УДК 51-7

В.М. Карнаухов
Московский государственный университет природообустройства

СТАТИСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ИНТЕРНЕТ-ЭКЗАМЕНА, ПРОВОДИМОГО В РАМКАХ АТТЕСТАЦИИ ВУЗА

Исследуется зависимость результатов интернет-экзамена от его параметров. Такой экзамен проводится в учебных заведениях в рамках аттестации вузов. В работе предлагается моделирование интернет-экзамена при помощи метода Монте-Карло. В качестве параметров экзамена рассмотрены уровень подготовленности группы, разнородность группы, уровень трудности теста, разнородность заданий теста, число студентов в группе, порог аттестации группы.

Ключевые слова: интернет-экзамен, метод Монте-Карло, аттестация, статистическое моделирование, тестирование.

Интернет-экзамен как часть аттестации вузов

В 2012 г. в МГУПриродообустройства проводилась аттестация университета. Аттестация вуза проводилась по многим параметрам, в числе которых был уровень подготовленности учащихся по многим фундаментальным дисциплинам. Математика – одна из таких фундаментальных дисциплин. Приведем описание тестирования студентов по математике в рамках аттестации вуза.

1. Вначале определяются учебные группы, которые должны пройти аттестацию. Эти группы выбирались равномерно по всем существующим факультетам. При аттестации вуза в тестировании принимали участие около 10 учебных групп.

2. На специальном известном сайте выкладываются примерные варианты теста. Каждый студент из выбранных групп имел возможность ознакомиться с задачами теста по математике, осуществить предварительную подготовку по дисциплине (табл. 1).

3. Задачи теста делятся по темам: по математике было 6 тем (табл. 2). Как видно из таблицы, по каждой теме предлагается от 4 до 10 задач. Подобные задачи студент получает во время тестирования на дисплее компьютера. Эти задачи могут быть перемешаны во время тестирования.

4. Для того чтобы студенту пройти аттестацию по математике, ему необходимо решить не менее половины задач по каждой теме. Эта специфика тестирования вводила в заблуждение многих студентов: при решении 92 % $(35/38 \cdot 100\%)$ всех задач теста учащийся мог и не пройти аттестацию, выполнив, например, по теме «Дифференциальные уравнения» всего лишь одну задачу из четырех (табл. 3).

5. В конце тестирования группы фиксировалось количество учащихся, не прошедших аттестацию. Если число таких студентов составляет процент, больший чем pr ($pr \approx 50$ %), то группа считается не прошедшей аттестацию, т.е. устанавливается факт отрицательной аттестации группы. В противном случае устанавливается факт положительной аттестации группы.

6. Если хотя бы одна из выбранных групп не проходит аттестацию, считается, что вуз по этому параметру не прошел аттестацию.

Описанное выше тестирование зависит от многих параметров, среди которых такие, как:

1) уровень подготовленности учащихся в тестируемой группе;

2) разнородность группы, определяемая диапазоном изменения уровня подготовленности;

3) уровень трудности заданий теста;

4) разнородность заданий теста, определяемая диапазоном изменения уровня трудности заданий;

5) количество учащихся в тестируемой группе;

6) порог аттестации группы pr (см. выше).

Целью данной работы является изучение зависимости вероятности положительной аттестации группы от вышеперечисленных параметров при помощи имитационного моделирования, использующего метод Монте-Карло.

Моделирование аттестационного тестирования при помощи метода Монте-Карло

Имитационное моделирование в представленной работе основано на использовании математической теории тестирования известного датского математика Г. Раша [1, 2]. Вкратце изложим основные положения этой теории, которые использовались в данной работе.

Естественно полагать, что успех участника тестирования в решении определенного тестового задания зависит, в основном, от двух факторов: трудности задания и подготовленности испытуемого. Таким образом, вероятность того, что определенный участник тестирования с уровнем подготовленности s верно решит определенное задание с уровнем трудности t, представляет некоторую функцию $p=p(s,t)$, которая называется функцией успеха. Для этой функции была выведена формула

$$p=\frac{s}{s+t},\quad \text{где}\quad s\in(0,\infty),\quad t\in(0,\infty).$$

Переменные s и t принято называть латентными (ненаблюдаемыми, т.е. недоступными для непосредственного измерения) параметрами, поскольку они призваны описывать некоторые скрытые характеристики участников тестирования и тестовых заданий.

