DOI: 10.34020/2073-6495-2020-3-144-154

УДК 519.711.3

КОЛИЧЕСТВЕННЫЙ МЕТОД ОЦЕНКИ РЕЗУЛЬТАТОВ ТЕСТИРОВАНИЯ

Ганичева А.В.
Тверская государственная сельскохозяйственная академия
E-mail: TGAN55@yandex.ru

Ганичев А.В.
Тверской государственный технический университет
E-mail: alexej.ganichev@yandex.ru

В статье рассмотрены такие важные вопросы тестирования, как согласованность оценки по каждому заданию теста со средней оценкой, определение необходимого числа заданий в тесте, прогнозирование процента отклонений оценок в данном диапазоне для генеральной совокупности при заданных значениях надежности и точности. Дано определение относительной ошибки оценивания в генеральной совокупности заданий и разработан метод расчета ее числовых характеристик (среднего значения, разброса, плотности распределения, процента возможных относительных ошибок в данном промежутке).

Ключевые слова: тест, согласованность оценки, число заданий, точность, надежность, относительная ошибка.

QUANTITATIVE METHOD FOR EVALUATING TEST RESULTS

Ganicheva A.V.
Tver State Agricultural Academy
E-mail: TGAN55@yandex.ru

Ganichev A.V.
Tver State Technical University
E-mail: alexej.ganichev@yandex.ru

The article considers such important testing issues as estimate conformity to average estimate regarding each test item, determination of the necessary number of test items, prediction of percentage of estimate deviations within the given range for general totality at specified reliability and accuracy values. The definition of relative estimate error in general totality of items is given; a method of calculation of its quantitative characters (average value, scatter, distribution density and percentage of probable relative errors within the given range) was developed.

Keywords: test, consistency of evaluation, number of tasks, accuracy, reliability, relative error.

ВВЕДЕНИЕ

Проблема оценки качества результатов тестирования является одной из самых важных во многих сферах жизни общества. Особенно она актуальна в образовательном процессе в связи с использованием дистанционных технологий обучения.

Вопросам тестирования посвящены работы многих авторов [1, 2, 6, 8–12], в которых исследованы характеристики надежности, валидности и сложности тестов. В работе [4] разработана интеллектуальная информационная система оптимального контроля знаний, использующая изоморфизм нечетких графов для решения заданий практического занятия, контрольной работы, при тестировании. Работа [5] посвящена методу определения оптимальных модулей и компетентности обучающихся. В работе [13] предложена математическая модель составляющих учебного процесса. Один из основных вопросов в теории тестирования – ошибка измерения знания. Для этого вычисляется стандартная ошибка индивидуального балла обучающихся [1], которая зависит от среднего балла всех испытуемых в данной группе по данному тесту. В данной статье будут рассмотрены отклонения оценок тестирования обучающегося, зависящие от его среднего балла.

Тестирование каждого обучающегося можно рассматривать как коллективное оценивание по разным фрагментам (темам, разделам) данной дисциплины. Коллектив экспертов в данном случае образует вопросы теста.

Все возможные оценки ответов обучаемого на разные вопросы (задания) по всему курсу, не только входящие в данный тест, образуют генеральную совокупность некоторой случайной величины X, которая представлена выборкой $\{x_i \mid i=\overline{1,n}\}$, средней выборочной $\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ и оценкой дисперсии $S_x^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\overline{x})^2$.

Пусть m_x и σ_x – соответственно математическое ожидание и среднее квадратическое отклонение случайной величины X. Величина $\overline{x}$ является случайной, так как определяется выборкой на основе статистических данных. Все дальнейшие рассуждения будем проводить для случайной величины X, аппроксимируемой нормальным распределением. Пусть аппроксимация проведена на уровне значимости α_0. До опыта все элементы x_i выборки представляют собой случайные величины, которые будем считать независимыми и имеющими тот же закон распределения, что и случайная величина X, и с теми же числовыми характеристиками, усеченным нормальным распределением в промежутке $[m_x-u, m_x+u]$, где $m_x-u_x \geq 0$.

Возможны два варианта. Первый вариант: если $u \geq 3\sigma_x$, то по правилу «3 сигм» с вероятностью, не меньшей $0{,}9973 \approx 1$, при заданной степени точности можно оперировать не с усеченным, а с обычным нормальным распределением X. Второй вариант: если u будет меньше, чем 3σ. В этом случае введем величину u_1 такую, что $u<u_1=3\sigma_x$, и рассмотрим новую случайную величину:

$$X_1=m_x-u+\frac{u}{u_1}(X-m_x+u_1),$$

которая находится в промежутке $[m_x - u_1, m_x + u_1]$, т.е. для X_1 с вероятностью, не меньшей 0,9973, можно оперировать не с усеченным, а с обычным нормальным законом распределения с $M[X_1] = m_x$ и $\sigma[X_1] = \frac{u}{u_1} \cdot \sigma[X]$.

