

КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ В ПЕДАГОГИЧЕСКОМ МОНИТОРИНГЕ

Евгений Драган¹, Юрий Жарких²

Киевский национальный университет имени Тараса Шевченко
ул. Владимирская, 64, Киев 01033, Украина, тел.: (380 44) 234-31-08
E-Mail : eugene_dragan@univ.kiev.ua¹, yzh@uninet.kiev.ua²

Abstract

The work presents results of computer testing in physics. Students were given test items of our design. It was shown that the mark for the testing has high reliability and good correlation with marks given by other monitoring methods.

It was ascertained that the level of knowledge of students is distributed by normal law of random variable. This is the consequence of superposition of latent parameters, which characterize aptitude of each student. Maximum value of distribution corresponds to mark which shows ability level of whole group. Dispersion of the distribution shows the level of group heterogeneity. With these data teacher can plan ways of learning, level and bulk of studying material for further training.

Тестовые методы установления уровня знаний находят все большее применение в системе образования. Развитие этих методов приводит к повышению надежности и точности оценивания подготовленности учащихся. Среди разных способов тестирования следует выделить те, которые основаны на использовании компьютерных технологий. Их основными достоинствами являются экспрессность и возможность применения в дистанционном образовании.

Компьютерное тестирование было применено нами для текущего мониторинга уровня знаний студентов при изучении курса физики. Программа обработки результатов тестирования выдавала оценку в виде процентного содержания правильно выполненных заданий. Таким образом, фактически, использовалась стобальная шкала оценок. Это дает возможность построить рейтинговое распределение студентов: расположить их в ряд по уровню знаний. Рейтинг студентов, установленный компьютерным тестированием, показан точками (•) на рисунке 1.

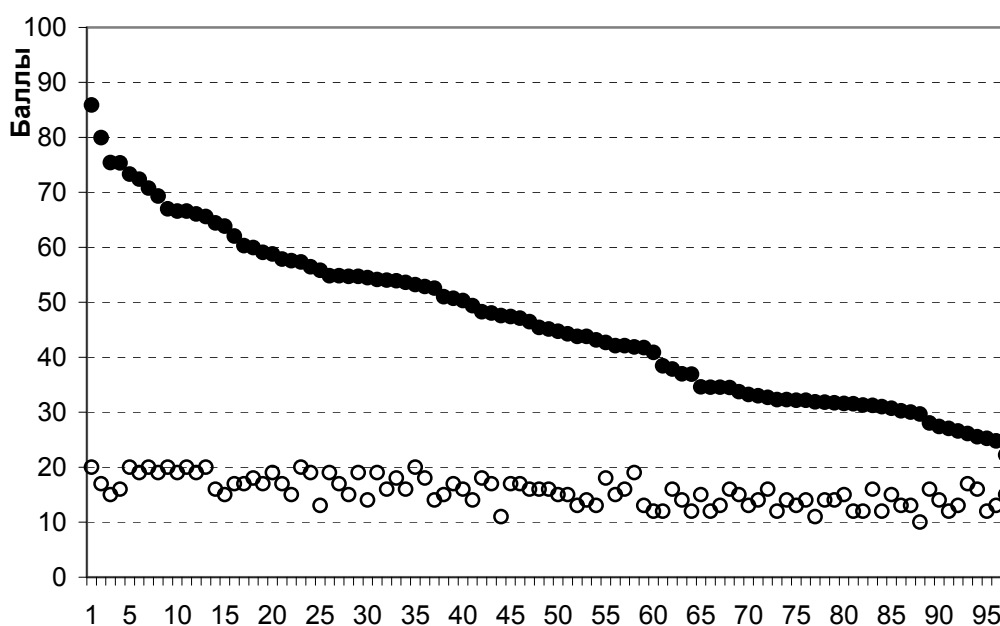


Рис.1 Рейтинги студентов, установленные компьютерным тестированием (•) и по семинарским занятиям (○), 2004 г.