На практике аргументы s и t удобно выражать в логарифмическом масштабе. Для этого вводятся переменные θ и δ по формулам

$$\theta=\ln(s)\in(-\infty,\infty),\quad \delta=\ln(t)\in(-\infty,\infty).$$

При этом функция успеха примет вид:

$$p=\frac{1}{1+\exp[-(\theta-\delta)]}.$$

Эта формула называется основной логистической моделью Раша, в которой аргументы θ и δ измеряются одной и той же шкалой с единицей измерения один логит.

Исходя из практической целесообразности, вероятность наступления некоторого события вычисляется с точностью до 3 знаков после запятой. Поэтому вероятность, равную 0,9999, можно принять за вероятность достоверного события. Решая уравнение

$$0{,}9999=\frac{1}{1+\exp[-(\Theta-\delta)]} \tag{1}$$

относительно величины $(\theta-\delta)$, получим, что $\theta-\delta=9{,}21$. Задача будет решена достоверно, если учащийся имеет максимально возможный уровень подготовленности

$\theta=\theta_{max}$, а задание – минимально возможный уровень трудности $\delta=\delta_{min}$. Учитывая, что измерение уровней подготовленности и уровней трудности осуществляется на одной симметричной шкале логитов ($\delta_{min}=-\theta_{max}$), приходим к равенству: $2\theta_{max}=9{,}21$. Отсюда получаем $\theta_{max}=4{,}6$, что позволяет считать, что значения латентных параметров реально меняются в пределах

Таблица 1

Требования ГОС к обязательному минимуму содержания основной образовательной программы

ИНТЕРНЕТ-ЭКЗАМЕН В СФЕРЕ ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
Специальность: 190207.65 – Машины и оборудование природообустройства и защиты окружающей среды.
Дисциплина: Математика
Время выполнения теста: 80 минут.
Количество заданий: 38.
ТРЕБОВАНИЯ ГОС К ОБЯЗАТЕЛЬНОМУ МИНИМУМУ СОДЕРЖАНИЯ ОСНОВНОЙ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Индекс	Дисциплина и ее основные разделы	Всего часов
ЕН.Ф	Федеральный компонент	1580
ЕН.Ф.01	Математика: аналитическая геометрия и линейная алгебра; дифференциальное и интегральное исчисления; векторный ангализ и элементы теория поля; гармонический анализ; дифференциальные уравнения; уравнения математической физики; функции комплексного переменного; численные методы; основы вычислительного эксперимента; элементы функционального анализа; элементы дискретного анализа; вероятность и статистика: теория вероятностей, случайные процессы, статистическое оценивание и проверка гипотез, статистические методы обработки экспериментальных данных; вариационное исчисление и оптимальное управление	700

от −4,6 до 4,6. В вышеупомянутой работе [2] этот интервал расширен до интервала (−5; 5). Поэтому связь между шкалой логитов и 100 % шкалой осуществлялась по формуле

$$B = \frac{\theta + \theta_{max}}{2 \cdot \theta_{max}} \cdot 100\%,$$

где θ – уровень подготовленности по шкале логитов, B – уровень подготовленности по 100 % – шкале , $\theta_{max} = 5$.

Для имитационного моделирования укажем диапазоны изменения вышеперечисленных параметров аттестационного тестирования.

1) Уровень подготовленности группы $\theta_{ср}$ определяется математическим ожиданием уровня подготовленности θ учащихся группы [1, 2]. Используя линейный способ перевода шкалы логитов в 100 % шкалу, будем измерять введенный параметр в процентах. Эти единицы измерения

