УДК 37.01(075.8)

МАТЕМАТИКО-СТАТИСТИЧЕСКОЕ ОБОСНОВАНИЕ ОЦЕНКИ ЗНАНИЙ СТУДЕНТА ПРИ ТЕСТИРОВАНИИ

БЫВШЕВ В. А.
доктор технических наук
БОГОМОЛОВ А. И.
кандидат технических наук
КОСТЮНИН В. И.
кандидат технических наук
Москва (Россия)

Педагогические тесты являются общепринятым инструментом обучения и контроля знаний студентов в Финуниверситете. Тесты дают возможность преподавателю осуществлять оперативную массовую оценку уровня полученных студентами знаний и навыков по соответствующей дисциплине, и в настоящее время используются при аттестации вузов[1]. Многие преподаватели вузов (в частности, Финуниверситета) используют тестовый контроль знаний на итоговых зачётах и экзаменах по дисциплинам. Чаще всего экзаменационный билет (тест) состоит из некоторого количества тестовых заданий с выбором в каждом задании одного правильного ответа из определённого числа предложенных ответов[2]. В процессе проверки результатов тестирования возникает естественная проблема объективной трансформации количества (R) правильных ответов студента в шкалу привычных оценок, например, в шкалу «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

В предлагаемой работе обсуждён формализованный подход к упомянутой выше трансформации, основанный на математико-статистической теории проверки гипотез[3]. Предлагаемый подход прошел многолетнюю (с 2004 г.) проверку на кафедре математического моделирования экономических процессов Финуниверситета и всегда проявлял объективный характер оценки знаний студентов.

ПРЕДПОСЫЛКИ И ОБОЗНАЧЕНИЯ

Ниже приняты следующие предположения.

1. Тест (билет) составлен в виде (n) заданий в тестовой форме с выбором в каждом задании одного правильного ответа из (m) предложенных ответов. Реальный пример, в котором $n = 10$ и $m = 4$, приведён в приложении.

2. В процессе тестирования проверяются следующие семантические (смысловые) гипотезы.

1) H_f: Студент совершенно не знает материал дисциплины и, следовательно, выбирает ответы наугад.

2) $\overline{H}_f$: Студент обладает определёнными знаниями и навыками и поэтому не выбирает ответы наугад.

3) H_e: Студент в полной мере освоил дисциплину и способен правильно ответить практически на любой вопрос.

Очевидно, в ситуации справедливой гипотезы 1) студент должен получить оценку «неудовлетворительно». Если же верна гипотеза 2), студент заслуживает, по крайней мере, оценку «удовлетворительно». Наконец, при справедливой гипотезе 3) преподаватель имеет основание оценить знания и навыки студента на «отлично».

3. Обозначим символом (p) вероятность правильного ответа студента на вопрос из тестового задания. Будем предполагать, что эта вероятность остаётся неизменной для всех заданий теста. Так, если имеет место равенство

$$p = p_c = \frac{1}{m}, \qquad (1)$$

то это равносильно справедливости гипотезы 1). Если же справедливо неравенство

$$p > p_c, \qquad (2)$$

то это равносильно истинности гипотезы 2). Наконец, когда упомянутая вероятность удовлетворяет неравенству

$$p \geq p_{\tilde{\Omega}} = 0,95, \qquad (3)$$

где $p_{\tilde{\Omega}} = 0,95$ – вероятность практически достоверного события [4], то такая ситуация равносильна справедливости гипотезы 3).

