

Русаков С. В., Русакова О. Л.

**АПОСТЕРИОРНАЯ ОЦЕНКА ТРУДОЕМКОСТИ УЧЕБНЫХ
ЗАДАНИЙ В РАМКАХ БАЛЬНОЙ-РЕЙТИНГОВОЙ СИСТЕМЫ**

Сергей Владимирович Русаков

доктор физико-математических наук, профессор,

заведующий кафедрой

прикладной математики и информатики

rusakov@psu.ru

ФГБОУ ВО «Пермский государственный национальный исследовательский

университет», Россия, Пермь

Ольга Леонидовна Русакова

кандидат физико-математических наук, доцент,

доцент кафедры

прикладной математики и информатики

rol58@yandex.ru

ФГБОУ ВО «Пермский государственный национальный исследовательский

университет», Россия, Пермь

**A POSTERIORI ASSESSMENT OF THE LABOR INTENSITY OF
TRAINING TASKS WITHIN THE SCORING SYSTEM**

Sergei Vladimirovich Rusakov

Perm State National Research University, Russia, Perm

Olga Leonidovna Rusakova

Perm State National Research University, Russia, Perm

Аннотация. В работе описывается методика оценки трудоёмкости (сложности) учебных заданий на основе статистического анализа результатов, показанных учебной группой в предыдущем учебном цикле и зафиксированных в рамках модульной балльно-рейтинговой системы. В качестве меры трудоёмкости задания служит средний балл набранный за его выполнение

всей учебной группой. Производится сравнение средних баллов для всех учебных заданий в рамках конкретной дисциплины, позволяющее выделить из них более сложные или наоборот, вызвавшие наименьшие затруднения у студентов. В качестве примера рассматриваются результаты освоения конкретной учебной дисциплины.

Abstract. *The paper describes a methodology for assessing the labor intensity (complexity) of training tasks based on a statistical analysis of the results shown by the training group in the previous training cycle and recorded within the framework of a modular scoring system. As a measure of the complexity of the task, the average score scored for it by the group is used. Comparison of average scores for all educational tasks within a particular discipline is made, which allows to single out more complex ones from them or vice versa, which caused the least difficulties for students. As an example, the results of mastering a specific academic discipline are considered.*

Ключевые слова: *модульная балльно-рейтинговая система, сложность задания, выборочное среднее, стандартное отклонение, статистические критерии различимости.*

Keywords: *modular scoring system, task complexity, sample mean, standard deviation, statistical criteria of distinguishability.*

В оценке выполнения различных учебных заданий в рамках балльно-рейтинговой системы (БРС) важным моментом является выбор максимального значения балла (разболовка), отводимых на ту или иную работу. В то время как для оценки сложности тестовых заданий существуют хорошо зарекомендовавшие себя методы, основанные на статистической обработке большого числа результатов тестирования (см. например, книгу [1]), баллы, отводимые в рамках БРС на ту или иную работу, выполняемую обучаемыми в период времени отводимого на самостоятельную работу, чаще всего устанавливаются непосредственно преподавателем, исходя из его экспертных оценок (априори).

Здесь так же присутствует некоторый субъективизм и при выставлении оценочных баллов. В то же время, если принять гипотезу, что средний балл показанный учебной группой тем ниже, чем сложнее задание, появляется возможность на основе реальных данных оценить трудоёмкость учебного задания.

Методику оценки трудоемкости (сложности) заданий разберём на примере конкретной дисциплины, в рамках которой студенты в течении триместра должны были выполнить и сдать 7 лабораторных работ. Отчеты по работам представлялись в электронном виде по E-mail. На выполнение каждой работы отводилось две недели, и она оценивалась максимально в 10 баллов. В тех случаях, когда отчет по лабораторной присылался после установленного срока на t дней начислялся штрафной балл $-0.1 \cdot t$. Кроме того, в случае наличия серьезных ошибок начислялось -0.5 и отчет отправлялся на доработку. Всё это позволяло хотя бы косвенным образом оценить затраты времени на выполнение задания конкретным студентом [2].