Таблица 2

Тематическая структура теста

ТЕМАТИЧЕСКАЯ СТРУКТУРА АПИМ			
№ ДЕ	Наименование дидактической единицы ГОС	№ задания	ТЕМА ЗАДАНИЯ
1	Алгебра и геометрия	1 2 3 4 5 6	Линейные операции над матрицами. Системы линейных уравнений, основные понятия. Прямая на плоскости. Прямая и плоскость в пространстве. Скалярное произведение векторов. Коллинеарность и перпендикулярность векторов
2	Математический анализ	7 8 9 10 11 12 13 14 15 16	Функции: основные понятия и определения. Предел функции. Производные первого порядка. Приложения дифференциального исчисления ФОП. Дифференциальное исчисление ФНП. Свойства определенного интеграла. Приложения определенного интеграла. Элементы теории множеств. Мера плоского множества. Отображение множеств
3	Теория функций комплексного перменного	17 18 19 20 21 22	Операции над комплексными числами. Определение функции комплексного переменного. Дифференцирование функции комплексного переменного. Периодические функции. Элементы гармонического анализа. Ряд Фурье. Теорема Дирихле
4	Дифференциальные уравнения	23 24 25 26	Типы дифференциальных уравнений. Дифференциальные уравнения первого порядка. Дифференциальные уравнения высших порядков. Линейные дифференциальные уравнения второго порядка
5	Теория вероятностей и математическая статистика	27 28 29 30 31 32	Теоремы сложения и умножения вероятностей. Полная вероятность. Формула Байеса. Дискретная случайная величина. Характеристики вариационного ряда. Точечные оценки параметров распределения. Интервальные оценки параметров распределения
6	Вычислительная математика, дискретная математика	33 34 35 36 37 38	Операции над высказываниями. Элементы теории множеств. Элементы комбинаторики. Численные методы решения алгебраических уравнений. Численные методы анализа. Численные методы решения дифференциальных уравнений

более привычны нам по оценкам ЕГЭ. Диапазон изменения этого параметра: $30 \le \theta_{ср} \le 70$. Значение $\theta_{ср} = 50$ будем считать средним значением.

2) Абсолютное значение максимального отклонения $\Delta\theta_{max}$ уровня подготовленности учащихся группы от величины $\theta_{ср}$. Можно назвать этот параметр неоднородностью группы. Единицами измерения этого параметра будут проценты (см. п. 1). Диапазон изменения этого параметра: $0 \le \Delta\theta_{max} \le 30$. Значение $\Delta\theta_{max} = 20$ будем считать средним значением.

3) Уровень трудности теста $\delta_{ср}$ определяется математическим ожиданием уровня трудности δ задания теста [1, 2]. Используя линейный способ перевода шкалы логитов в 100 % шкалу, будем измерять введенный параметр в процентах. Диапазон изменения этого параметра: $30 \le \delta_{ср} \le 70$. Значение $\delta_{ср} = 50$ будем считать средним значением.

4) Максимальное отклонение $\Delta\delta_{max}$ уровня трудности задания теста от величины $\delta_{ср}$. Можно назвать этот параметр неоднородностью теста. Единицами измерения этого параметра будут проценты (см. п. 3). Диапазон изменения этого параметра: $0 \le \Delta\delta_{max} \le 30$. Значение $\Delta\delta_{max} = 5$ будем считать средним значением.

Таблица 3

Задачи по теме «Дифференциальные уравнения»

Задача № 23 (выберите несколько вариантов ответа)

Среди перечисленных дифференциальных уравнений уравнениями первого порядка являются:

Варианты ответов:

1) $2x\dfrac{d^2y}{dx^2} + x\dfrac{dy}{dx} + y = 0.$ 2) $y^2\dfrac{\partial y}{\partial x} + x = 0.$

3) $x^2y' + 8y - x + 5 = 0.$ 4) $x\dfrac{d^2y}{dx^2} + xy\dfrac{dy}{dx} + x^2 = y.$

Задача № 24 (выберите один вариант ответа)

Общий интеграл дифференциального уравнения $\dfrac{dy}{y^2} = \dfrac{dx}{1+x^2}$ имеет вид…

Варианты ответов:

1) $-\dfrac{1}{y} = arctg\dfrac{1}{x} + C.$ 2) $\dfrac{1}{y} = -\ln(1+x^2) + C.$

3) $-\dfrac{1}{y} = arctgx + C.$ 4) $\dfrac{1}{y} = \ln(1+x^2) + C.$

Задача № 25 (выберите один вариант ответа)

Общее решение дифференциального уравнения $y''' = 2x + 1$ имеет вид …..

Варианты ответов:

1) $y = \dfrac{1}{12}x^4 + \dfrac{1}{6}x^3 + C.$ 2) $y = \dfrac{1}{24}x^4 + \dfrac{1}{6}x^3 + \dfrac{C_1}{2}x^2 + C_2x + C_3.$

3) $y = \dfrac{1}{12}x^4 + \dfrac{1}{6}x^3 + \dfrac{C_1}{2}x^2 + C_2x + C_3.$ 4) $y = x^4 + x^3 + C_1x^2 + C_2x + C_3.$

Задача № 26 (выберите один вариант ответа)

Дано дифференциальное уравнение $y'' - 4y' - 5y = 2e^{5x}$. Общим видом частного решения данного уравнения является…

Варианты ответов:

1) $y(x)_{частное} = C_0\cos 5x + C_1\sin 5x.$ 2) $y(x)_{частное} = C_0 + C_1x.$

3) $y(x)_{частное} = C_0xe^{5x}.$ 4) $y(x)_{частное} = C_0e^{5x}.$

5) Число n студентов в группе, принимающей участие в аттестации. Диапазон изменения этого параметра: $10 \le n \le 35$. Значение $n = 25$ будем считать средним значением.

