

9. Chepurnaja, Ju. V. "Situacionnoe upravlenie v adaptivnoj kompjuiterizirovannoj sisteme testirovanija obuchaemyh." *Jekspertnye ocenki jelementov uchebnogo processa: materialy XI mezhvuz. nauch.-prakt. konf.* Harkov: Narodnaja ukrainskaja akademija, 2009: 65–66.

10. Shirokih, G. V., and Chernyj, S. G. "Modelirovanie processov zagruzhennosti postuplenija informacii na osnove samoobuchajushhihsja setej." *Informacionnye tehnologii modelirovanija i upravlenija* 2(20) (2005): 279–282.

11. Chernyi, S. "The implementation of technology of multi-user client-server applications for systems of decision making support." *Metallurgical and Mining Industry* 3 (2015): 60–65.

12. Zhilenkov, A., and S. Chernyi. "Investigation performance of marine equipment with specialized information technology." *Procedia Engineering* 100 (2015): 1247–1252. DOI:10.1016/j.proeng.2015.01.490

ИНФОРМАЦИЯ ОБ АВТОРАХ

Журавлев Антон Евгеньевич —
кандидат технических наук.
ФГБОУ ВО «ГУМРФ имени
адмирала С. О. Макарова»
zhuravlevae@gumrf.ru
Васильченко Дмитрий Сергеевич —
кандидат технических наук.
ФГБОУ ВО «ГУМРФ имени
адмирала С. О. Макарова»
tillaus@mail.ru

INFORMATION ABOUT THE AUTHORS

Zhuravlev Anton Evgenevich — PhD.
Admiral Makarov State University
of Maritime and Inland Shipping
zhuravlevae@gumrf.ru
Vasilchenko Dmitriy Sergeevich — PhD.
Admiral Makarov State University
of Maritime and Inland Shipping
tillaus@mail.ru

Статья поступила в редакцию 20 октября 2016 г.

DOI: 10.21821/2309-5180-2016-8-6-251-259

УДК 371.3:681.51

Д. А. Печников

РАЗРАБОТКА МОДЕЛЕЙ И ПРОЦЕДУР ИНТЕРВАЛЬНОЙ ОЦЕНКИ РЕЗУЛЬТАТОВ КОНТРОЛЯ ЗНАНИЙ В КОМПЬЮТЕРНЫХ СИСТЕМАХ ТЕСТИРОВАНИЯ ВМФ

В данной статье рассматривается разработка моделей и процедур интервальной оценки результатов контроля знаний в компьютерных системах с применением методов, основанных на законах биномиального распределения и геометрического распределения как частного случая отрицательного биномиального распределения. Предложен новый переход от точечных оценок эмпирической частоты успешного выполнения тестовых заданий к интервальным оценкам вероятности этой успешности, позволяющим определять погрешность и доверительный интервал полученных оценок. Также разработаны модели критериально-ориентированного оценивания в компьютерных системах тестирования, доказано, что данные модели повышают результаты текущего педагогического контроля. Для того чтобы осуществлять военно-профессиональную подготовку специалистов ВМФ, необходима соответствующая учебно-материальная база. Основу такой базы раньше всегда составляли действующие учебные и тренировочные образцы военной техники и вооружения, которые ранее поступали в военно-морские учебные заведения в достаточном количестве и обеспечивали необходимый уровень как теоретической, так и практической подготовки будущих офицеров ВМФ. В результате резкого сокращения в 1990 гг. отраслей военно-промышленного комплекса, разрабатывающих технику и вооружения для ВМФ, новейшие образцы этой техники и вооружений теперь изготавливаются не сериями, а отдельными экземплярами. Вопрос о производстве учебных и тренировочных образцов не стоит вообще. В этих условиях в качестве единственного средства, способного обеспечить военно-профессиональную подготовку, могут рассматриваться только виртуальные аналоги образцов военной техники и вооружения, разрабатываемые средствами информационных технологий, т. е. тренажерно-обучающие системы ВМФ. Разработка и применение предлагаемых в данной статье моделей интервального оценивания результатов критериально-ориентированного те-

стирования обеспечивает переход от точечных оценок эмпирической частоты к оценкам вероятности правильного действия, которые имеют количественно оцениваемый доверительный интервал. Последнее обеспечивает применение этих моделей для решения задачи автоматизации управления процессом критериально-ориентированного тестирования в целях обеспечения его оперативности путем остановки процедуры предъявления технического задания после получения оценки заданной достоверности.