Не нарушая общности, будем рассматривать первый вариант.

Важной задачей выборочного метода является оценка параметров (характеристик) генеральной совокупности по данным выборки. В связи с этим актуальными являются задачи: 1) оценка показателя согласованности результатов тестирования по всем вопросам теста; 2) определение объема заданий при прогнозе для генеральной совокупности процента отклонений оценок в данном диапазоне; 3) определение для генеральной совокупности относительной ошибки оценивания, ее среднего значения, разброса, плотности распределения, процента возможных относительных ошибок в данном промежутке.

1. ОПРЕДЕЛЕНИЕ СОГЛАСОВАННОСТИ ОЦЕНОК ЗАДАНИЙ ТЕСТА СО СРЕДНЕЙ ОЦЕНКОЙ

При нахождении отклонений в оценке заданий по тесту важная роль отводится относительному отклонению (относительной ошибке) оценивания, которое можно определить либо как $\delta = |x - \bar{x}| / \bar{x}$, либо через $\delta_1 = (x - \bar{x}) / \bar{x}$ (случай $(x - \bar{x}) / \bar{x}$ симметричен δ_1). Пусть $\delta = |x - \bar{x}| / \bar{x}$ – относительное отклонение. Если $|x - \bar{x}| / \bar{x} > 1$, то оценка x не согласована с $\bar{x}$, если $|x - \bar{x}| / \bar{x} \le 1$, то x согласована с $\bar{x}$ и тем лучше, чем меньше относительное отклонение. Данный показатель можно рассматривать как показатель согласования в оценке.

Рассмотрим в качестве относительного отклонения $\delta_1 = |x - \bar{x}| / \bar{x}$. Это значение случайной величины Δ_1, представленной отношением величин $Y = X - \bar{x}$ и $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, которые на уровне a_0 аппроксимируются нормальным распределением, причем $M[Y] = 0$, $M[\bar{x}] = m_x$. Случайная величина Y представляет собой абсолютную ошибку тестирования.

Покажем, что Δ_1 имеет нормальное распределение, найдем соответствующие характеристики, выраженные через $\bar{x}$ и S_x^2, а также необходимое число наблюдений (вопросов в тесте), чтобы с заданными надежностью и точностью имело место нормальное распределение.

2. ОПИСАНИЕ АБСОЛЮТНОЙ ОШИБКИ ОЦЕНИВАНИЯ

Для описания случайной величины Y сначала оценим σ_x^2 через S_x^2. В работе [3] рассмотрен приближенный метод построения доверительного интервала для дисперсии σ_x^2, когда число наблюдений $n \ge 30$.

Это интервал $\left(S_x^2 - \varepsilon_1 - t_{\beta_1}\sqrt{\frac{2}{n-1}}(S_x^2 - \varepsilon_1),\, S_x^2 + \varepsilon_1 + t_{\beta_1}\sqrt{\frac{2}{n-1}}(S_x^2 + \varepsilon_1)\right)$, где

где ε_1 – точность оценки, t_{β_1} – аргумент функции Лапласа для доверительной вероятности β_1.

Отсюда находим

$$n \ge n_1 = \left\lfloor \frac{2t_{\beta_1}^2 (S_x^2 + \varepsilon_1)^2}{\varepsilon_1^2} \right\rfloor + 1. \quad (1)$$

Запись $\lfloor \ \rfloor$ обозначает целую часть числа.

Приведем алгоритм нахождения n, когда S_x не задано, и происходит итерационный процесс определения такого значения S_x, чтобы выполнялось неравенство (1).

Шаги алгоритма (Алгоритм 1).

1. Дано: $\alpha_1 = 1 - \beta_1, \varepsilon_1, n_1, n_{\max}$ – максимально возможное в данной ситуации число вопросов теста.

2. Находим правую часть неравенства (1), когда $S_x^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ и $\bar{x} = \frac{1}{n_1}\sum_{i=1}^{n} x_i$. Для этого рассматриваются оценки n_1 независимых экспертов.

3. Если полученная на шаге 2 правая часть (1) будет не меньше n_1 и не больше $n_{\max}$, то получаем оценку для числа экспертов n, удовлетворяющую неравенству: $n_1 \le n \le n_{\max}$. Переход на Конец.

4. Если правая часть (1) меньше, чем n_1, то полагаем $n_1 = n_1 + 1$. Если при этом $n_1 \le n_{\max}$, то переходим на шаг 2.

