УДК 378.14

ОЦЕНКА ПРОДОЛЖИТЕЛЬНОСТИ ТЕСТИРОВАНИЯ В ЗАВИСИМОСТИ ОТ СЛОЖНОСТИ ТЕСТА

Хузиахметова А.Р., Нуриев Н.К., Хузиахметова Р.Н.

ФГБОУ ВО «Казанский национальный исследовательский технологический университет», Казань, e-mail: huziah@yandex.ru

При создании тестов продолжительность тестирования является значимым фактором, определяющим качество получаемых в процессе тестирования результатов. Превышение или уменьшение оптимального времени тестирования снижает качественные показатели теста. В данной статье дается обоснование продолжительности тестирования в зависимости от сложности тестовых заданий, предлагаемых студентам при оценке полноты усвоенных знаний курса высшей математики. Для этого преподаватель определяет временной интервал, в течение которого он выполнит все тестовые задания, тем самым оценивая сложность предложенного теста. Продолжительность ответа студентов рассматривается как случайная величина. В результате обработки достаточно большого массива экспериментальных данных определяются продолжительность тестирования и его зависимость от сложности теста. Для описания полученной зависимости и реализации результатов предлагается преобразование координат. Вводимая функция имеет смысл обратного времени и определяет время, оставшееся до конца тестирования. На основании статистического анализа было установлено, что данная случайная величина – время до конца тестирования – подчиняется экспоненциальному закону. Для практического использования результатов исследования в заключении приводится соотношение, позволяющее установить допустимое значение продолжительности тестирования, чтобы объективно оценить качество знаний студента из освоенной им предметной области.

Ключевые слова: контроль знаний, трудоемкость теста, оценка продолжительности тестирования, обратное время, показательный закон распределения

THE ASSESSMENT OF THE TEST DURATION DEPENDING ON THE TEST COMPLEXITY

Khuziakhmetova A.R., Nuriev N.K., Khuziakhmetova R.N.

FGBOU VO «Kazan National Research Technological University», Kazan, e-mail: huziah@yandex.ru

While creating the tests, testing process duration is a significant factor that determines the quality of the results obtained in the process of testing. Exceeding or reducing of the optimal test time reduces the quality indicators of the test. In this article you can find the substantiation of test duration in dependence on the complexity of the test tasks which are offered to students during the assessment of completeness of the gained knowledge on the higher mathematics course. To do this, the teacher determines the time interval during which he completes all the tests, thereby assessing the complexity of the proposed test. The length of students' response is considered as a random variable. As a result of processing of the sufficiently large array of experimental data, the duration of the test and its dependence on the test complexity are determined. To describe the obtained dependence and implement the results, the coordinate transformation is proposed. The input function has the meaning of the inverse time and determines the time remaining until the end of the test. On the basis of statistical analysis it was found that this random variable - the time until the end of the test, is a subject to exponential law. In order to provide the practical usage of the study results, in conclusion you can find a ratio that allows you to set the acceptable value of the test duration to objectively assess the quality of student's knowledge from his subject area.

Keywords: knowledge control, test complexity, assessment of the test duration, return time, exponential law of distribution

В процессе математической подготовки студентов в метрическом компетентностном формате как в одной из форм педагогической технологии необходимо установить достаточно точные метрики средств обучения, в частности продолжительность тестового контроля знаний, оценку сложности тестовых задач.

Цель исследования. Педагогическое тестирование как инструмент проверки качества усвоенных знаний является широко распространенным средством контроля знаний студентов. К примеру, методика оценки продолжительности тестирования для дисциплины «Прикладная математика» направления подготовки «Информационные системы и технологии» рассмотрена в [1]. Какова продолжительность тестирования при изучении курса высшей математики студентами направления «Химическая технология» в технологическом университете, чтобы объективно, т.е. с учетом психологических и интеллектуальных особенностей обучаемого, оценить качество усвоенных им знаний?