Ключевые слова: доверительный интервал, интервальная оценка, компьютерная система тестирования, критериально-ориентированное тестирование, биномиальный закон распределения.

Вводная часть

В тенденциях развития тренажерно-обучающих систем (ТОС) ВМФ отмечается следующий ряд основных противоречий: 1) между ростом возможностей ТОС как средств моделирования управляемых объектов и неизменностью их дидактических характеристик; 2) между необходимостью индивидуализации процесса подготовки и ограниченными психофизиологическими возможностями обучающихся по управлению обучением в группах с количественным составом более пяти–семи обучаемых; 3) между ростом числа и сложности задач управления обучением и неизменностью психофизиологических возможностей руководителей обучения.

В качестве средства разрешения этих противоречий рассматривается автоматизация в ТОС ВМФ функций управления обучением. Ключевой функцией, без автоматизации которой нельзя автоматизировать все остальные функции управления обучением, является функция педагогического контроля качества подготовки обучаемых. Качество реализации в ТОС ВМФ функции контроля качества подготовки различно в отношении знаний и действий (умений и навыков) обучаемых. На достижение целей контроля усвоения знаний в ТОС ВМФ ориентирована компьютерная система тестирования военного назначения (КСТ ВН) «Система автоматизированного контроля (САК)», которая входит в состав комплекса программ инструментальных средств (КПИС) «Медиа-тор». Эта КСТ является типичным представителем современных КСТ и полностью соответствует современному уровню их развития.

ГОСТ РВ 51540-2005 и ГОСТ РВ 15.205-2004 определяют КСТ по функциональному назначению как продукцию двойного назначения и относят КСТ к комплектующим изделиям межотраслевого применения (КИМП), представляющим собой «изделие военной техники (ВТ), предназначенное для выполнения определенных технических функций в составе изделий ВТ или их составных частей, создаваемое не для конкретного изделия ВТ по самостоятельным комплектам и не подвергаемое изменениям в процессе создания изделий ВТ, в котором его применяют». В соответствии с классификациями, приведенными в ГОСТ РВ 51540-2005 и ГОСТ Р 50-605-80-93, КСТ должны быть отнесены к изделиям ВТ, которые входят в категорию учебно-тренировочных средств, принадлежащих к учебной ВТ как виду ВТ.

В настоящее время в компьютерных системах тестирования наибольшее распространение получили тестовые задания закрытого типа с единственным верным ответом, которые применяются не менее чем в 92 % случаев. В рассматриваемых тестовых заданиях успешность их выполнения оценивается путем установления факта совпадения ответа обучаемого с заранее заданным правильным вариантом ответа, любой другой вариант является неправильным. Для последующей статистической обработки результаты (верно – неверно) выполнения таких тестовых заданий представляются в виде дихотомической переменной (0 – 1).

Существует большое число компьютерных систем тестирования. По оценкам авторов, только в свободном доступе в Интернете их представлено не менее 50. Абсолютное большинство КСТ (VeralTest, iSpring QuizMaker, MyTest, AVELife TestGold Studio, UniTest System, Универсальный тестовый комплекс и т. д.) ориентировано на решение задач текущего педагогического контроля, который должен осуществляться в процессе обучения и обеспечивать адаптивность управления этим процессом. Текущий контроль, как показано в [2], [4], [7] – [9] и др., может быть реализован только методами критериально-ориентированного тестирования, идентифицирующими соотно-

шения «усвоил – не усвоил» между конкретными обучаемыми и конкретными познавательными объектами, в отношении которых реализуется рассматриваемый процесс обучения. При этом «сам тест должен полностью соответствовать содержанию обучения и разрабатываться тем преподавателем, который осуществлял обучение» [4].

В таких тестах преобладает форма закрытого технического задания (ТЗ) с единственным правильным ответом. Так, из 3581 ТЗ, ориентированных на решение задач текущего педагогического контроля и представленных в Интернете, эта форма используется в 3304 случаях, т. е. в 92,3 %. По данным других авторов [1], [4], [6], [10], доля заданий с единственным правильным ответом колеблется от 86 % до 98 %. Эти оценки и определяют целесообразность первоочередного решения задач интервального оценивания в отношении ТЗ закрытого типа с единственным верным ответом. Безошибочность (правильность) ответов обучаемого определяется в виде показателя частности B , который имеет вид

$$B = \frac{\sum_{i=1}^n j_i}{n}, \quad (1)$$

где i ($i = \overline{1, n}$) — номер ТЗ; n — число ТЗ в тесте; $(j_i = \overline{0, 1})$ — результат выполнения ТЗ ($j_i = 1$ — верный ответ; $j_i = 0$ — неверный ответ).

