

## МАТЕМАТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ АВТОМАТИЗИРОВАННОГО ТЕСТИРОВАНИЯ

Абылкасымова А.Е. г.Алматы, Алматинский государственный университет, Ашаев Ю.П., Корчагина Л.В. г.Алматы, Казахский национальный технический университет

В основу математического аппарата, используемого при обработке результатов автоматизированного тестирования, положены методы вероятностно-статистического анализа. Ответ студента на конкретное тестовое задание может быть оценен с двух позиций, "правильный" или "неправильный", "истинный" или "ложный", или в соответствии с методами дискретной математики, "1" или "0". Весь тест, состоящий из  $N$  тестовых заданий для конкретного студента, может быть оценен в диапазоне от 0 до  $N$  дискретных единиц. Исходя из того, что абсолютная оценка (в дискретных единицах) является не показательной, предлагается ввести относительную оценку результатов тестирования. Относительная оценка результатов тестирования определяется количеством правильных ответов  $N_1$ , данных в процессе тестирования на  $N$  тестовых заданий, и математически вычисляется по формуле :

$$R = \frac{N_1}{N} \cdot 100\% \quad (1)$$

Значение относительной оценки результатов тестирования может изменяться в диапазоне от 0% до 100%, т.е. при всех неправильных ответах тестирования её значение будет равно 0%, а при всех правильных ответах -100%.

Введение относительной оценки результатов тестирования позволяет сравнить между собой результаты, полученные при раз-

личных вариантах тестирования, и сопоставить результаты независимо от количества тестовых заданий, включенных в тест. Фактически, относительная оценка результатов тестирования отражает по 100-балльной шкале уровень знаний студента по дисциплине в целом или её разделам.

Ввиду того, что в системе высшего образования принята 4-балльная шкала оценок знаний, вводятся дополнительные ограничения, позволяющие переходить от одной шкалы к другой. Например, если значение относительной оценки  $Q$  превышает показатель  $d_1$ , то это соответствует оценке "отлично", если находится в диапазоне  $d_1 \geq Q > d_2$ , то соответствует оценке "хорошо", если  $d_2 \geq Q > d_3$ , то соответствует оценке "удовлетворительно", если  $Q \leq d_3$ , то соответствует оценке "неудовлетворительно".

Показатели  $d_1, d_2, d_3$  определяются в результате экспертного опроса специалистов, их предварительного анализа, и утверждаются соответствующими службами (деканатом, ученым советом, кафедрой и т.п.). На практике,  $d_1$  обычно задается в диапазоне от 70% до 90%,  $d_2$  - в диапазоне от 50% до 75%, и  $d_3$  - в диапазоне от 40% до 55%.

Введение показателя относительной оценки результатов тестирования позволяет не только оценить индивидуальный уровень знаний студента и сравнить уровни знаний студентов между собой, но и рассчитать условный рейтинг  $R$  учебной группы по данной дисциплине, по формуле, аналогичной выражению (1), заменив показатель  $N_1$  на сумму :

$$\sum_{i=1}^k N_1^i$$

где  $N_1^i$  - количество правильных ответов, данных  $i$ -ым студентом исследуемой учебной группы,  $k$  - количество студентов в учебной группе, а параметр  $N$  - на сумму :

$$\sum_{i=1}^k N^i$$

где  $N^i$  - общее количество тестовых заданий, включенных в тест, для  $i$ -го студента.

Общая формула тогда имеет вид :

$$R = \frac{\sum_{i=1}^k N_1^i}{\sum_{i=1}^k N^i} 100\% \quad (2)$$

Если количество тестовых заданий в каждом тесте постоянно и равно  $N$ , то выражение (2) упрощается :

$$R = \frac{\sum_{i=1}^k N_1^i}{k N} 100\% \quad (3)$$

Предложенный подход может быть использован не только для контроля знаний студентов, но и для анализа уровня знаний студентов, и степени усвоения ими конкретных разделов изучаемой дисциплины.

Если, в соответствии с принятой учебной программой, весь учебный курс дисциплины разбить на  $N$  тем (или разделов) и структуру теста сформировать в соответствии с этими темами (или разделами), то в каждый тест будут включены  $N$  тестовых заданий. Результатом опроса каждого студента будет массив, состоящий из  $N$  нулей и единиц:  $\{q_1^i, q_2^i, \dots, q_n^i, \dots, q_N^i\}$ . Далее, исходные массивы по всем  $k$  студентам для каждого  $n$ -го раздела (или темы) обрабаты-

ваются и рассчитывается результирующий  $G$  массив, состоящий из  $G_n$  элементов, определяемых по формуле :

$$G_n = \frac{\sum_{i=1}^k q_n^i}{k} \cdot 100\% \quad (4),$$

где  $q_n^i$  - конкретный ответ "истинный" - 1 или "ложный" - 0, данный  $i$  - ым студентом на  $n$  - ое тестовое задание.

