



Докладчик:

**Валентина
Сергеевна
Батасова
НИУ «МЭИ»**

Автор:

Валентина Сергеевна Батасова

***Национальный исследовательский
университет «МЭИ»***

Москва, Россия
14-17 апреля 2020 г.

Объективность результатов автоматизированного контроля знаний студентов (тестирования)

Под объективностью результатов тестирования обычно понимается соответствие этих результатов реальным знаниям студентов. Всегда остается сомнение в объективности тестирования, потому что очень трудно определить критерий объективности.

Предлагается в качестве критерия объективности результатов тестирования использовать их соответствие результатам традиционного контроля знаний студентов в форме устного экзамена и контрольных работ.

Приводятся результаты статистического анализа зависимости оценок, полученных на экзамене и контрольных работах, от оценок, полученных при тестировании.

Тестирование проводилось в системе дистанционного обучения «Прометей» НИУ «МЭИ». Расчеты были выполнены в среде MS Excel.



ИССЛЕДОВАНИЕ СТАТИСТИЧЕСКОЙ ЗАВИСИМОСТИ ЭКЗАМЕНАЦИОННЫХ ОЦЕНОК ОТ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ НА ПЕРВОМ КУРСЕ ФАКУЛЬТЕТА ЭЛЕКТРОННОЙ ТЕХНИКИ (ЭТФ) ИНСТИТУТА РАДИОТЕХНИКИ И ЭЛЕКТРОНИКИ (ИРЭ) НИУ «МЭИ»

Использовались оценки по дисциплине «Информатика» в зимнюю сессию 2018-2019 учебного года. Исследовалась зависимость оценок экзамена от результатов тестирования, проведенного в конце семестра. Рассматривалась выборка из 118 наблюдений.

Первичный анализ выборки показал:

- Результаты тестирования отличаются от результатов экзамена не более чем на 1 балл для 102 (из 118) наблюдений. Этот факт позволяет предположить наличие статистической зависимости экзаменационных оценок от оценок тестирования.
- Для 99 из 118 студентов (в 83,9 % случаев) оценка по тестированию не превосходила экзаменационную оценку. Таким образом, эксперт (преподаватель) ставил оценки менее строго, чем тест.

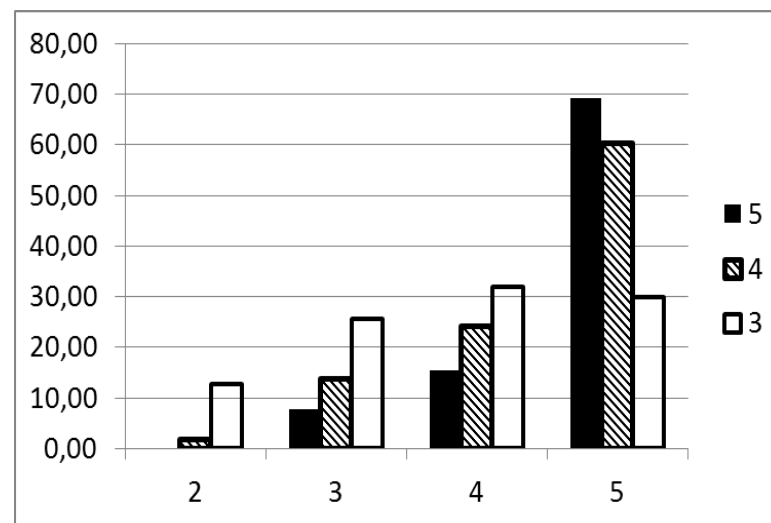
Москва, Россия
14-17 апреля 2020 г.

ЧАСТОТЫ ПОЛУЧЕНИЯ ОЦЕНОК НА ЭКЗАМЕНЕ ДЛЯ КАЖДОЙ ИЗ ДОПУСТИМЫХ ОЦЕНОК ТЕСТИРОВАНИЯ ПРЕДСТАВЛЕНЫ В ТАБЛИЦЕ И НА ГИСТОГРАММЕ:

Оценки		Экзамен			
		2	3	4	5
Т е с т	5	0,00	7,69	15,38	69,23
	4	1,72	13,79	24,14	60,34
	3	12,77	25,53	31,91	29,79

Выводы:

1. Более 68 % студентов, сдавших тест на «5», получили «5» на экзамене.
2. Более 84 % студентов, сдавших тест на «5», получили «4» или «5» на экзамене.
3. Более 84 % студентов, сдавших тест на «4», получили «4» или «5» на экзамене.
4. Более 87 % студентов, получивших «3» за тест, сдали экзамен на положительную оценку.



Гистограмма частот экзаменационных оценок. В легенде справа показаны оценки тестирования, по горизонтальной оси – экзаменационные оценки.