Материал и методы исследования. Пусть S – продолжительность временного интервала, в течение которого преподаватель (эксперт) выполнит все задания предложенного теста (S – экспертное время). Очевидно, что ни один из студентов не выполнит тестовые задания за время, меньшее S. Обозначим через V время окончания студентами всего теста, т.е. весь состав тестируемых студентов закончит работу на интервале $[S;V]$.

Тест на полноту усвоенных знаний содержит 10 вопросов (полнота знаний – это мера знаний теории в рамках рассматриваемой дисциплины). Варианты теста являются однородными по содержанию и одинаковыми по сложности. Сложность каждого задания случайно выбранного теста оценивает эксперт – преподаватель в минутах работы (мин/раб) [2].

Перейдем к оценке случайной величины X – времени выполнения студентом теста на полноту усвоенных знаний. Ее значение лежит в достаточно широком диапазоне: скорость реакции студентов на вопрос теста различна [3]. Нужно учесть и тот факт, что некоторые студенты не успевают ответить на вопрос по причине незнания материала. Таким студентам время тестирования можно увеличивать сколь угодно, но качество ответов от этого не улучшится [4].

К промежуточному экзамену (коллоквиуму) по теме «Дифференциальное исчисление функции одной переменной» были допущены 115 студентов. Усредненные данные представлены в таблице 1.

Таблица 1

Закон распределения случайной величины X теста на полноту усвоенных знаний

№	Срединное значение интервала, S	Значение эмпирической функции распределения $F^*(x)$
1	1,125	0
2	1,375	0,017
3	1,625	0,077
4	1,875	0,155

5	2,125	0,285
6	2,375	0,459
7	2,625	0,659
8	2,875	0,859
9	3,125	0,937
10	3,375	0,981
11	3,625	0,998

Пусть X – случайная величина – время выполнения студентом тестовых заданий. $P(X < 1{,}25 * S) = 0$, т.е. вероятность того, что студент выполнит тестовые задания за время, меньшее, чем $1{,}25 *$ S, равна нулю, где S – сложность теста. Ко времени $2{,}75 *$ S тестирование завершит чуть более половины студентов: $P(X < 2{,}75 * S) = 0{,}659$ и т.д. Из данных таблицы видно, что время выхода студентов по завершении тестирования начинается с момента времени $1{,}25 * S)$. В этом случае величина X будет распределена по нормальному закону с распределением Гаусса [5]. Полигон относительных частот, представленный на рисунке 1, иллюстрирует тот факт, что студенты стараются как можно полнее использовать отведенный лимит времени и заканчивают тестирование по его истечении. Возникает вопрос: сколько времени (величина V) нужно дать студентам, чтобы они ответили на поставленные 10 вопросов независимо от скорости (быстроты) реакции при достаточной подготовке по конкретному разделу?