5. В противном случае надо уменьшать значения α_1 и ε_1.

Таким образом, при $n > n_1$ с вероятностью $\beta_1 = 1 - \alpha_1$, с точностью ε_1 $D_x = S_x^2$ и с точностью $\frac{\varepsilon_1}{n}$ значение $D[\bar{x}] \approx \frac{S_x^2}{n}$.

При вычислении относительной погрешности считаем, что $\varepsilon_1 = \varepsilon_2 S_x^2$ и $\varepsilon_2 < 1$. Тогда относительная погрешность

$$\frac{\left|\sigma_x^2 - S_x^2\right|}{\sigma_x^2} \le \frac{\varepsilon_1}{S_x^2 - \varepsilon_1} = \frac{\varepsilon_2 \cdot S_x^2}{S_x^2 - \varepsilon_2 S_x^2} = \frac{\varepsilon_2}{1 - \varepsilon_2}.$$

Таким образом,

$$\frac{\left|\sigma_x^2 - S_x^2\right|}{\sigma_x^2} \le \frac{\varepsilon_2}{1 - \varepsilon_2}. \quad (2)$$

Неравенство (2) выполняется при $n \ge n_1$ с вероятностью $1 - \alpha_1$.

Найдем

$$D[Y] = D[X - \bar{x}] = D[X] + D[\bar{x}] - 2K_{X,\bar{x}} = \begin{cases} \frac{n+1}{n} D_x, \text{ если } x \notin \{x_1, ..., x_n\}, \\ D_x, \text{ если } x \in \{x_1, ..., x_n\}. \end{cases}$$

Тогда

$$f(y) = f(x - \bar{x}) = \begin{cases} \frac{\sqrt{n}}{\sqrt{2\pi} \cdot \sqrt{n+1}\ \sigma_x} e^{-\frac{(x-\bar{x})^2 \cdot n}{2(n+1)\sigma_x^2}}, \text{ если } x \notin \{x_1, ..., x_n\}, \\ \frac{\sqrt{n}}{\sqrt{2\pi} \cdot \sqrt{n+2}\ \sigma_x} e^{-\frac{(x-\bar{x})^2 \cdot n}{2(n+2)\sigma_x^2}}, \text{ если } x \in \{x_1, ..., x_n\}. \end{cases} \quad (3)$$

Заметим, что $\sqrt{n} \approx \sqrt{n+2}$, $\sqrt{n} \approx \sqrt{n+1}$ с точностью 3,3 % при $n \ge 40$.

Формула (3) при $n \geq \max\{40, n_1\}$ преобразуется в формулу

$$f_1(y) = f_1(x - \overline{x}) = \frac{1}{\sqrt{2\pi} \cdot S_x} e^{-\frac{(x-\overline{x})^2}{2S_x^2}}, \tag{4}$$

поскольку при этих значениях n с вероятностью $1-\alpha_1$ и с точностью ε_1 $\sigma_x^2 \approx S_x^2$.

Замечание. Если потребовать выполнение условия: $\varepsilon_2' S_x^2 \leq \varepsilon_1 \leq \varepsilon_2 S_x^2$, где $0 < \varepsilon_2' < \varepsilon_2 < 1$, то условие (1) преобразуется к виду:

$$n \geq n_1 = \left\lfloor \frac{2t_{\beta_1}^2 (1+\varepsilon_2')^2}{\varepsilon_2^2} \right\rfloor + 1.$$

Так, для $\alpha_1 = 0{,}1; \varepsilon_2 = 0{,}1; \varepsilon_2' = 0{,}09$ находим: $n \geq 647$.

3. ОЦЕНКА МАТЕМАТИЧЕСКОГО ОЖИДАНИЯ

Неизвестное значение m_x можно заменить точечной оценкой $\overline{x}$. Оценим соответствующие абсолютную и относительные погрешности.

Из теоремы Чебышева следует, что

$$P(|m_x - \overline{x}|) \leq \varepsilon_3) \geq 1 - \frac{D[\overline{x}]}{\varepsilon_3^2} = 1 - \frac{\sigma_x^2}{n\varepsilon_3^2} \geq 1 - \alpha_2, \tag{5}$$

где α_2 – сколь угодно малое число.

При этом

$$n \geq n_2 = \left\lfloor \frac{S_x^2 + \varepsilon_1}{\varepsilon_3^2 \alpha_2} \right\rfloor + 1. \tag{6}$$

Относительная погрешность

$$\varepsilon_4 = \left| \frac{\overline{x} - m}{m_x} \right| \leq \frac{\varepsilon_3}{m_x} \leq \frac{\varepsilon_3}{\overline{x} - \varepsilon_3},$$

где $\varepsilon_3 = \overline{x} \cdot \varepsilon_4$ и $0 < \varepsilon_4 < 1$, т.е.

$$\varepsilon_4 \leq \frac{\varepsilon_3}{1 - \varepsilon_3}. \tag{7}$$

Если S_x^2 не задано, то для формулы (6) используется Алгоритм 2, аналогичный Алгоритму 1.