6) Порог аттестации pr, представляющий собой процент такой, что если процент не прошедших аттестацию учащихся группы будет выше pr, то группа получает отрицательный аттестационный результат. Диапазон изменения этого параметра:

$50 \le pr \le 80$. Значение $pr = 50$ будем считать средним значением.

При исследовании зависимости результатов аттестации от вышеперечисленных параметров использовалась следующая гипотеза: уровень подготовленности учащегося группы и уровень трудности задания теста распределены по нормальному закону.

Используя выдвинутую гипотезу, а именно правило 3 сигм, можно утверждать, что

$\theta \approx N(\theta_{ср}, \Delta\theta_{max}/3)$,

$\delta \approx N(\delta_{ср}, \Delta\theta_{max}/3)$.

На основании этой гипотезы в программе создается одномерный массив уровней подготовленности учащихся, а также двумерный массив уровней трудности заданий теста. Для этого используется метод Монте-Карло, неоднократно использованный автором при исследовании различных сегментов учебного процесса [4, 5]. Этот метод позволяет моделировать на ЭВМ случайные величины, распределенные по заданному закону. Согласно выдвинутой гипотезе случайные величины θ и $\delta(i)$ (i – номер темы) распределены по нормальному закону (см. выше). Поэтому реализации этих случайных величин могут быть найдены исходя из следующих выкладок.

Если случайная величина θ распределена по нормальному закону $N(a, \sigma)$ с функцией распределения $F(x)=F_0((x-a)/\sigma)$, где $F_0(t)$ – функция распределения нормированного нормального распределения $N(0,1)$, то случайная величина $R=F(\theta)$ распределена равномерно на интервале (0, 1). Решая уравнение $R = F_0((\theta-a)/\sigma)$ относительно неизвестного θ, получим:

$\theta = F_0^{-1}(R) \cdot \sigma + a$,

где $F_0^{-1}(R)$ – обратная функция к функции $F_0(t)$.

Последняя формула позволяет находить реализации случайных величин θ и σ при помощи датчика случайных чисел R на (0, 1).

Для моделирования матрицы ответов учащихся использовалась формула (1) по следующей схеме: учащийся с уровнем подготовленности θ решает задание с уровнем трудности σ, если для очередной реализации r датчика случайных чисел выполняется следующее неравенство $r < p = 1/(1 + \exp(-(\theta - \delta)))$. В случае выполнения противоположного неравенства считается, что учащийся не решил задание.

На основании матрицы ответов устанавливается факт отрицательной или положительной аттестации группы. Вероятность положительной аттестации группы вычисляется при помощи проведения 1 000 статистических смоделированных испытаний.

Аттестационные области

Определим в области допустимых значений параметров системы (ОДЗПС) аттестационные области 4 типов: критическая, неудовлетворительная, удовлетворительная, положительная.

Будем считать, что учебное заведение успешно прошло аттестацию по тестированию, если все тестируемые группы прошли аттестацию. В этом случае будем говорить, что аттестация учебного заведения по тестированию студентов привела к положительному результату. В противном случае будем говорить, что учебное заведение не прошло аттестацию или аттестация учебного заведения привела к отрицательному результату.

Учитывая, что тестирование осуществляется примерно в 10 учебных группах, можно вычислить вероятность $P_{ат}$ положительного результата аттестации по формуле

$P_{ат} = p^{10}$,

где p – вероятность получения положительного результата при тестировании группы.

Разделим ОДЗПС на четыре области в зависимости от значения p.

Если $0 \le p < 0.63$, то $P_{ат} < 0.01$ (1 %). В этом случае можно считать достоверным событие отрицательного результата аттестации учебного заведения.

Определение 1. Область в ОДЗПС называется критической, если для любого набора параметров из этой области $0 \le p < 0.63$.