СРЕДНЯЯ ОЦЕНКА ЭКЗАМЕНА (y) ДЛЯ РАЗЛИЧНЫХ ОЦЕНОК ТЕСТИРОВАНИЯ (x)

x	y
3	3,85
4	4,34
5	4,83

Значения в таблице вычислены по формуле:

$$y = 0,491x + 2,374.$$

Формула получена по методу наименьших квадратов в предположении линейной зависимости $y(x)$.

По критерию Фишера незначимость линейной зависимости $y(x)$ отвергается с вероятностью ошибки 0,05.

Таким образом, подтверждена статистическая зависимость экзаменационных оценок от результатов тестирования.

АНАЛИЗ СТАТИСТИЧЕСКОЙ ЗАВИСИМОСТИ ОЦЕНОК ЗА КОНТРОЛЬНУЮ РАБОТУ ОТ ОЦЕНОК ТЕСТИРОВАНИЯ НА ВТОРОМ КУРСЕ ИНЖЕНЕРНО-ЭКОНОМИЧЕСКОГО ФАКУЛЬТЕТА (ИнЭИ) НИУ «МЭИ»

Использовались оценки по дисциплине «Объектно-ориентированный анализ и программирование» в осеннем семестре 2018-2019 учебного года. Исследовалась зависимость оценок за контрольную работу от оценок тестирования по этой же теме. Имелась выборка из 32 наблюдений.

Первичный анализ выборки показал:

- Оценки тестирования отличаются от оценок по контрольной работе не более чем на 1 балл для 29 (из 32) наблюдений. Это позволяет предположить наличие статистической связи оценок контрольной работы и оценок тестирования.
- Оценки тестирования в 29 наблюдениях (из 32) были больше или равны оценок за контрольную работу. Таким образом, в отличие от предыдущего исследования, экспертные оценки были в основном более строгими, чем оценки тестирования.

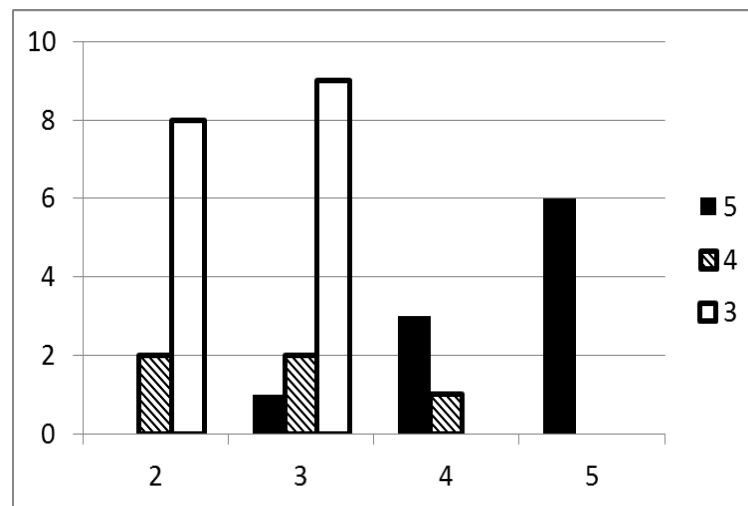
Москва, Россия
14-17 апреля 2020 г.

КОЛИЧЕСТВО ОЦЕНОК ЗА КОНТРОЛЬНУЮ РАБОТУ ДЛЯ КАЖДОЙ ИЗ ОЦЕНОК ТЕСТИРОВАНИЯ ПРЕДСТАВЛЕНО В ТАБЛИЦЕ И НА ГИСТОГРАММЕ:

Оценки		Контрольная работа				Всего
		2	3	4	5	
Т е с т	5	0	1	3	6	10
	4	2	2	1	0	5
	3	8	9	0	0	17

ВЫВОДЫ:

1. Все студенты, получившие «5» за тест, написали контрольную работу на положительную оценку, причем 9 из десяти студентов получили «4» и «5».
2. Из 17 студентов, получивших за тестирование «3», ни один не получил за контрольную оценку «4» или «5».
3. Из 32 студентов, сдавших тест, 24 студента (75 %) получили положительную оценку по контрольной работе.



Гистограмма количества оценок за контрольную работу. В легенде справа - оценки тестирования, по горизонтальной оси - оценки по контрольной работе.

СРЕДНЯЯ ОЦЕНКА ЗА КОНТРОЛЬНУЮ РАБОТУ (y) ДЛЯ РАЗЛИЧНЫХ ОЦЕНОК ТЕСТИРОВАНИЯ (x)

x	y
3	2,44
4	3,40
5	4,35

Значения в таблице вычислены по формуле:

$$y = 0,955x - 0,422.$$

Формула получена по методу наименьших квадратов в предположении линейной зависимости $y(x)$.

Из таблицы очевидно, что оценки за контрольную работу в среднем ниже, чем оценки за тест.

По критерию Фишера незначимость линейной зависимости $y(x)$ отвергается с вероятностью ошибки 0,05.

Таким образом, подтверждена статистическая зависимость оценок за контрольную работу от оценок за тест.