Все существующие процедуры обработки результатов выполнения тестов в КСТ основаны на оценке вида (1). Такие оценки являются точечными и не несут информации о величине своей погрешности, а потому не могут быть использованы в целях обеспечения оперативности процедуры тестирования (прекращения тестирования при получении оценки, имеющей заданную погрешность). Далее предлагаются модели и процедуры интервального оценивания результатов выполнения систем ТЗ закрытого типа с единственным верным ответом, которые, в отличие от точечных оценок, несут в себе информацию как о своей наиболее вероятной величине, так и о ее погрешности.

Постановка и решение задачи интервальной оценки результатов критериально-ориентированного тестирования

Предлагаемые модели и процедуры базируются на том факте, что показатель «правильность» является дихотомической переменной, а процедура его определения в процессе решения обучаемым тестирующей выборки из n заданий соответствует схеме Бернулли.

Для корректного применения схемы Бернулли должны быть выполнены следующие условия:

- 1) каждое испытание имеет ровно два исхода, условно называемых успехом и неудачей;
- 2) независимость испытаний: результат очередного эксперимента не должен зависеть от результатов предыдущих экспериментов;
- 3) вероятность успеха должна быть постоянной (фиксированной) для всех испытаний.

Первое условие всегда выполняется за счет того, что в ТЗ закрытого типа с единственным верным ответом результаты представлены дихотомической переменной ($j_i = \overline{0, 1}$).

Второе условие в критериально-ориентированном тестировании выполняется также всегда за счет возможности предъявления всех ТЗ в случайном порядке.

Третье условие выполняется только в случае, когда обучаемый никаких новых сведений об объектах тестирования из формулировки ТЗ не получает. Выполнение этого требования ограничено существованием у обучаемого возможности угадывания верного решения. Возможность угадывания правильного решения имеет место во всех формах ТЗ закрытого типа, но в форме задания с единственным верным ответом вероятность такого угадывания наиболее высока.

Для исключения этой возможности и достижения истинности результатов тестирования в тестологии используются два метода. Суть первого метода состоит в снижении вероятности угадывания путем увеличения в ТЗ числа дистракторов (альтернатив решений). В тестологии [1], [5], [6], [10] и др. считается, что истинность оценки выполнения систем закрытых ТЗ с единственным верным ответом достигается, если во всех ТЗ число дистракторов равно или больше

пяти. Суть второго метода состоит во введении поправки на угадывание в результат (1) тестирования. В случае критериально-ориентированного тестирования используются фиксированные поправки, виды которых приведены в [1], [5], [6], [10] и др. Оба метода реализуемы в любой из КСТ.

С учетом изложенного примем допущение о том, что вероятность правильного выполнения ТЗ определяется только знаниями обучаемого и не зависит от фактора угадывания. В соответствии с этим допущением безошибочность выполнения ТЗ обучаемым характеризуется некоторой постоянной вероятностью $p = \text{const}$ правильного выполнения ТЗ, что и определяет полное соответствие процедуры критериально-ориентированного тестирования схеме Бернулли.

Пусть в результате решения обучаемым выборки из n ТЗ получен вектор (примеры векторов приведены в табл. 1)

$$\vec{j} = (j_1, \dots, j_i, \dots, j_n), \quad (2)$$

в котором $j_i = \overline{0,1}$ есть оценки успешности выполнения соответствующего задания ($j_i = 1$ — ТЗ выполнено без ошибок, $j_i = 0$ — ТЗ выполнено с ошибкой). Требуется оценить вероятность p безошибочного решения ТЗ.