2) Если $0.63 \le p < 0.93$, то $0.01 \le P_{ат} < 0.5$ (от 1 до 50 %). В этом случае аттестация может привести к положительному результату, а может и не привести. Причем вероятность отрицательного результата выше, нежели положительного. Мож-

но сказать, что скорее всего в этом случае аттестация приведет к отрицательному результату.

Определение 2. Область в ОДЗПС называется неудовлетворительной, если для любого набора параметров из этой области $0.63 \le p < 0.93$.

3) Если $0.93 \le p < 1$, то $0.5 \le P_{ат} < 1$ (от 50 до 100 %). В этом случае аттестация может привести к положительному результату, а может и не привести. Причем вероятность положительного результата выше, нежели отрицательного. Можно сказать, что скорее всего в этом случае аттестация приведет к положительному результату.

Определение 3. Область в ОДЗПС называется удовлетворительной, если для любого набора параметров из этой области $0.93 \le p < 1$.

Если $p = 1$, то $P_{ат} = 1$. В этом случае в любом случае аттестация приведет к положительному результату.

Определение 4. Область в ОДЗПС называется положительной, если для любого набора параметров из этой области $p = 1$.

Значимость параметров аттестационного тестирования

Определенные выше параметры системы тестирования обладают разной степенью влияния на изменение вероятности p. На рис. 1–4 приведены графики зависимости вероятности успешной аттестации вуза от каждого параметра при фиксированных средних значениях других 5 параметров.

Как видно из рис. 1 и 2, изменение параметров системы $\theta_{ср}$, $\delta_{ср}$ в диапазоне возможных значений приводит к изменению вероятности аттестации вуза от 0 до 1. Поэтому параметры $\theta_{ср}$, $\delta_{ср}$ можно считать значимыми для вероятности p.

Для исследования значимости неоднородности по уровню подготовленности и уровню трудности заданий теста среднее значение уровня подготовленности было заменено на значения $\theta_{ср}$=55 и $\theta_{ср}$=60.

Как видно из рис. 3, изменение параметров системы $\Delta\theta_{max}$, $\Delta\delta_{max}$ в диапазоне возможных значений практически не изменяет величину p. Поэтому можно считать параметры $\Delta\theta_{max}$, $\Delta\delta_{max}$ не значимыми для вероятности p. При этом из рис. 3 наблюдается небольшое уменьшение вероятности p с ростом неоднородности.

Для исследования значимости числа студентов в группе среднее значение уровня подготовленности было заменено на значения $\theta_{ср}$=55 и $\theta_{ср}$=60.

Как видно из рис. 4, значимость числа студентов выше значимости параметров $\Delta\theta_{max}$, $\Delta\delta_{max}$, но не выше значимости параметров $\theta_{ср}$, $\delta_{ср}$. Изменение числа студентов может привести к переходу набора параметров из одной аттестационной области в другую. Из рис. 4 также видно:

1) в группах, средний уровень подготовленности $\theta_{ср}$ которых равен 60, с ростом числа n студентов в группе, вероятность положительной аттестации все-таки незначительно увеличивается.

2) в группах, средний уровень подготовленности $\theta_{ср}$ которых равен 55, с ростом числа n студентов в группе вероятность положительной аттестации значительно уменьшается.

Поэтому можно сделать следующий вывод: с ростом среднего уровня подготовленности группы меняется ориентация монотонности зависимости вероятности p от числа студентов (с убывания на возрастание).

Для исследования значимости порога аттестации, помимо среднего значения уровня подготовленности, были использованы его значения $\theta_{ср}$=55 и $\theta_{ср}$=60.

Как видно из рис. 5, значимость порога аттестации выше значимости параметров $\Delta\theta_{max}$, $\Delta\delta_{max}$, но не выше значимости параметров $\theta_{ср}$, $\delta_{ср}$. Изменение порога аттестации может привести к переходу набора параметров из одной аттестационной области в другую. Из рис. 5 также видно, что при любом среднем уровне подготовленности группы с ростом порога аттестации увеличивается значение вероятности p.

Влияние параметров аттестации на ее результаты

Определенные выше аттестационные области для постоянного набора параметров: n=25, pr=50 %, $\delta_{ср}$ = 50 %, $\Delta\delta_{max}$ = 5 % (средние значения параметров системы) – изображены на рис. 6.

Точные значения вероятности p в зависимости от изменения параметров $\theta_{ср}$ и $\Delta\theta_{max}$ даются в табл. 4.