На оси координат указана оценка студента, а на оси абсцисс место студента в рейтинге (с каждым номером связана фамилия конкретного студента). На этом же рисунке точками (○) показаны оценки этих студентов, выставленные преподавателем за работу на семинарах по физике. Эта оценка наиболее надёжно характеризует уровень знаний студента потому, что в ней один и тот же преподаватель обобщает результаты работы студента за весь семестр при решении задач, выступлениях с докладами, написании контрольных работ. Видно качественное согласие между рейтингом и оценками студентов за семинары. Оба распределения спадают от первого студента до девяносто седьмого. В таком простом сравнении мы

говорим только о качественном согласии потому, что рейтинг и семинары оценивают в разных шкалах: 100 и 20 баллов соответственно. Количественное согласие в таком случае можно оценить из расчета коэффициента корреляции k между рейтингом и семинарскими оценками [1]. Определенный по данным, представленным на рисунке 1 коэффициент $k = 0,67$. Таким образом, компьютерное тестирование характеризует рейтинг студента с высокой степенью надежности.

Рейтинг студента является его индивидуальной характеристикой и поэтому имеет большое значение для каждого студента. Для преподавателей же более важными являются данные, характеризующие успешность учебы академической группы, потока, курса. Эти данные дают представление об уровне учебного потенциала студенческого контингента в целом. Поэтому преподаватель может планировать способы, уровень и объемы подачи учебного материала так, чтобы он наиболее эффективно воспринимался большей частью группы.

Общее представление о распределении студентов по уровню знаний в произвольной их выборке можно получить из анализа процентиля [2]. Процентилью характеризует процентное содержание в группе студентов с разным уровнем учебной успешности. Она устанавливает соответствие между любым балльным порогом и частью студенческой группы, которая не смогла преодолеть этот порог. Точками на рисунке 2 показаны значения процентиля, полученные по результатам компьютерного тестирования, рассмотренного выше. По оси абсцисс отложена столбальная шкала оценок (B), а по оси ординат (левая шкала) – относительное число студентов в процентах (P). Кривой 1 показан ход логистической функции:

$$P(B) = \frac{e^{A(B-\mu)}}{1 + e^{A(B-\mu)}} \quad (1)$$

где A – параметр крутизны, μ – среднее значение балла, при их значениях $A = 0,125$, $\mu = 63$. Видно, что значения точек процентиля хорошо укладываются на логистическую кривую. Отметим, что такая же статистическая функция описывает характеристическую кривую успешности студента в вероятностной модели тестирования [3].

Более детальную информацию, которая характеризует учебную успешность студентов можно получить при дифференцировании функции, описывающей процентиль. Дифференциал этой функции показан кривой 2 на рисунке 2. По оси ординат (правая шкала) указана плотность вероятности $C(B)$, которая определяет число студентов приходящихся на единичный интервал баллов, показанных на оси абсцисс. Пунктиром (кривая 3) на этом же рисунке показана кривая нормального распределения случайных величин. Нормальное распределение задает связь плотности вероятности обнаружения некоторой величины x в ансамбле их случайных реализаций (например, ошибки измерений, распределение молекул газа по скоростям и т.п.) [4]. Для рассматриваемого распределения уровней студентов можно записать:

$$C(B) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(B-\mu)^2}{2\sigma^2}} \quad (2)$$

где μ – наиболее вероятное значение величины B , σ – дисперсия (разброс значений B от её средней величины).

Из того факта, что кривые 2 и 3 практически совпадают можно заключить, что оценки студентов распределены по нормальному закону (2). Поэтому успешность обучения некоторой группы студентов в целом можно характеризовать параметрами μ и σ . Максимумы кривых 2 и 3 располагаются против оценки $m = 63$ баллов, которая с наибольшей вероятностью характеризует успешность учебы, данной группы студентов. Дисперсия распределения $\sigma = 15,3$ балла характеризует степень неоднородности группы. Показательным является сам факт распределения рейтинговых оценок по нормальному закону распределения случайных величин. Он свидетельствует о том, что рейтинговая оценка является следствием наложения скрытых параметров, характеризующих каждого студента. Такими параметрами могут быть, например: уровень природных способностей студента, уровень способности к самоорганизации, уровень его физического здоровья.