4. РЕАЛИЗАЦИЯ ОЦЕНКИ ЧИСЛА ВОПРОСОВ

Приведем пример по оценке числа вопросов теста согласно формуле (6). Рассмотрим сначала первый вариант определения промежутка задания X (об этом говорилось во Введении), когда $\overline{X} \in \left[\overline{x} - \varepsilon_3 + 3\sqrt{S_x^2 + \varepsilon_1}, \overline{x} + \varepsilon_3 + 3\sqrt{S_x^2 + \varepsilon_1} \right]$.

Если, например, средний балл $\overline{x}$ равен 5, то при $\varepsilon_4 = 0{,}1$ имеем: $\varepsilon_3 \geq 5 \cdot 0{,}1 = 0{,}5$. При максимальном балле 10 величина $S_x = \frac{5}{3}$, тогда формула (6) преобразуется к виду

$$n_2 = \left\lfloor \frac{2{,}78 + \varepsilon_1}{0{,}25 \cdot \alpha_2} \right\rfloor + 1.$$

Поскольку $\varepsilon_1 = S_x^2 \cdot \varepsilon_2$, то при $\varepsilon_2 = 0{,}1$ имеем $\varepsilon_1 = 2{,}78 \cdot 0{,}1 = 0{,}278$. При $\alpha_2 = 0{,}1$ получаем $n_2 = 124$.

Во втором варианте используем X_1 и согласно формуле (6) при $\varepsilon_3 = 0{,}5$, $\alpha_2 = 0{,}1$, $\varepsilon_1 = S_{x_1}^2 \cdot \varepsilon_2$, $\varepsilon_2 = 0{,}1$ имеем: $n_2 \geq 44 \cdot S_{x_1}^2$, т.е. n_2 определяется величиной разброса случайной величины X при данных $\varepsilon_1, \varepsilon_2, \alpha_2$.

Так, при тестировании в Тверской государственной сельскохозяйственной академии по теории вероятностей и математической статистике максимальный балл был 10 с учетом сложности. Использовались подсказки (не более 5), каждая из которых снижает оценку на $0{,}7 \cdot k$ баллов, где k – число полезных подсказок. Максимальное количество подсказок, равное 5, определено на основе пробных тестирований, при которых допускалось консультирование. Снижение балла происходит в арифметической прогрессии. Параметр 0,7 выбран из условия максимального «штрафа» за подсказку, так как в этом случае при пяти подсказках максимально снижается балл. Существуют разные варианты определения числа подсказок и баллов за них.

Отметим, что результаты данной статьи справедливы и при традиционном тестировании без подсказок. В то же время умело определенные подсказки во многом способствуют использованию теста не только для контроля, но и для обучения, т.е. повышают обучающий потенциал тестовых заданий.

В случае неправильного ответа по данному заданию возможен двоякий подход. Первый подход (стандартный) – за данное задание ставится оценка 0 баллов, независимо от того, сколько подсказок студент использовал. Второй подход – использование оценок онтологий и различных фрагментов данного задания и методов нечеткого управления в адекватной системе оценивания качества обучения.

При выборке объема $n = 150$ заданий теста у группы студентов в среднем получилось: $\bar{x} = 5{,}1$; $S_{x'}^2 = 3{,}06$ и $n \geq 3{,}06 \cdot 44 = 134{,}6$.

При аппроксимации статистических данных нормальным законом распределения уровень значимости составил 0,1.

5. ПОСТРОЕНИЕ ФУНКЦИИ ПЛОТНОСТИ ОТНОСИТЕЛЬНОЙ ОШИБКИ ОЦЕНИВАНИЯ И ЕЕ ИСПОЛЬЗОВАНИЕ

Произведем разложение функции $Y_1 = \vartheta(\bar{x}) = \dfrac{1}{\bar{x}}$ в ряд Тейлора в окрестности точки $m_{\bar{x}}$, ограничиваясь тремя слагаемыми:

$$Y_1 = \frac{1}{m_{\bar{x}}} - \frac{1}{(m_{\bar{x}})^2} \cdot (\bar{x} - m_{\bar{x}}) + \frac{1}{(m_{\bar{x}})^3} \cdot (\bar{x} - m_{\bar{x}})^2. \tag{8}$$

Погрешность такого представления равна остаточному члену $R(\bar{x})$ ряда Тейлора, причем $R(\bar{x}) = \dfrac{1}{2\xi^4}(\bar{x} - m_{\bar{x}})^3$, ξ заключено между $\bar{x}$ и $m_{\bar{x}}$. При этом, как следует из (5), $\bar{x}$ отличается от $m_{\bar{x}}$ не более чем на ε_3, с вероятностью, не меньшей 1 – α_2, где ε_3 и α_2 – сколь угодно малые числа. Поэтому $R(\bar{x})$ будет сравнительно небольшим. Пусть $e = m_{\bar{x}}$, если $\xi \geq m_{\bar{x}}$, $e = m_{\bar{x}} - \xi$, если $m_{\bar{x}} > \xi$.