ПРИМЕР, КОГДА ОТСУТСТВУЕТ СТАТИСТИЧЕСКАЯ ЗАВИСИМОСТЬ ОЦЕНОК , ПОСТАВЛЕННЫХ ПРЕПОДАВАТЕЛЕМ, ОТ ОЦЕНОК ТЕСТИРОВАНИЯ

Исследуется еще одна группа потока, данные по которому только что были изложены. Эта группа рассматривается отдельно, так как является самой слабой группой потока по данной дисциплине. Контрольную работу подавляющее число студентов группы написали неудовлетворительно, поэтому оценки были получены в результате многократного переписывания. Тест также многократно передавался. Выборка состоит из 19 наблюдений.

Первичный анализ выборки показал:

- Оценки тестирования отличаются от оценок по контрольной работе не более чем на 1 балл для 16 (из 19) наблюдений. Поэтому было решено попытаться выявить статистическую зависимость оценок контрольной работы от оценок тестирования.
- Оценки тестирования в основном (в 18 наблюдениях из 19), как и в предыдущем исследовании, были больше или равны оценок за контрольную работу. Автор считает, что это свойство выборки в данном случае объясняется тем, что у большинства студентов группы отсутствуют необходимые навыки решения задач.

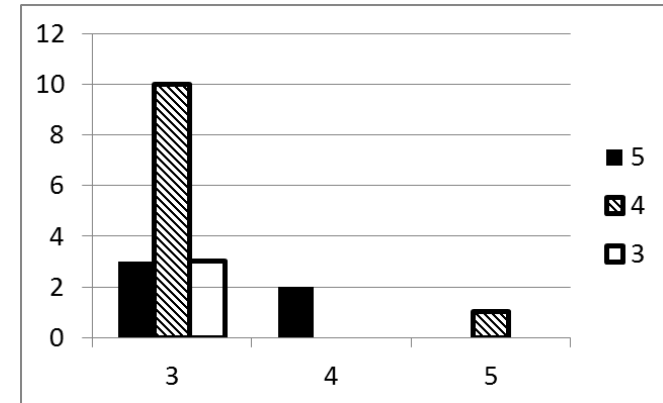
КОЛИЧЕСТВО ОЦЕНОК ЗА КОНТРОЛЬНУЮ РАБОТУ ДЛЯ КАЖДОЙ ИЗ ОЦЕНОК ТЕСТИРОВАНИЯ :

Оценки		Контрольная работа			
		3	4	5	Всего
Т е с т	5	3	2	0	5
	4	10	0	1	11
	3	3	0	0	3

ВЫВОДЫ:

1. Связь между оценками тестирования и контрольной работы не является очевидной.
2. Большинство студентов (11 из 19) сдали тест на оценку «4», в результате многократных пересдач.
3. Большинство студентов (16 из 19) получили по контрольной работе оценку «3», в результате многократных переписываний.

По критерию Фишера гипотеза о незначимости линейной зависимости оценок контрольной работы от оценок тестирования не отклоняется. Таким образом, нет оснований считать эти оценки статистически связанными.



Гистограмма количества оценок за контрольную работу. В легенде справа - оценки тестирования, по горизонтальной оси - оценки по контрольной работе.

ЗАКЛЮЧЕНИЕ

1. Опыт автора показывает (и это подтверждается статистическим анализом), что хорошие тесты, многократно опробованные на разных потоках, часто теряют свою объективность в следующих ситуациях:

- когда один и тот же тест многократно пересдается с целью улучшения оценки; «коллективный разум» неуспевающих студентов в процессе многократных пересдач по крупицам собирает тест и распространяет его в своем «коллективе»;
- когда тестирование проводится не в часы занятий в присутствии преподавателя, а в любое время с любого компьютера с доступом в Интернет; без сомнения, в этой ситуации плохо успевающие студенты сдают тесты с чужой помощью.

2. На современном уровне автоматизации образования объективную оценку знаний дают традиционные, проводимые преподавателем, контрольные мероприятия: экзамены, защиты лабораторных и практических работ, контрольные работы. Тестирование можно использовать только в сочетании с этими мероприятиями, например, для самостоятельной проверки знаний студентами, а также, при определенных обстоятельствах, для оперативного (но не окончательного!) оценивания знаний.

БЛАГОДАРНОСТИ

Автор выражает глубокую благодарность:

- Администраторам и организаторам системы дистанционного обучения «Прометей» НИУ «МЭИ» за безотказную и оперативную организацию тестирования на потоках очного, очно-заочного и заочного обучения.
- Деканату ЭТФ и лично заместителю директора ИРЭ Денису Анатольевичу Зезину за предоставление данных об оценках студентов первого курса по дисциплине «Информатика».

Спасибо за внимание!



Контакты докладчика:

Батасова Валентина Сергеевна

BatasovaVS@mpei.ru