Таблица 1

Некоторые варианты векторов результатов тестирования для $n = 12$

Вариант	j_1	j_2	j_3	j_4	j_5	j_6	j_7	j_8	j_9	j_{10}	j_{11}	j_{12}
1	1	1	1	1	1	1	1	1	1	1	1	1
2	0	0	0	0	0	0	1	1	1	1	1	1
3	1	1	1	1	1	1	0	0	0	0	0	0
4	1	0	1	0	1	0	1	0	1	0	1	0
5	0	1	0	1	0	1	0	1	0	1	0	1
6	1	1	0	0	0	0	0	0	0	1	1	1
7	1	1	0	0	1	1	1	1	1	0	1	1
8	1	1	1	1	0	1	1	1	0	1	1	1
9	1	1	1	1	0	0	0	1	1	1	1	0
10	1	1	0	1	1	0	0	1	1	0	1	0

Эмпирические данные, полученные по схеме Бернулли, подчиняются законам таких распределений как биномиальное, геометрическое и отрицательное биномиальное. В целях интервальной оценки вероятности p безошибочного выполнения ТЗ целесообразно использовать те законы распределения, которые обеспечивают обработку любого числа ($i, i = \overline{1, +\infty}$) ТЗ. Такому требованию соответствуют только законы биномиального и геометрического распределения. Биномиальное распределение имеет место в тех случаях, когда последовательность испытаний Бернулли обрывается после проведения фиксированного числа n испытаний. При этом под биномиальной случайной величиной X понимается число успехов в серии из n испытаний Бернулли.

Вероятность $P(J)$ появления вектора J при биномиальном распределении оценивается как

$$P(J) = C_n^k p^{n-k} q^k \quad \text{при} \quad C_n^k = \frac{n!}{k!(n-k)!}, \quad (3)$$

где J — рассматриваемый вектор результатов тестирования; p — вероятность безошибочного выполнения ТЗ; $q = 1 - p$ — вероятность выполнения задания с ошибкой; k — число заданий, выполненных с ошибкой; n — общее число заданий; $n - k$ — число заданий, выполненных без ошибок.

Геометрическое распределение возникает при обрыве испытаний сразу же после первого успеха или неудачи. При этом вероятность $P(J; j_i = 1)$ успеха оценивается как

$$P(J; j_i = 1) = p^h q, \quad (4)$$

а вероятность $P(J; j_i = 0)$ неудачи — как

$$P(J; j_i = 0) = q^h p, \quad (5)$$

где p — вероятность безошибочного выполнения задания; $q = 1 - p$ — вероятность выполнения задания с ошибкой; $h(h = \overline{1, n})$ — номер i последнего ТЗ, выполненного без ошибки или, наоборот, с ошибкой (число «неудач» до первого «успеха» или число «успехов» до первой «неудачи»).

Для обеспечения возможности использования геометрического распределения в отношении векторов (2) любой конфигурации введем понятие вектора $\vec{h}$ непрерывного результата (успеха или неудачи), под которым будем понимать любую часть вектора (2), каждый из элементов которой равен 1 или 0. Если все элементы вектора равны 1 ($j_i = 1$) — вектор $\vec{h}^+$ непрерывного успеха, если все элементы вектора равны 0 ($j_i = 0$) — вектор $\vec{h}^-$ непрерывной неудачи. Например, в табл. 1 для шестого варианта вектора $\vec{j}$ — это вектора $\vec{h}_1^+ = (j_1, j_2)$, $\vec{h}_2^- = (j_3, j_4, j_5, j_6, j_7, j_8, j_9)$, $\vec{h}_3^+ = (j_{10}, j_{11}, j_{12})$.

Вероятность появления вектора (1) можно оценивать с двух позиций [11] – [13]:

1) как вероятность события L , заключающегося в том, что при выполнении n заданий обучаемым было достигнуто $(n - k)$ успехов (допущено k ошибок);

2) как вероятность события H , заключающегося в появлении в векторе (1) m векторов непрерывного результата.

Вероятность $P(L)$ наступления события L оценивается по биномиальному закону. В соответствии с (3), при определении доверительного интервала (ДИ) вероятности p успешного выполнения ТЗ корректно использование следующих формул:

1) оценка точного ДИ биномиального распределения [3]:

$$\left(\frac{(n - k)}{(n - k) + (k + 1)F_{2(k+1), 2(n-k), 1-\varepsilon/2}}, \frac{(n - k + 1)F_{2(n-k+1), 2k, 1-\varepsilon/2}}{k + (n - k + 1)F_{2(n-k+1), 2k, 1-\varepsilon/2}} \right), \quad (6)$$

где n — число испытаний; k — число ошибок; $F_{f, g, \alpha}$ — квантиль порядка α распределения F с f, g степенями свободы;