Таблица 1 — Результаты выполнения лабораторных работ

№ студента	Лаб.1	Лаб.2	Лаб.3	Лаб.4	Лаб.5	Лаб.6	Лаб.7	Средний балл студента
1	8,25	7,25	5,25	4,8	6,10	6,25	6,65	6,36
2	5,75	5,5	6,00	5,60	5,10	6,25	8,30	6,07
3	8,15	8,75	7,75	5,75	8,10	7,45	8,75	7,81
4	7,5	9,00	8,25	6,75	9,50	8,00	7,75	8,11
5	6,75	8,00	7,20	4,45	8,35	5,85	6,40	6,71
6	8,00	8,65	8,25	6,80	6,00	7,75	8,15	7,66
7	10,0	10,0	9,50	6,30	7,60	7,30	9,00	8,53
8	7,75	8,45	6,80	5,05	5,00	7,50	8,50	7,01
9	5,50	6,5	6,00	5,45	5,60	7,65	7,75	6,35
10	9,00	9,75	6,75	6,65	7,90	10,0	9,00	8,44
11	8,25	9,25	9,25	7,05	8,60	9,75	9,50	8,81
12	7,75	8,75	7,45	8,00	7,00	8,50	7,50	7,85
13	9,75	9,75	9,5	8,40	9,00	9,75	7,00	9,02
14	7,50	8,25	5,7	7,25	8,50	8,50	8,25	7,71
15	8,0	9,50	9,5	8,00	8,25	7,75	8,50	8,50
16	8,25	7,75	4,15	5,85	6,70	7,50	8,75	6,99
Среднее	7,88	8,44	7,33	6,38	7,33	7,86	8,11	
Ст. отклонение	1,20	1,23	1,65	1,19	1,43	1,23	0,88	

В таблице 1 представлены конкретные данные балльных оценок семи лабораторных работ, выполненных группой из 16 студентов механико-математического факультета ПГНИУ. Как видно из таблицы, средний балл варьируется в широких пределах от 6.38 до 8.44 и, наверное, можно было по крайней мере считать, что лабораторная работа 2 проще, а лабораторная 4 сложнее на фоне остальных. Кроме того, можно отметить значительную величину стандартных отклонений (за исключение лабораторной 7), что во-первых, свидетельствует о различном уровне подготовки студентов, а во-вторых — наличии заметных «выбросов», возможно имеющих случайный характер. Можно так же отметить, что результаты отдельных студентов (средний балл студента) так же варьируются в широких пределах от 6.07 до 9.02. Таким образом явно видно наличие в группе как «сильных» так и «слабых» студентов. Объективность такого утверждения косвенно подтверждается таблицей парных линейных корреляций столбцов таблицы 1.

Таблица 2 — Парные линейные корреляции

	Лаб.1	Лаб.2	Лаб.3	Лаб.4	Лаб.5	Лаб.6	Лаб.7
Лаб.1	1,000	0,825	0,462	0,422	0,431	0,473	0,236
Лаб.2	0,825	1,000	0,712	0,583	0,666	0,618	0,284
Лаб.3	0,462	0,712	1,000	0,555	0,520	0,371	0,158
Лаб.4	0,422	0,583	0,555	1,000	0,506	0,730	0,194
Лаб.5	0,431	0,666	0,520	0,506	1,000	0,462	-0,001
Лаб.6	0,473	0,618	0,371	0,730	0,462	1,000	0,399
Лаб.7	0,236	0,284	0,158	0,194	-0,001	0,399	1,000

Из таблицы 2 видно, что имеет место значимая корреляция результатов выполнения всех лабораторных за исключением седьмой, то есть «сильные» студенты стабильно выполняют задания лучше «слабых». Особое место седьмой лабораторной связано, возможно, с тем что эту тему (сетевое планирование) студенты ранее частично рассматривали в разделе «Теория графов» курса «Дискретной математики» и отчасти сработали остаточные знания.