После завершения тестирования в распоряжении преподавателя оказывается информация о знаниях и навыках студента в виде количества (R) его правильных ответов. Множество возможных значений этой величины состоит, очевидно, из следующих констант:

$$MR = (0, 1, 2, \ldots, n). \qquad (4)$$

По наблюдённому значению R преподаватель должен оценить уровень знаний и навыков студента по стандартной шкале оценок - «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

4. Будем постулировать, что величина R является биномиальной случайной переменной, вероятностная функция которой имеет вид[5]

$$P_R(k \mid p) = P(R = k) = C_n^k \cdot p^k \cdot (1-p)^{n-k}, \quad k = 0,1,2\ldots,n. \qquad (5)$$

---

[1] http://www.fepo.ru

[2] Аванесов В. С. Современные методы обучения и контроля знаний / В. С. Аванесов. – М. : Наука, 1998. – С. 9 – 10..

[3] Нейман Ю. Вводный курс теории вероятностей и математической статистики / Ю. Нейман. – М. : Наука, 1968. – С. 320.

[4] Бывшев В. А. Эконометрика / В. А. Бывшев. – М. : Финансы и статистика, 2008.. – С. 69.

Данный постулат позволяет придать упомянутым выше семантическим гипотезам 1), 2) и 3) характер статистических гипотез[6] и, в итоге, объективизировать трансформацию величины R в оценку знаний и навыков студента по стандартной шкале. Конкретно, гипотезы 1), 2) и 3) принимают характер следующих статистических гипотез о параметре (p) закона распределения (5) случайной переменной R:

$$\begin{aligned} &H_f : p = p_c = \frac{1}{m}, \\ &\bar{H}_f : p > p_c, \\ &H_e : p = p_{\tilde{\Omega}} = 0{,}95. \end{aligned} \qquad (6)$$

Результаты проверки этих гипотез и позволяют объективизировать трансформацию величины R в оценку знаний и навыков студента по стандартной шкале.

ПРОЦЕДУРА ПРОВЕРКИ СТАТИСТИЧЕСКИХ ГИПОТЕЗ (6) И ТРАНСФОРМАЦИЯ КОЛИЧЕСТВА R ПРАВИЛЬНЫХ ОТВЕТОВ СТУДЕНТА В ОЦЕНКУ ЕГО ЗНАНИЙ И НАВЫКОВ ПО СТАНДАРТНОЙ ШКАЛЕ

Процесс проверки статистических гипотез (6) и трансформации величины R в оценку знаний и навыков студента по стандартной шкале удобно осуществить в итоге следующих шагов.

1. Задаёмся уровнем значимости α (вероятностью отвергнуть гипотезу в ситуации, когда она верна), например, α = 0, 075, и исследуем гипотезу $H_e : p = p_{\tilde{\Omega}}$ против альтернативы $\bar{H}_e : p < p_{\tilde{\Omega}}$. Стандартная наиболее мощная критическая область гипотезы H_e против альтернативы $\bar{H}_e$ имеет вид[7]

$$M_R(\bar{H}_e) = (0, 1, \ldots, T(n)). \qquad (7)$$

Правая точка этого множества вычисляется как корень уравнения

$$P(R \in M_R(\bar{H}_e) / H_e) = \sum_{k=0}^{T(n)} C_n^k \cdot p_{\tilde{\Omega}}^k \cdot (1 - p_{\tilde{\Omega}})^{n-k} = \alpha, \qquad (8)$$

согласно принятому уровню значимости α. Если величина R попадает в множество (7), то гипотеза H_e отклоняется и, следовательно, испытуемый не получает оценку «отлично». Напротив, если количество R оказывается в множестве принятия гипотезы H_e:

$$M_R(H_e) = (T(n) + 1, \ldots, n), \qquad (9)$$

то испытуемый получает оценку «отлично». Предположим, однако, что количество R оказалось в критической области (7). Для решения вопроса об оценке испытуемого нужно сделать следующий шаг.

