9	86,0				
2008 г. — весна	160	14,0	49,5	77,0				
— осень	190	18,0	57,0	84,0				
2009 г. — весна	707	16,0	53,8	92,0				
— осень	524	18,0	52,7	90,0				

Окончание

Период	Колич.	Полученные баллы (% правильных ответов н	а все задания теста)
	тестов	наим.	средн.	наиб.
2010 г. — весна	428	11,0	51,6	93,0
— осень	318	12,0	58,6	90,0
2011 г. — весна	253	11,0	53,1	85,0
— осень	385	18,0	54,0	92,0
2012 г. — весна	150	23,0	53,9	92,0

Таблица 2
Правильные ответы на каждое задание теста (%)

Период	Колич.	% правильн	% правильных ответов (на каждое задание теста)						
	тестов	наим.	средн.	наиб.					
2005 г. — весна	190	25,3	60,2	87,4					
— осень	171	24,6	58,4	91,2					
2006 г. — весна	79	21,5	50,2	86,1					
— осень	280	28,9	54,6	77,1					
2007 г. — весна	258	29,8	47,2	76,0					
— осень	325	32,6	52,9	79,4					
2008 г. — весна	160	28,1	49,5	84,4					
— осень	190	33,0	56,9	78,2					
2009 г. — весна	707	31,1	50,3	69,0					
— осень	524	29,4	49,5	70,0					
2010 г. — весна	428	37,6	51,6	68,2					
— осень	318	37,1	58,6	82,4					
2011 г. — весна	253	30,4	53,1	74,7					
— осень	380	39,0	54,0	75,8					
2012 г. — весна	150	25,3	50,8	74,7					

По итогам статистической обработки результатов компьютерного тестирования студентов выявляются ошибки в тестовых заданиях (длинная формулировка или отсутствие правильного ответа) и определяются хорошо или плохо освоенные студентами разделы курса.

Для иностранных студентов определяются задания с наиболее сложными для них формулировками на русском языке. Это позволяет специально (на русском языке) разработать и использовать тестовые задания, имеющие сжатые (короткие) и понятные им формулировки; обеспечить этих студентов соответствующими методическими материалами для подготовки к тестам; дать им возможность в любое время проходить репетиционное тестирование через Интернет.

ЛИТЕРАТУРА

- [1] Нардюжев В.И., Нардюжев И.В. Модели и алгоритмы информационно-вычислительной системы компьютерного тестирования: Монография. М.: Прометей, 2000.
- [2] Нардюжев В.И., Нардюжев И.В. Свидетельство РОСПАТЕНТ № 2000610068 от 27 января 2000 г. об официальной регистрации программы для ЭВМ «Оперативный статистический анализ результатов тестирования на компьютерах через Интернет (СТАТИНФО)».
- [3] Куринин И.Н., Нардюжев В.И., Нардюжев И.В. Компьютерное тестирование в оценке учебных достижений студентов: учебно-методическое пособие. М.: Изд-во РУДН, 2008.

STATISTICAL ANALYSIS OF RESULTS OF COMPUTER-BASED TESTING IN THE CREDIT SYSTEM OF EDUCATION

I.N. Kurinin, V.I. Nardyuzhev, I.V. Nardyuzhev

Chair of computer technologies Peoples' Friendship University of Russia Miklukho-Maklaya str., 6, Moscow, Russia, 117198

In this article a program for statistical analysis of results of computer-based testing is described. Some examples of its usage for assessment of students' achievements in the credit system of education are demonstrated. Consideration is given to the evaluation of the quality of the tests.

Key words: student, achievements, assessment, computer-based testing, statistical analysis, quality evaluation.