Из формулы (8) с точностью

$$\varepsilon_5 = M\left[\left|R_2(\bar{x})\right|\right] \leq M\left[\frac{1}{2\xi^4}\cdot\varepsilon_3^3\right] = \frac{1}{2}\varepsilon_3^3 M\left[\frac{1}{\xi^4}\right] < \frac{1}{2}\varepsilon_3^3/(e)^4 \leq \frac{1}{2}\varepsilon_3^3/(\bar{x}-\varepsilon_3)^4$$

найдем

$$m_{y_1} = \frac{1}{m_{\bar{x}}} + \frac{1}{(m_{\bar{x}})^3}\cdot D_{\bar{x}}, \tag{9}$$

причем, например, при $\bar{x} = \overline{1,5}$, $\varepsilon_3 = 0,01$ получим $\varepsilon_5 = 0$ с точностью до шестого знака после запятой. При $0 \leq \bar{x} = \overline{1,5} < 0$ и $\varepsilon_3 = 0,01$ аналогично $\varepsilon_5 = 0$ с точностью до пятого знака. Вырожденный случай, когда $\bar{x} = 0$, не рассматривается.

По определению, $\Delta_1 = Y \cdot Y_1$, или $\delta_1 = y \cdot y_1$. Разлагая функцию δ_1 в ряд Тейлора в окрестности точки (m_y, m_{y_1}), получим

$$\delta_1 = m_y \cdot m_{y_1} + m_{y_1}\cdot(y - m_y) + m_y \cdot (y_3 - m_{y_1}).$$

Данное представление δ_1 является точным, так как все частные производные высших порядков равны нулю. Отсюда, поскольку $m_y = 0$, с учетом (7) находим

$$\delta_1 = m_{y_1}\cdot y = \left(\frac{1}{m_{\bar{x}}} + \frac{1}{(m_{\bar{x}})^3}\cdot D_{\bar{x}}\right)y. \tag{10}$$

Следовательно,

$$m_{\delta_1} = \left(\frac{1}{m_{\bar{x}}} + \frac{1}{(m_{\bar{x}})^3}\cdot D_{\bar{x}}\right)\cdot m_y = 0, \tag{11}$$

$$D_{\delta_1} = \left(\frac{1}{m_{\bar{x}}} + \frac{1}{(m_{\bar{x}})^3}\cdot D_{\bar{x}}\right)^2 \cdot \frac{n+1}{n}\sigma_x^2 \approx \left(\frac{1}{m_{\bar{x}}} + \frac{1}{(m_{\bar{x}})^3}\cdot D_{\bar{x}}\right)^2 \cdot \frac{n+1}{n}S_x^2. \tag{12}$$

Равенство (12) выполняется с относительной точностью $\frac{\varepsilon_2}{1-\varepsilon_2}$ при $n \geq n_1$ с вероятностью $(1-\alpha_1)$.

Из (10)–(12) с учетом того, что Y имеет нормальное распределение (на уровне значимости α_0), находим

$$f_1(\delta_1) = \frac{1}{\sqrt{2\pi}\cdot\sigma_x\cdot c}e^{-\frac{(\delta_1)^2}{2\sigma_x^2 c^2}}, \tag{13}$$

где $c = \left(\frac{1}{m_{\bar{x}}} + \frac{1}{(m_{\bar{x}})^3}\cdot\frac{\sigma_x^2}{n}\right)\cdot\sqrt{\frac{n+1}{n}}$. Точность данной формулы при замене σ_x^2 на S_x^2 определяется так:

$$\varepsilon_6 = \left|-\frac{1}{2} - \gamma_1 + \gamma_2 + \gamma_3\right|\cdot\frac{\varepsilon_2}{1-\varepsilon_2}, \tag{14}$$

где $\gamma_1=\dfrac{\sigma_x^2}{m_x^2 n+\sigma_x^2}$; $\gamma_2=\dfrac{\delta_1^2}{2\sigma_x^2\cdot\left(\dfrac{1}{m_x}+\dfrac{1}{m_x^3}\cdot\dfrac{\sigma_x^2}{n}\right)^2}$; $\gamma_3=\dfrac{\delta_1^2}{\left(m_x^2+\dfrac{\sigma_x^2}{n}\right)^3 n}$.