Перед тем как сделать выводы из рис. 6 и табл. 4, сделаем следующие замечания.

1) Учебные группы обычно бывают разнородными (есть «отличники», есть и «двоечники»). Разнородность, оцениваемая в 10%, соответствует примерно одному баллу по пятибалльной шкале; исходя из этих данных, при $\Delta\theta_{max}$<10 в группе будут только учащиеся, которые учатся примерно

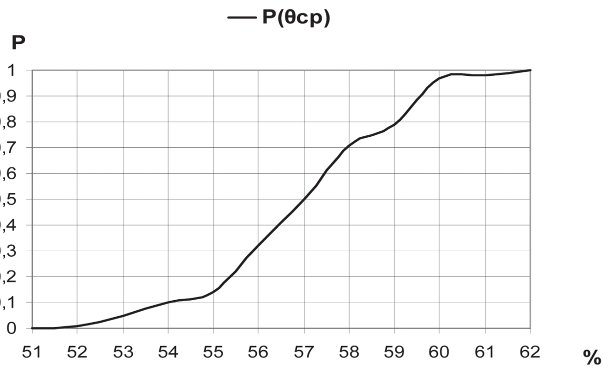

Рис. 1. Значимость среднего уровня подготовленности группы

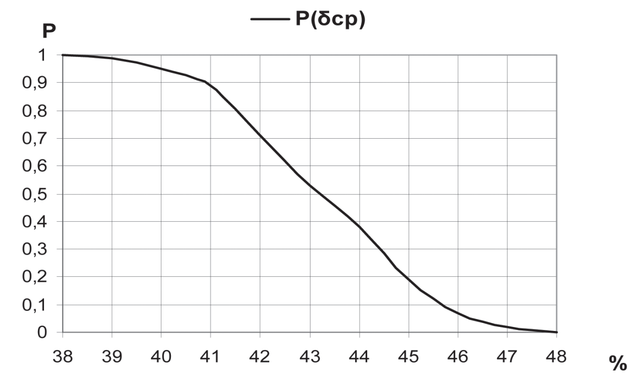

Рис. 2. Значимость среднего уровня трудности заданий теста

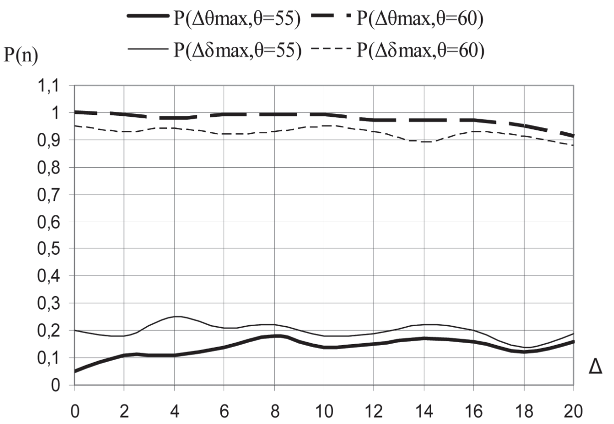

Рис. 3. Значимость неоднородности группы и неоднородности теста

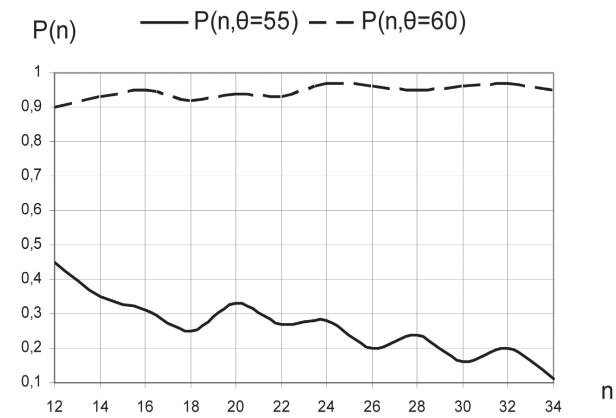

Рис. 4. Значимость числа студентов в группе на результаты аттестации

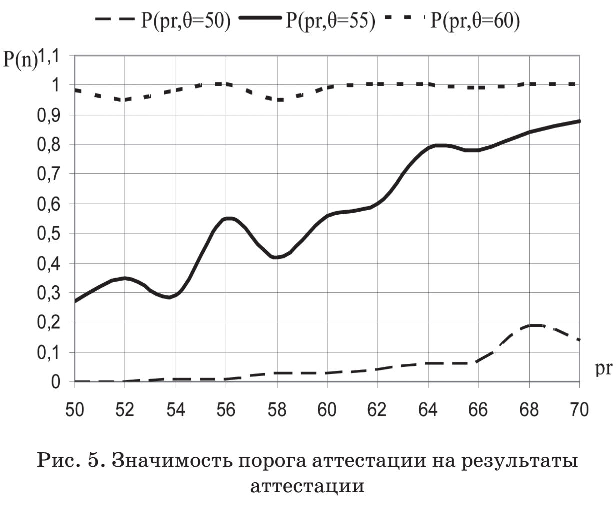

Рис. 5. Значимость порога аттестации на результаты аттестации

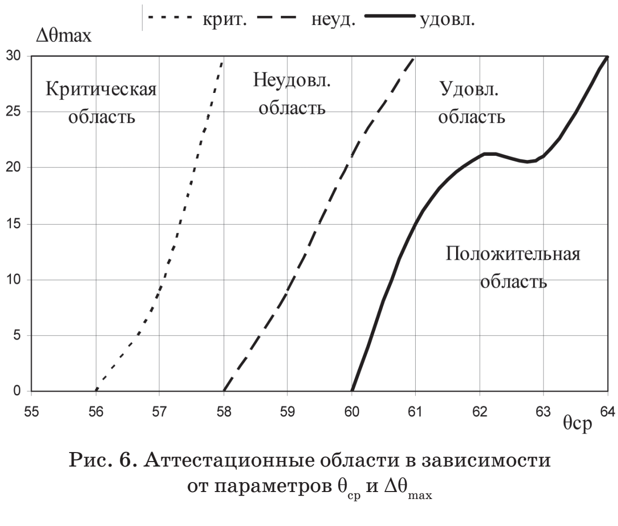

Рис. 6. Аттестационные области в зависимости от параметров $\theta_{ср}$ и $\Delta\theta_{max}$

Таблица 4

Значения вероятности p в зависимости от параметров $\theta_{ср}$ и $\Delta\theta_{max}$

$\theta_{ср}$ / $\Delta\theta_{max,1}$	0	3	6	9	12	15	18	21	24	27	30
55	0,13	0,12	0,11	0,07	0,18	0,14	0,22	0,26	0,16	0,23	0,29
56	0,36	0,41	0,34	0,29	0,31	0,34	0,34	0,32	0,37	0,4	0,4
57	0.74	0,57	0,64	0,63	0,59	0,55	0,59	0,55	0,58	0,56	0,47
58	0,9	0,86	0,86	0,81	0,81	0,79	0.72	0,75	0,61	0,66	0,75