Естественно, что кроме оценок компьютерного тестирования, студенты получали оценки преподавателей при текущем и рубежном контроле. Анализ массива оценок, полученных этой же группой студентов при текущем и рубежном контроле, показал, что они распределены по нормальному закону таким же образом, как и оценки компьютерного тестирования.

Следовательно, компьютерное тестирование обнаруживает те же рейтинговые особенности группы студентов, что и классический способ оценивания их знаний. Поэтому значение наиболее вероятной оценки, характеризующей уровень знаний группы студентов, можно определять по положению макси-

муна μ кривой $C(V)$ для обоих методов оценивания. Значения μ , найденные по кривым $C(V)$ за 2004-2007 годы показаны на рисунке 3.

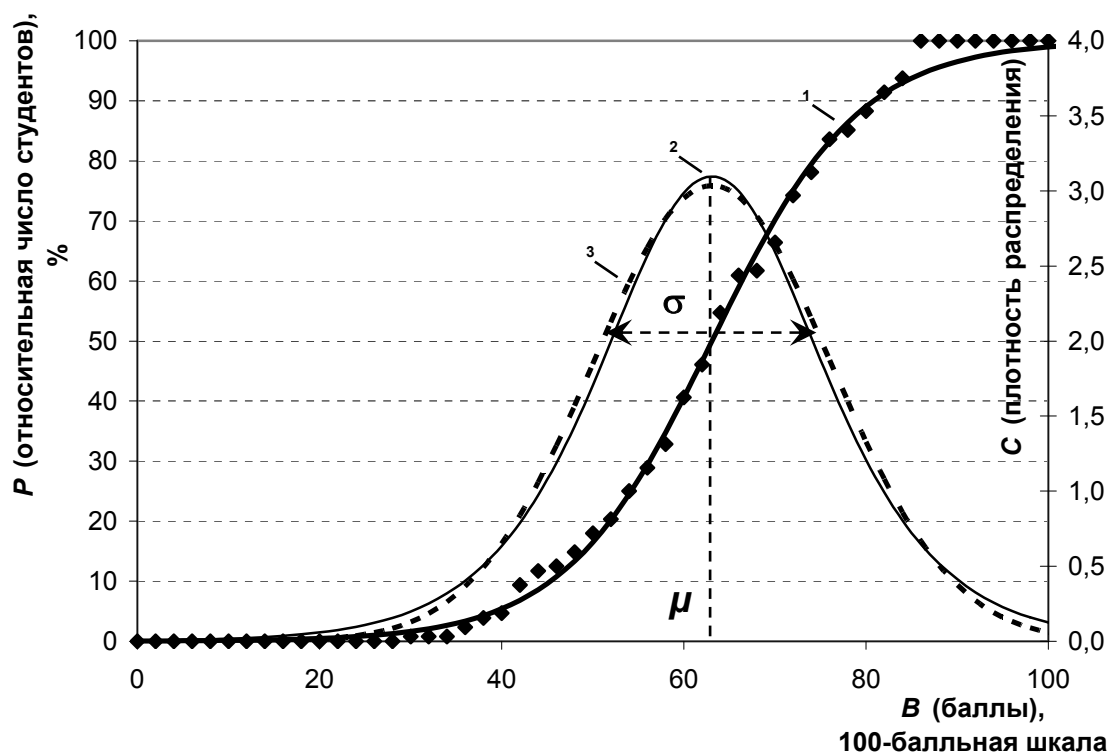


Рис.2. Распределения плотности вероятности $C(V)$, полученные по результатам компьютерного тестирования знаний, 2005 г