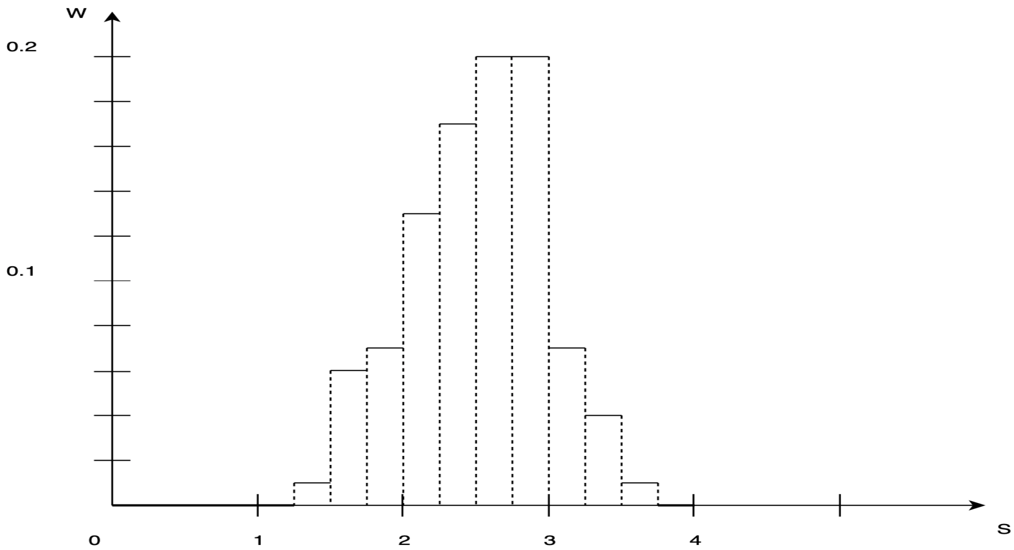

Рис. 1. Полигон относительных частот времени тестирования

Тест на полноту содержит вопросы теоретического характера: определения, свойства, методы вычисления соответствующих характеристик, приложения. Соответственно, независимо от продолжительности тестирования V оценка за тест, как правило, полностью зависит от степени подготовленности студента. Понятно, что многие студенты, даже выполнив все задания, не покидают аудиторию до окончания тестирования (параметр V). Однако неподготовленным студентам выделенного для ответов времени всегда будет недостаточно. Цель данной работы – обосновать выбор времени V, так как излишнее увеличение времени тестирования не приведет к улучшению качества ответов студентов.

Авторами выполнен значительный объем наблюдений при проведении тестирования со студентами института нефти и нефтехимии по следующим разделам курса высшей математики: линейная и векторная алгебра, аналитическая геометрия, дифференциальное и интегральное исчисления функции одной переменной, комплексные числа, дифференциальные уравнения. Полученные результаты дают основание считать, что продолжительность тестирования не должна превышать величины $3S$, а функция плотности вероятности распределения времени тестирования имеет вид, представленный на рисунке 2 ($V = 3S$).

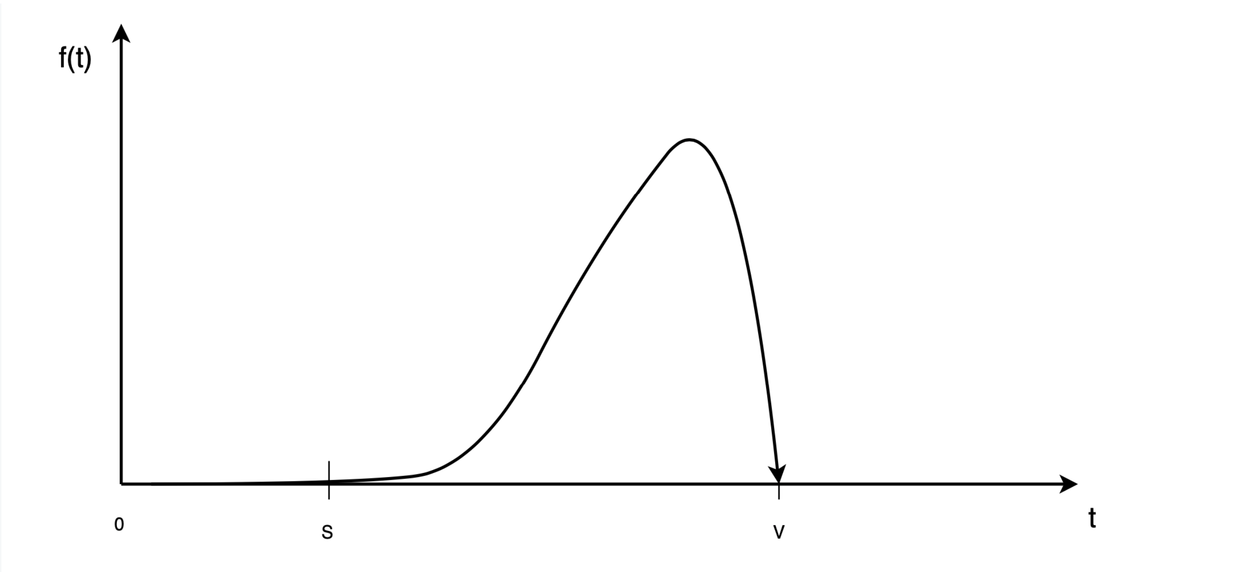

Рис. 2. Вид функции плотности вероятности случайной величины X

В таблице 2 представлен статистический ряд данных на полноту усвоенных знаний по *линейной, векторной алгебре* (объем выборки $n = 77$).