59	0,96	0,97	0,97	0,94	0,91	0,92	0,89	0,84	0,85	0,81	0,74
60	0,99	1	0,99	0,97	0,96	0,98	0,98	0,96	0,89	0,89	0,87
61	1	1	1	1	1	1	0,99	1	0,95	0,93	0,9
62	1	1	1	1	1	0,99	1	1	0,98	0,99	0,99
63	1	1	1	1	1	1	1	1	0,99	0,99	0,97
64	1	1	1	1	1	1	1	1	1	1	1

на один и тот же балл. Такое явление наблюдается крайне редко.

2) Если принять 30 баллов за минимальную оценку, при которой можно поступить в вуз, то для непрестижных вузов ($\theta_{ср} \leq 50$) наибольшее отклонение $\Delta\theta_{max}$ не может быть больше 20.

Выводы:

1) при $\theta_{ср} \leq 56$ аттестация учебного заведения приведет к отрицательному результату;

2) при $\theta_{ср} \geq 64$ аттестация учебного заведения приведет к положительному результату;

3) если предположить, что $\Delta\theta_{max}$ обычно бывает не меньше 10 (замечание 1), то аттестация учебного заведения приведет к положительному результату при $\theta_{ср} \geq 62$.

Сделаем основные выводы из проведенных исследований.

1) Если определенные выше параметры системы расположить в порядке убывания их значимости для вероятности *p* положительной аттестации группы учащихся, то получится следующая последовательность:

а) средний уровень подготовленности учащихся группы и средний уровень трудности заданий теста (наибольшая значимость),

б) число студентов в группе и порог аттестации (средняя значимость),

в) неоднородность группы и неоднородность теста (низкая значимость).