Таким образом, массив  $G = (G_1, G_2, \dots, G_n, \dots, G_{N-1}, G_N)$  будет характеризовать степень усвоения студентами соответствующих разделов (или тем) изучаемой дисциплины. Варьируя показателем  $k$  (количество студентов) можно получить результирующий массив для учебной группы, специальности, потока, ВУЗа, или всех студентов, изучающих данную дисциплину, в соответствии с утвержденной учебной программой.

Значение величины  $G_n$  также может изменяться в диапазоне от 0% до 100%. Наиболее наглядно массив  $G$  может быть представлен графически в виде диаграммы, условный вид которой приведен на рис. 1.

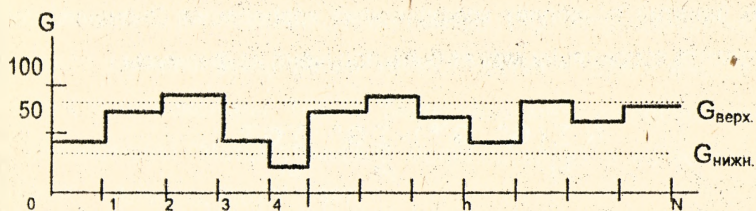


Рис. 1. Эскиз диаграммы, отражающей уровень знаний и степень усвоения студентами разделов изучаемой дисциплины.

Задав соответствующие нижние -  $G_{нижн.}$  и верхние -  $G_{верх.}$  ограничения можно легко выявить разделы, степень усвоения которых

студентами ниже предельно допустимого -  $G_{\text{нижн}}$  и выше рекомендуемого  $G_{\text{верх}}$ . Показатели  $G_{\text{нижн}}$  и  $G_{\text{верх}}$  также могут быть получены на основании предварительного экспертного опроса и обобщения.

Данные массива  $G$  могут быть получены для всего потока. В этом случае они могут служить основой для корректировки учебной программы, с целью более углубленного изучения разделов (или тем), по которым были получены наиболее низкие показатели  $G_n$ . В рамках утвержденной программы это может быть достигнуто путем перераспределения лекционных часов между темами и разделами, по которым получены высокие и низкие значения показателя  $G_n$ .

Кроме того, результаты тестового опроса в учебной группе позволяют выработать рекомендации преподавателям, ведущим практические и лабораторные занятия в этой группе, по более углубленному изучению разделов и тем, по которым студентами в результате тестирования были показаны наиболее низкие знания.

Одной из сторон преподавательской деятельности является согласование лекционных и практических форм изучения дисциплины. Оценить степень согласования тематики лекционных и практических занятий позволяет коэффициент корреляции. Значение коэффициента корреляции может быть получено из формулы :

$$R = \frac{N \sum_{j=1}^N (G_j^n \cdot G_j^{np}) - \sum_{j=1}^N G_j^n \cdot \sum_{j=1}^N G_j^{np}}{\sqrt{[N \sum_{j=1}^N (G_j^n)^2 - (\sum_{j=1}^N G_j^n)^2] \cdot [N \sum_{j=1}^N (G_j^{np})^2 - (\sum_{j=1}^N G_j^{np})^2]}} \quad (5),$$

где  $G_j^n$  - показатель, характеризующий восприятие студентами  $j$  - го раздела (или темы) на лекционных занятиях,  $G_j^{np}$  - показатель,



характеризующий восприятие студентами  $j$  - го раздела (или темы) на практических занятиях.

Чем выше значение коэффициента корреляции, тем выше согласование лекционных занятий с практическими, и, как правило, выше степень усвоения данного материала.

Данные принципы были реализованы в разработанной авторами компьютерной системе индивидуального контроля знаний ТЕСТ. Система обладает широким набором функций индивидуального тестирования с использованием персонального компьютера, с последующей регистрацией результатов тестирования, и их обработкой. Система была апробирована в Казахском национальном техническом университете и в Академии гражданской авиации, и внедрена в учебный процесс для очной и заочной форм обучения. В течение двух последних лет система используется при сдаче экзаменов и зачетов по дисциплине ИНФОРМАТИКА.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ашаев Ю.П., Корчагина Л.В., Ланков П.Ю. Повышение эффективности изучения дисциплины "ИНФОРМАТИКА" путем внедрения компьютерной системы "ТЕСТ". Материалы международной научной конференции "Математическое моделирование в естественных науках". - Алматы. КазГУ. 1997.

2. Абылкасымова А.Е., Ашаев Ю.П., Корчагина Л.В. Методологические аспекты автоматизированного тестирования на примере компьютерной системы "ТЕСТ". Профессиональная направленность обучения математике и информатике в условиях дифференциации школ. - Алматы. АГУ. 1997.