Возможны два варианта преобразования выражения (14).

Первый вариант предполагает, что сумма положительных членов в (14) не меньше суммы отрицательных. Тогда

$$\varepsilon_6=\left|-\frac{1}{2}-\gamma_1+\gamma_2+\gamma_3\right|\cdot\frac{\varepsilon_2}{1-\varepsilon_2}$$

и

$$\varepsilon_6\le\left(-\frac{1}{2}-\frac{S_x^2(1-\varepsilon_2)}{\overline{x}^2(1+\varepsilon_3)^2\cdot n+S_x^2(1+\varepsilon_3)}+\delta_1^2/2\cdot(1+\varepsilon_2)\left(\frac{1}{\overline{x}(1+\varepsilon_3)}+\frac{(1-\varepsilon_2)\cdot S_x^2}{\overline{x}^3(1+\varepsilon_3)^3\cdot n}\right)^2+\right.$$
$$\left.+\delta_1^2/\left(\overline{x}^2(1+\varepsilon_3)^2+\frac{S_x^2(1-\varepsilon_2)}{n}\right)^3\cdot n\right)\cdot\frac{\varepsilon_2}{1-\varepsilon_2}.$$

Например, пусть $S_x=1{,}67,\ \overline{x}=5,\ \varepsilon_2=0{,}1,\ \varepsilon_3=0{,}01\ \ n\ge n_3=\max\{n_1,n_2\}\ge 647.$ Тогда

$$\varepsilon_6\le(-0{,}5+11{,}6\delta_1^2)\cdot\frac{\varepsilon_2}{1-\varepsilon_2}=-0{,}05+1{,}16\delta_1^2.$$

Поскольку $\delta_1=\dfrac{|x-\overline{x}|}{\overline{x}}=\left|\dfrac{x}{\overline{x}}-1\right|$, то, чем ближе x к $\overline{x}$, тем ближе ε_6 к нулю, но при этом должно быть либо $x\ge 1{,}2\overline{x}$, либо $x\le 0{,}8\overline{x}$, но это противоречивое неравенство.

При втором варианте сумма положительных членов в (13) меньше суммы отрицательных и $\varepsilon_6=\left(\dfrac{1}{2}+\gamma_1-\gamma_2-\gamma_3\right)\cdot\dfrac{\varepsilon_2}{1-\varepsilon_2}$. Оценка ε_6 в этом случае отличается от оценки первого варианта только знаками в скобках, содержащих ε_2 и ε_3: знаки перед ε_2 и ε_5 меняются на противоположные.

Для этого варианта для тех же данных, что и в первом варианте, получаем

$$\varepsilon_6\le\left(0{,}5-11{,}6\delta_1^2\right)\cdot\frac{\varepsilon_2}{1-\varepsilon_2}=0{,}05-1{,}16\delta_1^2.$$

При этом, чем меньше x отличается от $\overline{x}$, тем ближе ε_6 к нулю при $0{,}8\overline{x}\le x\le 1{,}2\overline{x}$.

Так, если $\overline{x}=5,\ x=4{,}5$, то $\delta_1=0{,}1$ и $\varepsilon_6\le 0{,}4\ \%$.

При замене в (12) σ_x на S_x получаем плотность

$$f_1(\delta_1)=\frac{1}{\sqrt{2\pi}S_x c_1}\cdot e^{\frac{-\delta_1^2}{2S_x^2c_1}}, \tag{15}$$

где $c_1=\left(\dfrac{1}{m_{\overline{x}}}+\dfrac{1}{(m_{\overline{x}})^3}\right)\sqrt{\dfrac{n+1}{n}}$.

Уточним в формуле (15) c_1, заменив $m_{\bar{x}}$ на $\bar{x}$. Относительная погрешность при этом будет

$$\varepsilon_7 \le \gamma \cdot \left|1-\delta_1^2 \cdot \sqrt{\frac{n}{n+1}} \Big/ 2(S_x^2-\varepsilon_1)\left(1\Big/\bar{x}(1+\varepsilon_3)(1+\varepsilon_3) + (S_x^2-\varepsilon_1)\Big/\bar{x}^3(1+\varepsilon_3)^3 \cdot n\right)\right| \times$$

$$\times \frac{\varepsilon_4^4}{1-\varepsilon_4}, \qquad (16)$$

где $\gamma = \left(\frac{1}{1-\varepsilon_3}+\frac{3}{\bar{x}^3(1-\varepsilon_3)^3}\cdot\frac{S_x^2+\varepsilon_1}{n}\right)\Big/\left(\frac{1}{\bar{x}(1+\varepsilon_3)}+\frac{1}{\bar{x}^3(1+\varepsilon_3)^3}\cdot\frac{S_x^2-\varepsilon_1}{n}\right)$ и в (16) под знаком модуля вычитаемое меньше, чем 1. Если вычитаемое больше 1, то знаки слагаемых меняются и в знаменателе второго члена будет стоять вместо $1+\varepsilon_3$ выражение $1-\varepsilon_3$.