2. Исследуем гипотезу $H_f : p = p_c = \frac{1}{m}$ против альтернативы $\bar{H}_f : p > p_c$. Стандартная равномерно наиболее мощная критическая область гипотезы H_f против альтернативы $\bar{H}_f$ имеет вид[8]

$$M_R(\bar{H}_f) = (t(n), t(n) + 1, \ldots, n). \qquad (10)$$

Левая точка $t(n)$ этого множества вычисляется как корень уравнения

$$\begin{aligned} &P(R \in M_R(\bar{H}_f) / H_f) = \\ &= \sum_{k=t(n)}^{n} C_n^k \cdot p_c^k \cdot (1 - p_c)^{n-k} = \alpha, \end{aligned} \qquad (11)$$

согласно принятому уровню значимости α. Если величина R попадает в множество (11), то гипотеза H_f отклоняется и, следовательно, испытуемый не получает оценку «неудовлетворительно». Напротив, если количество R оказывается в множестве принятия гипотезы H_f:

$$M_R(H_f) = (0, 1, .., t(n) - 1), \qquad (12)$$

то испытуемый получает оценку «неудовлетворительно». Предположим, однако, что количество R оказалось в критической области (10), а точнее, в следующем подмножестве этой области:

$$M_R(\bar{H}_f \cap \bar{H}_e) = (t(n), t(n) + 1, \ldots, T(n)). \qquad (13)$$

В этой ситуации испытуемый заслуживает либо оценку «удовлетворительно», либо оценку «хорошо», и для решения вопроса об оценке испытуемого нужно сделать последний шаг.

3. Вычисляем середину множества (13):

$$A(n) = \frac{1}{2} \cdot (t(n) + T(n)). \qquad (14)$$

Если количество R оказывается меньше величины $A(n)$, то испытуемый получает оценку «удовлетворительно»; если же количество правильных ответов R равно или больше величины $A(n)$, то студент заслуживает оценку «хорошо».

Подводя итог, сформулируем алгоритм трансформации количество R правильных ответов студента в оценку по стандартной шкале, полагая известными (см. пример) обсуждённые выше величины $t(n)$, $T(n)$ и $A(n)$.

1) Ситуация соответствует оценке «неудовлетворительно».

2) Ситуация соответствует оценке «удовлетворительно».

3) Ситуация соответствует оценке «хорошо».

4) Ситуация соответствует оценке «отлично».

ПРИМЕР

1. В приложении размещён один из реальных вариантов множества тестов по эконометрике, которые используются на кафедре ММЭП Финуниверситета в процессе итогового контроля знаний студентов. Видно, что тест составлен в виде n = 10 заданий в тестовой форме с выбором в каждом задании одного правильного ответа из m = 4 предложенных ответов. Следовательно, в множестве (4) величина n = 10. В свою очередь, вероятность выбрать наугад правильный ответ в задании p_c = 0,25.

[5] Нейман Ю. Вводный курс теории вероятностей и математической статистики / Ю. Нейман. – М. : Наука, 1968. – С. 239.
[6] Там же. – С. 321.
[7] Там же. – С. 419.
[8] Там же.

2. Определим величины $t(10)$ и $T(10)$ критических областей (7) и (10). Примем уровень значимости $\alpha = 0{,}075$ и, используя функцию БИНОМРАСП Excel, находим при $p = 0{,}95$ величину $T(10) = 8$, которая приближённо является корнем уравнения (8). Затем при $p = 0{,}25$ отыскиваем величину $t(10) = 5$, которая приближённо служит корнем уравнения (11). Наконец, вычисляем согласно (14) значение $A(10) = 6{,}5$.

3. В данном примере обсуждённый выше алгоритм трансформации количество R правильных ответов студента в оценку по стандартной шкале выглядит так:

1) ситуация $0 \le R \le 4$ соответствует оценке «неудовлетворительно»;

2) ситуация $5 \le R \le 6$ соответствует оценке «удовлетворительно»;

3) ситуация $7 \le R \le 8$ соответствует оценке «хорошо»;

4) ситуация $9 \le R \le 10$ соответствует оценке «отлично».

ПРИЛОЖЕНИЕ: ПРИМЕР ТЕСТА

13.1. Укажите номер правильного ответа.