2) оценка приближенного ДИ биномиального распределения [3]:

$$\left(p^* - u_{1-\varepsilon/2} \sqrt{\frac{p^*(1-p^*)}{n}}, p^* + u_{1-\varepsilon/2} \sqrt{\frac{p^*(1-p^*)}{n}} \right) = \left(\frac{n-k}{n} - u_{1-\varepsilon/2} \sqrt{\frac{(n-k)k}{n^2}}, \frac{n-k}{n} + u_{1-\varepsilon/2} \sqrt{\frac{(n-k)k}{n^2}} \right), \quad (7)$$

где p^* — оценка эмпирической частоты правильного выполнения ТЗ; $u_{1-\varepsilon/2}$ — квантили стандартного нормального распределения порядка $1 - \varepsilon / 2$.

Геометрическое распределение является единственным дискретным распределением со свойством отсутствия памяти (последствия): знание об отсутствии успеха (неудачи) в предыдущих опытах никак не влияет на распределение оставшегося числа опытов до появления успеха (неудачи). Используем это свойство для оценки вероятности события H , которое заключается в появлении в векторе (2) m векторов непрерывного результата.

В соответствии с ним события $h_s \in H$ появления в составе вектора (2) отдельных векторов $\vec{h}_s$ непрерывного результата относятся к категории попарно несовместных событий. Поэтому вероятность $P(H)$ любого из вариантов их реализаций (см. табл. 1) оценивается по формуле

$$P(H) = P(h_1 \cup \dots \cup h_s \cup \dots \cup h_m) = P(h_1) + \dots P(h_s) + \dots P(h_m), \quad (8)$$

где $s (s = \overline{1, m}; m \leq n)$ — номер вектора непрерывного результата; n — число выполненных ТЗ; $P(h_s)$ — вероятность появления отдельных векторов $\vec{h}_s$ непрерывного результата, оцениваемая в соответствии с (4), (5).

Для оценки в (8) неизвестного параметра p предлагается использовать метод правдоподобия, при котором в качестве оценки вероятности p принимается ее значение

$$\hat{p} = \arg \max_{p \in (0,1)} f(\vec{j}|p) = \arg \max_{p \in (0,1)} P(H), \quad (9)$$

максимизирующее вероятность $P(H)$ появления при выполнении n ТЗ вектора $\vec{J} = (j_1, \dots, j_i, \dots, j_n)$, включающего m ($m = \overline{1, n}$) векторов $\vec{h}_s$ непрерывного результата.

Соответствующая функция $W(p_s, p)$ правдоподобия для любой из реализаций события $H = \cup h_s$ может быть представлена в виде

$$W(p_s, p) = \prod_{s=1}^{s=m} P(h_s) = \prod_{s=1}^{s=l} P(h_s^+) \prod_{s=l+1}^{s=m} P(h_s^-) = p^{(n-k)+(m-l)} q^{k+l}, \quad (10)$$

где k ($k = \overline{0, n}$) — число ТЗ, выполненных с ошибкой; $(n-k)$ ($(n-k) = \overline{0, n}$) — число ТЗ, выполненных правильно; s ($s = \overline{1, m}$) — номера векторов $\vec{h}_s$ непрерывного результата s ($s = \overline{1, l}$); ($s = \overline{l+1, m}$) — соответственно номера векторов непрерывного успеха и непрерывной неудачи; l , ($m-l$) — соответственно число векторов непрерывного успеха и непрерывной неудачи ($m+l \leq n$).

Поиск оценки $\hat{p}$ наибольшего правдоподобия целесообразно реализовать путем решения следующей задачи нелинейного программирования:

$$\left. \begin{aligned} p^{(n-k)+(m-l)} (1-p)^{k+l} &\rightarrow \max_p; \\ p &\in (0, 1). \end{aligned} \right\} \quad (11)$$

Для решения (11) можно использовать стандартную функцию «Solve (Поиск решения)» Microsoft Excel. ДИ оценки $\hat{p}$ может быть определен как

$$\left(\hat{p} - u_{1-\varepsilon/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + u_{1-\varepsilon/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right). \quad (12)$$

Сравним оценки обучаемых, полученные обычным путем на основе предлагаемых моделей.

Пусть обучаемыми выполнено 13 заданий и получены результаты, приведенные в табл. 1. Оценки этих результатов тестирования приведены в табл. 2.