При $S_x = 1{,}67$; $\bar{x} = 5$; $n \ge 647$; $\varepsilon_3 = 0{,}01$; $\varepsilon_1 = 0{,}28$; $\delta_1 = 0{,}1$ имеем:

$$\varepsilon_7 \le 4{,}6 \cdot 0{,}01 = 0{,}046 \text{ или } 4{,}6~\%.$$

Относительная погрешность при замене в (14) σ_x^2 на S_x^2 и m_x на $\bar{x}$

$$\frac{\Delta f}{f} \approx \frac{f'_{\sigma_x^2}\cdot\sigma_x^2}{f}\cdot\frac{\Delta\sigma_x^2}{\sigma_x^2}+\frac{f'_{m_x}\cdot m_x}{f}\cdot\frac{\Delta m_x}{m_x}. \qquad (17)$$

Относительная погрешность составляет $\varepsilon_6+\varepsilon_7$. Для рассматриваемого примера это 5 %.

Погрешность самой формулы не более $0{,}1(\varepsilon_6+\varepsilon_7)$, поскольку при замене приращение функции ее дифференциалом погрешность была определена как $0{,}1 \cdot df$. Для рассматриваемого примера эта погрешность составляет 0,5 %.

Таким образом, формула (14) преобразовалась в формулу (18):

$$f_1(\delta_1) = \frac{1}{\sqrt{2\pi}S_x c_2}\cdot e^{\frac{-\delta_1^2}{2S_x^2c_2}}, \qquad (18)$$

где $c_2 = \left(\frac{1}{\bar{x}}+\frac{1}{\bar{x}^3}\cdot\frac{S_x^2}{n}\right)\sqrt{\frac{n}{n+1}}$.

Относительная погрешность такого преобразования не более $\varepsilon_8 = 1{,}1(\varepsilon_6+\varepsilon_7)$ при вероятности $(1-\alpha_0)(1-\alpha_1)(1-\alpha_2)$ и $n \ge n_3$.

Для рассмотренного примера $\varepsilon_8 = 1{,}1(0{,}4+4{,}6) = 5{,}5~\%$, т.е. погрешность не более 5,5 % при вероятности, не меньшей 0,73, при $n \ge 647$.

Из (18) находим

$$P(-d<\delta_1<d) = \Phi\left(\frac{d}{S_x\cdot c_2}\right) = \gamma. \qquad (19)$$

Погрешность этой формулы не превосходит суммы погрешности функции Лапласа и величины $\varepsilon_8\gamma$.

Итак, формулы (10)–(19) при соответствующем объеме выборки дают достаточно точное значение соответственно относительного отклонения в оценке, его среднего значения, разброса, плотности распределения и вероятности попадания на участок $(-d, d)$.

ЗАКЛЮЧЕНИЕ

В работе получены следующие результаты: приведена оценка согласованности результатов тестирования, определен объем заданий, при котором процент возможных отклонений оценки будет в данном интервале, для генеральной совокупности определено относительное отклонение оценивания и его характеристики.

Полученные результаты используются в учебном процессе в учебных заведениях г. Твери.

Разработанный метод анализа качества процесса тестирования может использоваться не только в образовании, но и в психологическом консультировании, системах оценивания профессиональной подготовки, в медицинской диагностике, социологических опросах, выборочном контроле оценки качества продукции, системах тестирования аппаратуры, программного обеспечения и баз данных.