2) Престижные вузы ($\theta_{ср} \geq 62$) будут успешно проходить аттестацию.

3) Аттестация непрестижных вузов ($\theta_{ср} \leq 56$) всегда будет приводить к отрицательному результату.

ЛИТЕРАТУРА

1. *Rasch G.* Probabilistic Models for Some Intelligence and Attainment Tests. – Copengagen Denmark: Danish Institute for Educational Research, 1968.

2. *Нейман Ю.М., Хлебников В.А.* Введение в теорию моделирования и параметризации педагогических тестов. – М., 2000. – 169 с.

3. *Карнаухов В.М.* Использование метода Монте-Карло в теории моделирования и параметризации педагогических тестов // Информатика и образование. – 2009. – № 7. – С. 109–111.

4. *Карнаухов В.М.* Электронное тестирование с двумя и более попытками для решения одного задания. – М.: ФГБОУ ВПО, 2011.

V.M. Karnaukhov
Moscow State University of Environmental Engineering, Russia

STATISTICAL MODELING OF INTERNET EXAMINATION CONDUCTED IN THE FRAMEWORK OF HIGHER SCHOOL CERTIFICATION

Key words: the Internet-exam, Monte-Carlo method, certification, statistical modeling, testing.

Currently many Universities of Russia are undergone the procedure of certification which is carried out in several directions including the test of students of their knowledge qualification level on fundamental disciplines. Mathematics is one of the fundamental disciplines. We present a description of students' testing in mathematics in the framework of certification of higher education institution.

1) First, the learning groups put to the certification are defined. These groups were selected evenly from all the faculties. Thus, ten groups took part in the testing during the university certification.

2) The modal patterns of the test were laid out in the special website. Each student from the selected groups

had the opportunity to familiarize with the tasks on mathematics and practice in the course.

3) The tasks of the test were categorized by the topics: 6 topics on mathematics. As the table shows, there are from 4 to 10 tasks for each topic. The student solves similar tasks during the test which are displayed on the computer screen. These tasks can be shuffled during the test.

4) The student should solve at least a half of all the tasks to pass the mathematics test. This test specificity has misled a number of students. Having solved 92% ($35/38 \cdot 100\%$) of all the tasks the student could not pass the test on condition that he solved, for example, only one task of the four on the topic "Differential equations".

5) In the end of the test the students who did not pass the test were recorded. If the number of those students was greater than pr (pr $\approx 50\%$), the group was considered not to pass the certification, and the fact of negative certification in the group was determined. Otherwise, the fact of positive certification of the group is settled.

6) If at least one of the selected groups will not pass the test, it is believed that the university did not pass the certification.

The testing depends on many parameters, such as:

1) knowledge qualification level (θcp) of students in the tested group;

2) heterogeneity of the group, determined by the range of knowledge qualification level;

3) level of test tasks difficulty;

4) heterogeneity of the test tasks, determined by the range of changes of difficulty level;

5) number of students in the tested group,

6) certification threshold of the pr group (see above).

The objective of this work is to study the dependence of probability of positive certification of the group from the above parameters with the help of simulation using the Monte Carlo method.

The following results are obtained:

1) If the determined parameters of the system be arranged in descending order of their importance for the probability p of positive certification of the students' group, the result will be in the following sequence:

a) the average level of the knowledge qualification level of students and the average level of difficulty of the test tasks (the highest value),

b) the number of students in the group and the threshold certification (medium importance),

c) the heterogeneity of the group and the heterogeneity of the test (low importance).

2) Prestigious universities (θav ≥ 62) will successfully pass the certification.

3) Certification of low-prestige universities (θav ≤ 56) will always lead to negative results.

REFERENCES

1. *Rasch G.* Probabilistic Models for Some Intelligence and Attainment Tests. – Copengagen Denmark: Danish Institute for Educational Research, 1968.

2. *Nejman Ju.M., Hlebnikov V.A.* Vvedenie v teoriju modelirovanija i parametrizacii pedagogicheskih testov. – M., 2000. – 169 s.

3. *Karnauhov V.M.* Ispol'zovanie metoda Monte-Karlo v teorii modelirovanija i parametrizacii pedagogicheskih testov // Informatika i obrazovanie. – 2009. – № 7. – S. 109–111.

4. *Karnauhov V.M.* Jelektronnoe testirovanie s dvumja i bolee popytkami dlja reshenija odnogo zadanija. – M.: FGBOU VPO, 2011.