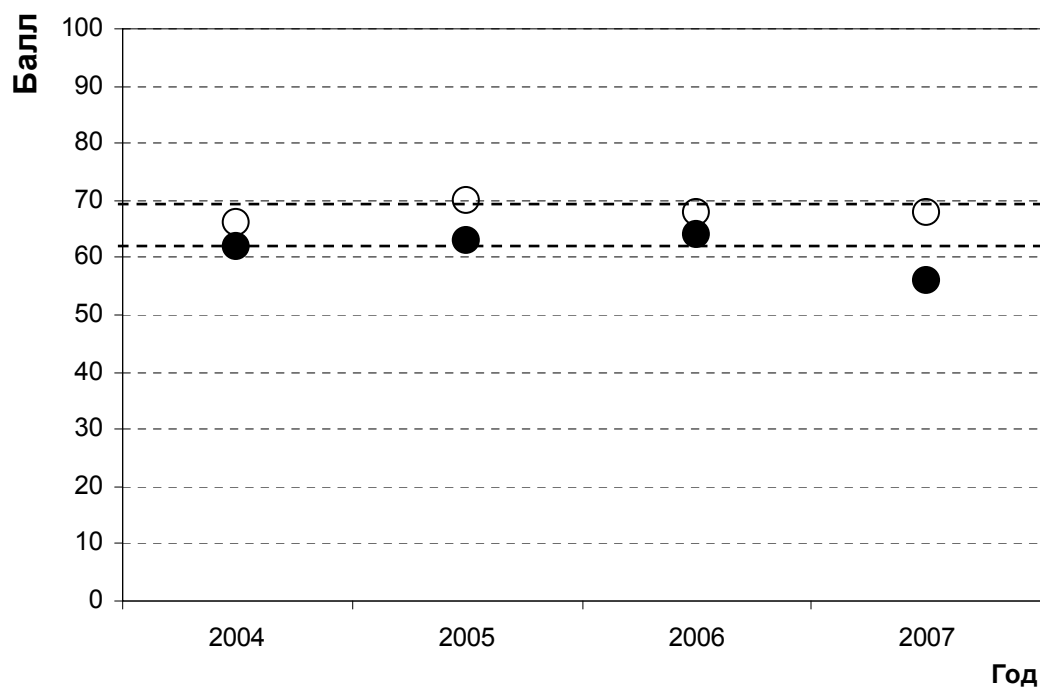


Рис.3 Наиболее вероятная оценка, характеризующая успеваемость студентов в разные годы. (○) – суммарная оценка за семестр, (●) – оценка компьютерного тестирования.

Точки (о) получены по результатам классического оценивания преподавателем, а точки (•) получены при компьютерном тестировании. Видно, что результаты оценивания двумя методами согласуются в том, что среднее значение μ слабо изменяется в рассмотренный период. Однако значение μ классического оценивания выше того, которое получено компьютерным тестированием 69 и 63 соответственно). По-видимому, это может свидетельствовать о том, что субъективно оценивая знания студентов, преподаватели склонны к некоторому завышению оценок.

В заключение сформулируем основные выводы:

1. Компьютерное тестирование является эффективным инструментом для проведения педагогического мониторинга уровня знаний студентов. Оценка такого тестирования имеет высокую степень надёжности и хорошо коррелирует со значениями, полученными другими методами контроля.

2. Результаты компьютерного тестирования и классического оценивания знаний обнаруживают одинаковые статистические особенности. Распределение студентов по уровню знаний в некоторой их группе (потоке, курсе) описывается законом нормального распределения случайных величин. Это вызвано статистическим наложением латентных параметров, характеризующих каждого студента и являющихся случайными величинами. Балльная оценка, соответствующая максимуму функции распределения, позволяет делать выводы об успешности группы в целом. А размытие функции распределения дает представление о разбросе уровня знаний студентов в группе.

3. Сравнение результатов педагогического мониторинга за 2004-2007 годы показало, что средний уровень знаний студентов за это время существенно не изменился. Однако при субъективном оценивании знаний студентов наблюдается тенденция небольшого завышения оценок преподавателями. Очевидно, что этот вывод носит частный характер в соответствии с рассмотренной конкретной ситуацией.

Литература:

- [1] Клайн П. "Справочное руководство по конструированию тестов" К.: ПАН Лтд, 1994.
- [2] Анастаси А., Урбина С. "Психологическое тестирование", СПб.: Питер, 2001.
- [3] Baker F. B. "The Basics of Item Response Theory. ERIC Clearinghouse on Assessment and Evaluation", 2001.
- [4] Гласс Дж., Стенли Дж. "Статистические методы в педагогике и психологии", М.: Прогресс, 1976.