Количество параметров в простой макромодели Кейнса, $\begin{cases} C = a_0 + a_1 \cdot Y + u \\ Y = C + I \\ 0 \le a_1 < 1;\; E(u) = 0;\; E(u^2) = \sigma_u^2 \end{cases}$, где Y – доход, C – уровень потребления, I – объём инвестиций, равно 1) единице; 2) двум; 3) трём; 4) четырём.

13.2. Укажите номер правильного ответа.

Пусть $(A, \vec{m}, Cov)$ – параметрическая модель Марковица фондового рынка и $\vec{x}$ – портфель активов из списка A. Тогда по правилу $\vec{x}^T \cdot Cov \cdot \vec{x}$ вычисляется: 1) доходность портфеля; 2) ожидаемая доходность портфеля; 3) дисперсия доходности портфеля; 4) ковариация портфеля

13.3. Укажите номер правильного ответа.

В зависимости $y = f(x) + u$ между экономическими переменными (x, y), где u – случайная переменная с нулевым ожидаемым значением, величина $f(x)$ отражает: 1) влияние неучтённых факторов; 2) экономический закон; 3) случайные причины; 4) влияние эндогенной переменной.

13.4. Укажите номер правильного ответа.

В модели годовой доходности (r) на обыкновенную акцию $\begin{cases} r = a_0 + u \\ E(u) = 0;\; E(u^2) = \sigma^2 \end{cases}$ параметр σ^2 имеет смысл: 1) ожидаемой доходности; 2) меры несистематического риска; 3) меры систематического риска; 4) ковариации.

13.5. Укажите номер правильного ответа.

В рамках оценённой инвестиционной модели $\begin{cases} I_t = 200 + 0{,}18 \cdot Y_t - 19 \cdot R_t + u \\ \quad (40) \quad (0{,}01) \quad (4) \quad (50) \\ R^2 = 0{,}82 \end{cases}$ доход (Y_t) и реальная ставка процента (R_t) не объясняют: 1) 0,01 величины I_t; 2) 4% величины I_t; 3) 50% величины I_t; 4) 18% величины I_t.

13.6. Укажите номер правильного ответа.

В рамках модели $\begin{cases} y = a_0 + a_1 \cdot x + u \\ E(u) = 0;\; E(u^2) = \sigma_u{}^2 \end{cases}$ по формулам $\left(\tilde{\sigma}_u \cdot \sqrt{\dfrac{1}{\sum_1^n (x_i - \bar{x})^2}} \right)$ вычисляется: 1) оценка $\tilde{y}$; 2) характеристика точности оценки коэффициента функции регрессии; 3) оценка σ_y; 4) оценка $\tilde{u}$.

13.7. Укажите номер правильного ответа.

Переменная u в рамках модели $\begin{pmatrix} y = a_0 \cdot \exp(a_1 \cdot x) \cdot x^{a_2} \cdot (1 + u) \\ E(u) = 0;\; E(u^2) = \sigma_u^2 \end{pmatrix}$ обладает физической размерностью: 1) переменной x; 2) переменной y; 3) нулевой; 4) как у параметра a_1.

13.8. Укажите номер правильного ответа.

Для модели из задания *13.7*: 1) выполняются предпосылки теоремы Гаусса – Маркова; 2) не выполняются предпосылки теоремы Гаусса – Маркова; 3) выполняются условия теста Голдфелда – Квандта; 4) выполняются условия теста Дарбина – Уотсона.

13.9. Укажите номер правильного ответа.

Модель из задания *13.7*: 1) является линейной по коэффициентам; 2) не является линейной по коэффициентам; 3) образует схему Дарбина – Уотсона; 4) образует схему Голдфелда – Квандта.

13.10. Укажите номер правильного ответа.

Функция регрессии модели из задания *13.7* является: 1) степенной функцией; 2) показательной функцией; 3) кинематической функцией; 4) функцией Кобба – Дугласа.