Прежде, чем устанавливать статистическую значимость различий средних баллов за выполнение каждой лабораторной, проведем «очистку» исход-

ных данных с помощью процедуры сжатия информационного пространства [3], идея которой состоит в следующем: будем исключать из выборки те значения случайной величины X , которые не попадают в интервал: $X_{cp,j}-k\cdot\sigma_j \leq X_{i,j} \leq X_{cp,j}+k\cdot\sigma_j$. Где $X_{i,j}$ — рейтинговый балл i -го студента за j -ую лабораторную, $X_{cp,j}$ — средний балл все группы за j -ую лабораторную, а σ_j — соответствующее стандартное отклонение. Величина k устанавливается и выбирается таким образом, чтобы уменьшение всей выборки («сжатие информационного пространства») имело наперед заданную величину. В нашем случае значение $k=1.6$ обеспечило «сжатие» примерно на 10%. После применения этой процедуры произошла корректировка и средних значений и стандартных отклонений, результаты которой представлены на рисунках 1 и 2. Как мы видим из этих диаграмм для всех лабораторных, кроме четвертой и шестой, средние значения немного увеличились, а стандартные отклонения уменьшились во всех случаях.

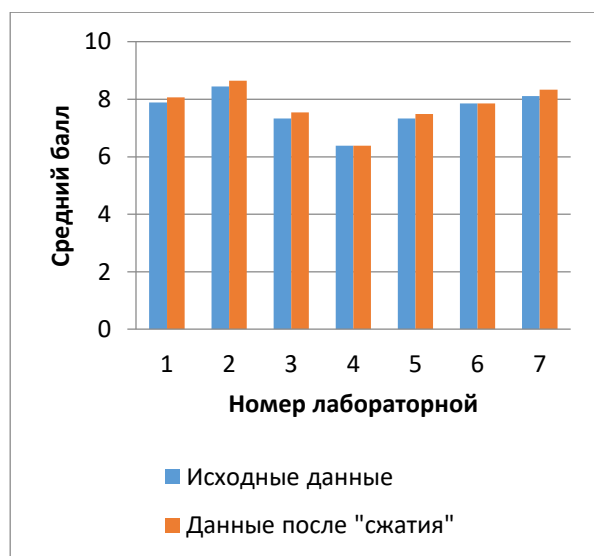


Рисунок 1 — Среднее значение баллов

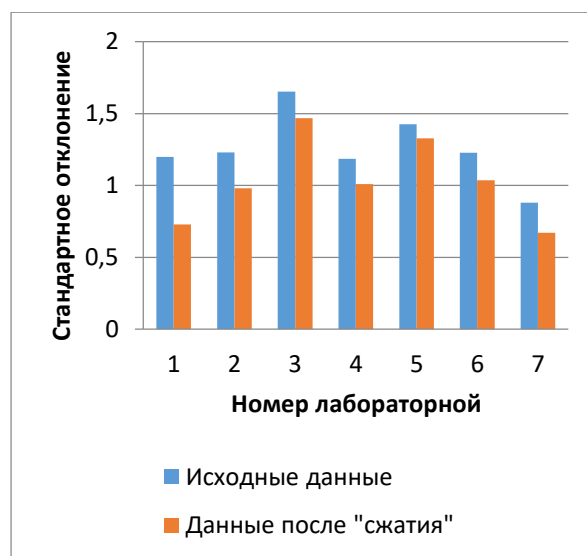


Рисунок 2 — Стандартное отклонение

Статистическое различие средних значений по каждой из лабораторных будем устанавливать с помощью t -критерия, подразумевающего нормальное распределение выборочных данных, и непараметрического U -критерия (критерий Манна-Уитни) [4]. Отметим, что в случае t -критерия статистическая различимость средних с заданным уровнем значимости α фиксируется при значении $t > t_{критич}$, а в U -критерии — при $U < U_{критич}$. В таблице 3 представлены

значения этих критериев рассчитанных по всей выборке, при этом, над главной диагональю значения U , а под диагональю t . В свою очередь, в таблице 4 — аналогична информация, по данным после процедуры сжатия информационного пространства.