Таблица 2

Наблюдаемые частоты, полученные в ходе тестирования

Номер интервала	Интервал, S	Частота n_i
1	[1; 1,25)	2
2	[1,25;1,5)	2

3	[1,5; 1,75)	8
4	[1,75; 2)	8
5	[2; 2,25)	9
6	[2,25; 2,5)	11
7	[2,5; 2,75)	16
8	[2,75; 3)	17
9	[3; 3,25)	4

Описание зависимости, представленной на рисунке 2 и соответственно в таблице 2, достаточно сложно, поэтому предлагается другой путь, более простой в реализации и интерпретации результатов. А именно вводится линейная функция вида:

$$\tau = V - t,$$

представляющая собой преобразование координат. При этом начало координат переносится в точку $t = V$, ось абсцисс меняет свое направление на противоположное, а величина τ, имея смысл обратного времени, показывает, сколько времени осталось до конца тестирования. Точка S оси t отображается на оси τ в точку $V - S$. Полигон относительных частот «обратного времени» представлен на рисунке 3, вариационный ряд – в таблице 3.

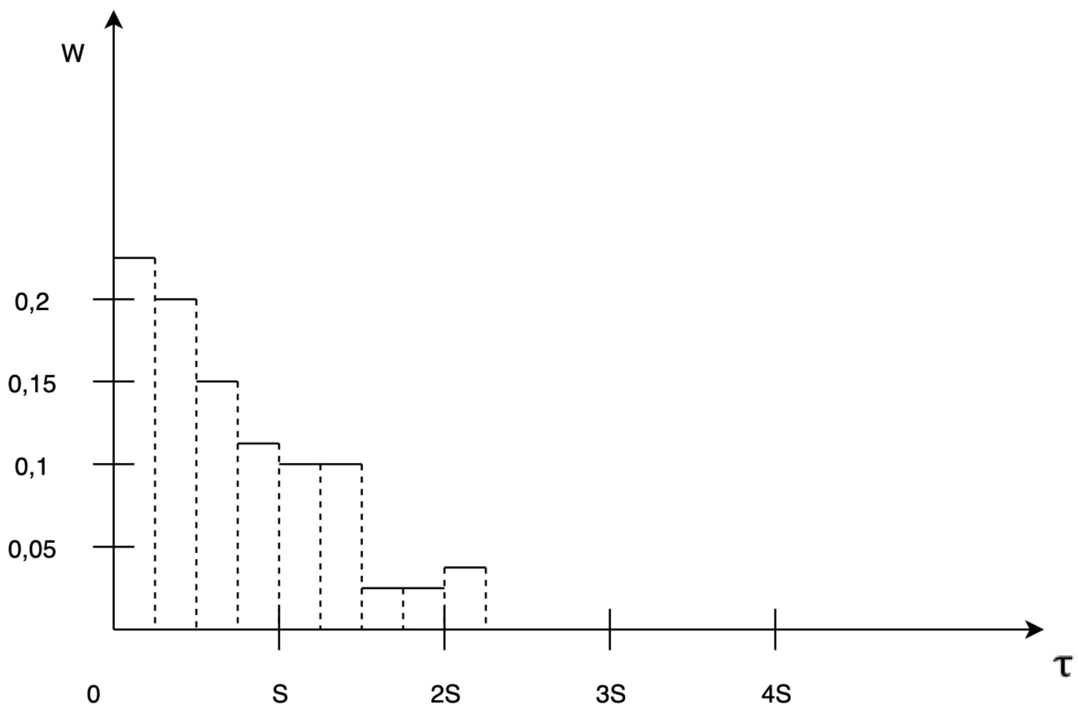

Рис. 3. Полигон относительных частот «обратного времени»

Таблица 3

Вариационный ряд «обратного времени»

Номер интервала	1	2	3	4	5	6	7	8	9
Интервал «обратное время», τ (S)	(0;0,25]	(0,25;0,5]	(0,5;0,75]	(0,75;1]	(1;1,25]	(1,25;1,5]	(1,5;1,75]	(1,75;2]	> 2
Частота n_i	17	16	11	9	8	8	2	2	4