Литература

1. *Аванесов В.С.* Научные основы тестового контроля знаний. М.: Исследовательский центр, 1994. 135 с.
2. *Бочко С.Б., Изимов М.У.* Математическая модель оценки результатов тестирования // Вестник Томского государственного педагогического университета. 2004. № 6. С. 88–89.
3. *Вентцель Е.С.* Теория вероятностей: учебник для вузов. М.: Высшая школа, 1999. 576 с.
4. *Ганичева А.В.* Интеллектуальная информационная система оптимального контроля знаний // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 101. С. 358–374.
5. *Ганичева А.В.* Метод определения оптимальных модулей и компетентности обучаемых // Качество. Инновации. Образование. 2013. № 10 (101). С. 19–23.
6. *Ганичева А.В.* Оценка эффективности процесса обучения // Бизнес. Образование. Право. Вестник Волгоградского института бизнеса. 2014. № 4. С. 301–304.
7. *Ганичева А.В.* Модель менеджмента качества учебных планов // Качество. Инновации. Образование. 2012. № 4 (83). С. 37–41.
8. *Ганичева А.В. и др.* Тестовые технологии в обучении / отв. ред. А.В. Ганичева. Тверь: ТГСХА, 2011. 130 с.
9. *Звонников В.И., Челышкова М.Б.* Современные средства оценивания результатов тестирования: учебное пособие для студ. высш. учеб. заведений. М.: Издательский центр «Академия», 2007. 224 с.
10. *Кузьмин О.В., Бочко С.Б.* Тестирование, как способ мониторинга качества подготовки технических специалистов // Современные технологии. Системный анализ. Моделирование. 2010. № 1. С. 247–260.
11. *Полещук О.М., Рыбников К.К.* Модели анализа тестирования в образовательном процессе // Ярославский педагогический вестник. 2002. № 2 (31). С. 124–127.
12. *Челышкова М.Б.* Теория и практика конструирования педагогических тестов: учебное пособие. М.: Логос, 2002. 432 с.
13. *Ganicheva A.V.* Optimization Models of Components of Educational Process. British Journal of Mathematics and Computer Science. 2016. № 14 (5). P. 7–11.
14. *Safargaliev E.R., Eremina I.I., Savitsky S.K., Camelina V.A.* Mathematical Model and Qualimetric Assessment of Graduate Education Quality in Environment Saturated with Information and Communication Technologies. International Education Studies. 2015. Vol. 8. № 2. P. 78–83.

Bibliography

1. *Avanesov V.S*. Nauchnye osnovy testovogo kontrolja znanij. M.: Issledovatel'skij centr, 1994. 135 p.
2. *Bochko S.B., Izimov M.U*. Matematicheskaja model' ocenki rezul'tatov testirovanija // Vestnik Tomskogo gosudarstvennogo pedagogicheskogo universiteta. 2004. № 6. P. 88–89.
3. *Ventcel' E.S*. Teorija verojatnostej: uchebnik dlja vuzov. M.: Vysshaja shkola, 1999. 576 p.
4. *Ganicheva A.V*. Intellektual'naja informacionnaja sistema optimal'nogo kontrolja znanij // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta. 2014. № 101. P. 358–374.
5. *Ganicheva A.V*. Metod opredelenija optimal'nyh modulej i kompetentnosti obuchaemyh // Kachestvo. Innovacii. Obrazovanie. 2013. № 10 (101). P. 19–23.
6. *Ganicheva A.V*. Ocenka jeffektivnosti processa obuchenija // Biznes. Obrazovanie. Pravo. Vestnik Volgogradskogo instituta biznesa. 2014. № 4. P. 301–304.
7. *Ganicheva A.V*. Model' menedzhmenta kachestva uchebnyh planov // Kachestvo. Innovacii. Obrazovanie. 2012. № 4 (83). P. 37–41.
8. *Ganicheva A.V. i dr*. Testovye tehnologii v obuchenii / otv. red. A.V. Ganicheva. Tver': TGSHA, 2011. 130 p.
9. *Zvonnikov V.I., Chelyshkova M.B*. Sovremennye sredstva ocenivanija rezul'tatov testirovanija: uchebnoe posobie dlja stud. vyssh. ucheb. zavedenij. M.: Izdatel'skij centr «Akademija», 2007. 224 p.
10. *Kuz'min O.V., Bochko S.B*. Testirovanie, kak sposob monitoringa kachestva podgotovki tehnicheskih specialistov // Sovremennye tehnologii. Sistemnyj analiz. Modelirovanie. 2010. № 1. P. 247–260.
11. *Poleshhuk O.M., Rybnikov K.K*. Modeli analiza testirovanija v obrazovatel'nom processe // Jaroslavskij pedagogicheskij vestnik. 2002. № 2 (31). P. 124–127.
12. *Chelyshkova M.B*. Teorija i praktika konstruirovanija pedagogicheskih testov: uchebnoe posobie. M.: Logos, 2002. 432 p.
13. *Ganicheva A.V*. Optimization Models of Components of Educational Process. British Journal of Mathematics and Computer Science. 2016. № 14 (5). P. 7–11.
14. *Safargaliev E.R., Eremina I.I., Savitsky S.K., Camelina V.A*. Mathematical Model and Qualimetric Assessment of Graduate Education Quality in Environment Saturated with Information and Communication Technologies. International Education Studies. 2015. Vol. 8. № 2. P. 78–83.