Таблица 2

Оценки результатов тестирования, приведенных в табл. 1

Номер варианта	$n-k$	k	l	$m-l$	Точечная оценка по (1)	Интервальные оценки					
						биномиального распределения			геометрического распределения		
						вероятнейшее значение по (3)	границы ДИ по (6)		вероятнейшее значение по (11)	границы ДИ по (12)	
							нижняя	верхняя		нижняя	верхняя
1	12	0	1	0	1,000	1,000	0,735	1,000	0,923	0,754	1,000
2	6	6	1	1	0,500	0,500	0,211	0,789	0,500	0,182	0,818
3	6	6	1	1	0,500	0,500	0,211	0,789	0,500	0,182	0,818
4	6	6	6	6	0,500	0,500	0,211	0,789	0,500	0,182	0,818
5	6	6	6	6	0,500	0,500	0,211	0,789	0,500	0,182	0,818
6	5	7	2	1	0,417	0,417	0,152	0,723	0,400	0,089	0,711
7	9	3	3	2	0,750	0,750	0,428	0,945	0,647	0,343	0,951
8	10	2	4	2	0,833	0,833	0,516	0,979	0,667	0,367	0,966
9	8	4	2	2	0,667	0,667	0,349	0,901	0,625	0,317	0,933
10	7	5	4	4	0,583	0,583	0,277	0,848	0,550	0,234	0,866

Выводы

На современном уровне развития информационных технологий тестирование является единственным универсальным техническим средством текущего контроля знаний обучаемых [14] – [17]. Процедуры контроля знаний в современных компьютерных системах тестирования не соответствуют предъявляемым к ним требованиям по следующим характеристикам:

- 1) отсутствию возможности оценки погрешности результатов тестирования;
- 2) отсутствию возможности прекращения тестирования при достижении заданной достоверности его результатов

Оценки, приведенные в табл. 2, свидетельствуют о том, что применение предлагаемых моделей интервального оценивания результатов критериально-ориентированного тестирования обеспечивает переход от точечных оценок эмпирической частоты к оценкам вероятности правильного действия, которые имеют количественно оцениваемый ДИ. Последнее обеспечивает применение этих моделей для решения задачи автоматизации управления процессом критериально-ориентированного тестирования в целях обеспечения его оперативности путем остановки процедуры предъявления ТЗ после получения оценки заданной достоверности. Разработанный в статье новый переход от точечных оценок эмпирической частоты успешного выполнения тестовых заданий к интервальным оценкам вероятности этой успешности, позволяет с достаточной точностью определять погрешность и доверительный интервал полученных оценок.

Для эффективного внедрения средств оперативного критериально-ориентированного педагогического контроля знаний в процесс подготовки специалистов ВМФ и обеспечения возможности совместного анализа и обработки оценок результатов обучения, целесообразно реализовать следующие практические рекомендации:

- формулировку системы предпочтений преподавателя в отношении качества подготовки обучаемых и системы предпочтений преподавателя в отношении значимости единичных тестовых заданий в тесте следует рассматривать в качестве наиболее ответственных этапов подготовки теста к практическому использованию;
- преподавателю, впервые производящему ввод своих систем предпочтений, следует проверить их фактическое соответствие на выборке из пяти – десяти таких результатов выполнения теста, которые покрывают весь спектр возможных результатов (от полностью верных до полностью неверных).

СПИСОК ЛИТЕРАТУРЫ

1. *Аванесова Т. П.* Электронное обучение: учеб. пособие / Т. П. Аванесова, А. Н. Печников, А. Н. Шиков. — СПб.: Изд-во Военной академии связи, 2014. — 73 с.
2. *Евсигнеев А. Е.* Концепция развития тестовой технологии контроля уровня обученности студентов в системе профессионального образования России. [Электронный ресурс] / А. Е. Евсигнеев. — Режим доступа: http://www.vevivi.ru/best/downloads.html&req=download&code=confirm_download&id=156184 (дата обращения: 07.11.2016).
3. *Наумов В. Н.* Теоретические основы представления содержания профессиональной подготовки специалистов ВМФ в интеллектуальных обучающих системах: дис. ... д-ра воен. наук / В. Н. Наумов. — СПб.: ВВМИРЭ им. А. С. Попова, 1994. — 296 с.
4. *Переверзев В. Ю.* Технология разработки тестовых заданий: справ. рук. / В. Ю. Переверзев. — М.: Е-Медиа, 2005. — 265 с.
5. *Печников А. Н.* О соответствии результатов критериально ориентированного тестирования целям текущего педагогического контроля / А. Н. Печников, Д. А. Печников // *Фундаментальные и прикладные исследования в современном мире.* — 2015. — № 10-5. — С. 74–85.
6. *Аванесов В. С.* Композиция тестовых заданий: учебная книга для преподавателей вузов, учителей школ, аспирантов и студентов педвузов / В. С. Аванесов. — 4-е изд. — М.: Центр тестирования, 2002. — 237 с.
7. *Ким В. С.* Тестирование учебных достижений: монография / В. С. Ким. — Уссурийск: Изд-во УГПИ, 2007. — 214 с.
8. *Печников А. Н.* Решение задач текущего педагогического контроля на основе анализа результатов критериально-ориентированного тестирования / А. Н. Печников, Д. А. Печников // *Образовательные технологии и общество.* — 2015. — Т. 18. — № 2. — С. 489–513.
9. *Иванов В. Н.* Метод автоматизированной оценки качества профессиональной деятельности в корабельных автоматизированных системах: дис. ... канд. воен. наук / В. Н. Иванов. — СПб.: ВМИРЭ, 2009. — 187 с.