Функция плотности вероятности распределения «обратного времени» $g(\tau)$ с достаточной степенью точности аппроксимируется кривой, приведенной на рисунке 4, описание которой задается однопараметрической функцией:

$$g(\tau) = \begin{cases} \lambda e^{-\lambda\tau}. \tau \geq 0 \\ 0, \tau < 0 \end{cases},$$

где λ интерпретируется как интенсивность выхода студентов по окончании выполнения тестовых заданий и определяется как величина, обратная среднему времени тестирования, т.е. $\lambda = \dfrac{1}{\overline{x_в}(\tau)}$.

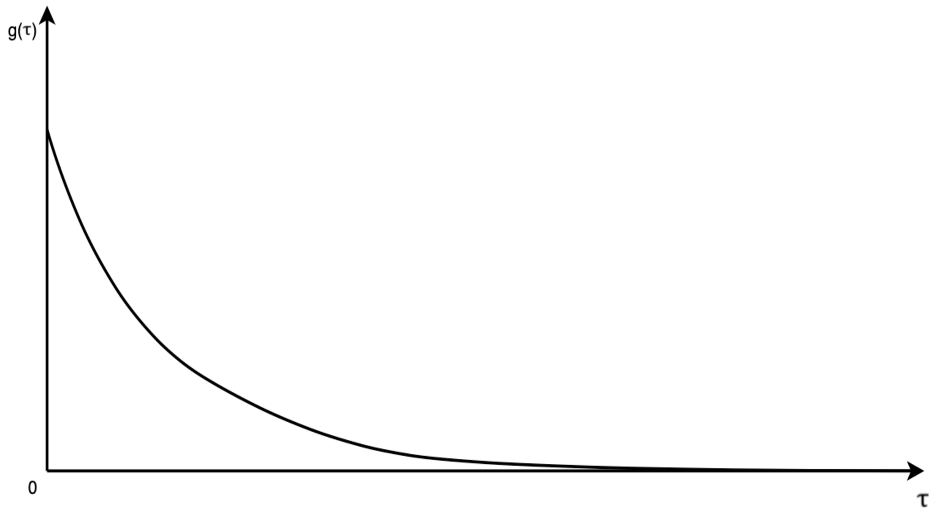

Рис. 4. Вид функции плотности вероятности распределения «обратного времени»

Определение экспертного времени S не вызывает затруднений и является, как правило, достаточно надежной величиной. Значение интенсивности λ как величины, обратной среднему времени тестирования, не обладает подобной устойчивостью, поскольку степень готовности студентов к тестированию различна. Проверим гипотезу о том, что время, оставшееся до конца тестирования, при уровне значимости 0,05 распределено по показательному закону, т.е. функция $g(\tau) = \lambda e^{-\lambda\tau}$ может быть использована для описания данной величины (случайная величина τ – «обратное время»). Проверку выполним по классическому критерию согласия χ^2 . Результаты вычислений сведены в таблицу 4.

Таблица 4

Наблюдаемые и теоретические частоты, полученные в ходе эксперимента

Номер интервала	n_i	n_i'	$\frac{(n_i - n_i')^2}{n_i'}$
1	17	21,637	0,994
2	16	15,554	0,013
3	11	11,165	0,002
4	9	8,085	0,104
5	8	5,775	0,857
6	8	4,235	3,308
7	2	2,926	0,293
8	2	2,233	0,024
9	4	5,39	0,358
	n = 77	n_i' = 77	$\chi^2_{набл.}$ = 5,953

Теоретические значения частот n_i' вычислялись по формуле $n_i' = n \cdot P_i$, где n – объем выборки, P_i – относительная частота соответствующего разряда экспоненциальной функции $g(\tau) = \lambda e^{-\lambda\tau}$; для нахождения выборочной средней $\overline{x}_в$ в качестве «представителя» *i*-го интервала принималась его середина $x_i^* = (\tau_i + \tau_{i+1})/2$:

– выборочная средняя, где x_i^* – середина интервалов таблицы 3:

$\overline{x}_в = (17 \cdot 0{,}125 + 16 \cdot 0{,}375 + 11 \cdot 0{,}625 + 9 \cdot 0{,}875 + 8 \cdot 1{,}125 + 8 \cdot 1{,}375 + 2 \cdot 1{,}625 + 2 \cdot 1{,}875 + 4 \cdot 2{,}125)/77 =$ 0,76;

– оценка параметра предполагаемого показательного распределения:

$$\lambda = 1 / \overline{x}_в = 1 / 0{,}76 = 1{,}32.$$

Таким образом, дифференциальная функция предполагаемого показательного распределения имеет вид $f(\tau) = 1{,}32 \cdot e^{-1{,}32\tau}$ $(\tau > 0)$.;

– вероятности попадания τ в каждый из интервалов: $P_i = e^{-\lambda\tau_i} - e^{-\lambda\tau_{i+1}}$. Например, для первого интервала $P_1 = P(0 < \tau < 0{,}25) = e^{-1{,}32 \cdot 0} - e^{-1{,}32 \cdot 0{,}25} = 1 - 0{,}719 = 0{,}281$. Аналогично вычисляются вероятности в остальных интервалах: $P_2 = 0{,}202$, $P_3 = 0{,}145$, $P_4 = 0{,}105$, $P_5 =$ 0,075,

$P_6 = 0{,}055$, $P_7 = 0{,}038$, $P_8 = 0{,}029$, $P_9 = 0{,}07$;

– теоретические частоты $n_i^{'} = 77 \cdot P_i$, $i = \overline{1,9}$: $n_1^{'} = 21{,}637$, $n_2^{'} = 15{,}554$, $n_3^{'} = 11{,}165$, $n_4^{'} = 8{,}085$, $n_5^{'} = 5{,}775$, $n_6^{'} = 4{,}235$, $n_7^{'} = 2{,}926$, $n_8^{'} = 2{,}233$, $n_9^{'} = 5{,}39$.;

– определяем значение критерия $\chi^2_{.}$ по формуле :

$$\chi^2_{.\ набл.} = \sum_{i=1}^{n} \frac{(n_i - n_i^{'})^2}{n_i^{'}} = 5{,}953;$$

– по таблице критических точек распределения χ^2, по заданному значению α =0,05 и числу степеней свободы $k = s - 2 = 9 - 2 = 7$ находим $\chi^2_{крит.} = \chi^2(0{,}05;7) = 14{,}1$.

Так как $\chi^2_{набл.} < \chi^2_{крит.}$, то нет оснований отвергнуть гипотезу о показательном распределении «обратного времени».

Результаты исследования, выводы. Итак, при подготовке к тестированию студентов исходя из полученных результатов необходимо действовать по следующему алгоритму:

– эксперт (преподаватель) оценивает ***S*** – сложность (трудоемкость в мин/раб) теста;

– продолжительность теста V полагается равной $3S$: $V = 3S$.

Как показывает опыт, полученную эвристическую формулу $V = 3S$ (обоснованная продолжительность тестирования) можно использовать только в ограниченном диапазоне измерений S. При $S > 25$ мин результат тестирования ухудшается: начинает действовать новый фактор – усталость студента. В нашем случае $S = 8$ мин.

Список литературы

1. Печеный Е.А., Старыгина С.Д. Дидактическая инженерия: модель построения оптимального расписания для поточного тестирования // Образовательные технологии и общество. 2017. № 4. С. 430-442.
2. Хузиахметова А.Р. Математическая подготовка студентов в метрическом компетентностном формате // Образовательные технологии и общество. 2014. № 4. С. 636-644.
3. Чошанов М.А. Дидактика и инженерия. М.: БИНОМ. Лаборатория знаний, 2011. 248 с.
4. Беспалько В.П. Педагогика и прогрессивные технологии обучения. М.: Изд-во Института профессионального образования МО России, 1995. 342 с.
5. Гмурман В.Е. Теория вероятностей и математическая статистика. Учебное пособие для вузов. М.: Изд-во Юрайт, 2012. 478 с.