Таблица 3 — Рассчитанные значения t и U критериев для исходной выборки

t \ U	Лаб.1	Лаб.2	Лаб.3	Лаб.4	Лаб.5	Лаб.6	Лаб.7
Лаб.1	---	81(16,16)	106(16,16)	52(16,16)	111(16,16)	122(16,16)	100(16,16)
Лаб.2	1.302	---	79(16,16)	31(16,16)	69(16,16)	91(16,16)	102(16,16)
Лаб.3	1.084	2.160	---	83(16,16)	125(16,16)	97(16,16)	88(16,16)
Лаб.4	3.559	4.821	1.862	---	77(16,16)	49(16,16)	32(16,16)
Лаб.5	1.187	2.362	0.000	2.042	---	108(16,16)	84(16,16)
Лаб.6	0.058	1.345	1.026	3.458	1.122	---	96(16,16)
Лаб.7	0.605	0.884	1.662	4.674	1.857	0.662	---

Таблица 4 — Рассчитанные значения t и U критериев после применения процедуры сжатия информационного пространства на 10 %

t \ U	Лаб.1	Лаб.2	Лаб.3	Лаб.4	Лаб.5	Лаб.6	Лаб.7
Лаб.1	---	52(13,15)	79(13,15)	18(13,14)	84(13,15)	76(13,14)	58(13,14)
Лаб.2	1.760	---	65(15,15)	14(15,14)	55(15,15)	60(15,14)	85(15,14)
Лаб.3	1.224	2.405	---	56(15,14)	110(15,15)	84(15,14)	66(15,14)
Лаб.4	5.010	6.108	2.503	---	52(14,15)	30(14,14)	10(14,14)
Лаб.5	1.462	2.703	0.111	2.537	---	95(15,14)	62(15,14)
Лаб.6	0.639	2.104	0.653	3.803	0.824	---	56(14,14)
Лаб.7	0.985	0.980	1.890	6.040	2.193	1.472	---

И в той и другой таблице, значения удовлетворяющие критерию статистической различимости при уровне значимости $\alpha=5\%$ выделены полужирным курсивом, а соответствующие ячейки «раскрашены». В части таблиц, где представлены данные U -критерия, в круглых скобках указано количество элементов в сравниваемых векторах. Можно отметить, что процедура сжатия информационного пространства позволила более четко выделить две лабораторные, средние результаты выполнения которых отличаются от остальных: это лабораторная 2 (средний балл выше) и лабораторная 4 (средний балл ниже). Таким образом можно считать лабораторную 2 проще, а лабораторную 4 сложнее. Для остальных лабораторных уровень сложности примерно одинаков, что свидетельствует о достаточной объективности первичной экспертной оценки.

Заключение

В настоящей статье на конкретном примере показано, как на основе анализа среднего по группе значения рейтингового балла можно оценить трудоёмкость (сложность) заданий, выполняемых в режиме самостоятельной работы. Полученную информацию можно использовать для корректировки заданий на следующий учебный цикл. Кроме того, в определенных случаях преподаватель может рассчитать поправочные коэффициенты, чтобы таким уровнем выровнять показатели. Так в рассмотренном примере баллы набранные при выполнении лабораторных 1 и 2 входят в первую контрольную точку [2], а лабораторных 3 и 4 во вторую. Причем на каждую из этих контрольных точек отводится по 20 баллов. Тогда поправочные коэффициенты (1.035; 0.965) и (0.916; 1.084) для результатов лабораторных 1–2 и 3–4 соответственно, позволят выполнить это выравнивание.

Список литературы

1. Самылкина, Н. Н. Современные средства оценивания результатов обучения / Н. Н. Самылкина. Москва: БИНОМ. Лаборатория знаний, 2007. 172 с. Текст: непосредственный.
2. Русаков, С. В. Анализ эффективности учебной деятельности студентов в рамках балльно-рейтинговой системы / С. В. Русаков, О. Л. Русакова, М. Д. Смольяков. Текст: электронный // Новые информационные технологии в образовании и науке. 2022. № 5. 2022. С. 74–79. URL: <https://elar.rsvpu.ru/handle/123456789/41613?ysclid=lhuvb2po4z751692665>.
3. Гусев, А. Л. Метод отбора информационного пространства для построения нейронной сети, как модели управления, в условиях зашумленных и не полных данных / А. Л. Гусев. Текст: непосредственный // Нейрокомпьютеры: разработка, применение. 2011. № 7. С. 55–57.
4. Айвазян, С. А. Прикладная статистика. Основы моделирования и первичная обработка данных / С. А. Айвазян, И. С. Енюков, Л. Д. Мешалкин; под ред. С. А. Айвазяна. Москва: Финансы и статистика, 1983. 471 с. Текст: непосредственный.