10. Печников А. Н. Проектирование и применение компьютерных технологий обучения. / А. Н. Печников, Ю. А. Ветров. — СПб.: Изд-во БГТУ, 2002. — Ч. 1. Концепция систем автоматизированного обучения и моделирование процессов деятельности. Кн. 1. — 195 с.
11. Ван дер Варден Б. Л. Математическая статистика / Б. Л. Ван дер Варден: пер. с нем. — М.: Иностранная литература, 1960. — 436 с.
12. Печников А. Н. Комплекс методов и моделей оценки качества и эффективности деятельности операторов вооружения и военной техники в тренажерно-обучающих системах ВМФ: дис. ... д-ра техн. наук / А. Н. Печников. — СПб.: СПбГЭТУ, 2011. — 343 с.
13. Печников А. Н. Генератор оценки / А. Н. Печников, А. М. Стручков, Е. В. Хекерт, [и др.]. — Свидетельство о государственной регистрации программ для ЭВМ №2009613362 от 26.06.2009.
14. Платов А. Н. Автоматизация управления контролем подготовки на тренажерах ВМФ: дис. ... канд. техн. наук / А. Н. Платов. — СПб.: 1 ЦНИИ МО РФ, 2000. — 125 с.
15. Пузырев В. Ю. Диагностика деятельности операторов корабельных электромеханических систем в специализированных тренажерах ВМФ: дис. ... канд. техн. наук / В. Ю. Пузырев. — Новороссийск: НГМА, 2005. — 129 с.
16. Чельщикова М. Б. Адаптивное тестирование в образовании (теория, методология, технология) / М. Б. Чельщикова. — М.: Исслед. центр проблем качества подготовки специалистов, 2001. — 165 с.
17. Чельщикова М. Б. Теория и практика конструирования педагогических тестов: учеб. пособие / М. Б. Чельщикова. — М.: Логос, 2002. — 432 с.

DEVELOPMENT OF MODELS AND PROCEDURES FOR INTERVAL ESTIMATION OF THE RESULTS OF CONTROL OF KNOWLEDGE IN COMPUTER SYSTEMS TESTING, NAVY

This article discusses the development of models and procedures for interval estimation of knowledge of test results in the computer systems using methods based on the laws of the binomial distribution and geometric distribution, as a special case of the negative binomial distribution. A new transition from the point estimates of the empirical relative frequency successful execution of tests for interval estimates ve probability of success, allowing determination of the error and the confidence interval of the estimates. Also in the article are designed criterion-based model otse-nivaniya computer testing systems, proven data model that increases the results of the re-curent pedagogical control. To carry out military training experts of the Navy requires appropriate training facilities. The basis of such a framework had always constituted existing training and tre-nirovochnye samples of military equipment and weapons that had previously acted in naval schools in sufficient numbers and to ensure the necessary level of both theoretical and practical training of future naval officers. As a result of the sharp contraction in 1990, branches of the military-industrial complex, developing a techno-ku and weapons for the Navy, the newest samples of the equipment and weapons are now manufactured Xia not series and individual pieces. The issue of education and training samples production is not worth at all. In these circumstances, as the only means capable of providing military training, can be considered only the virtual counterparts of military equipment and weapons, developed by means of information and technologies, ie, trenazherno-learning systems (TOS) of the Navy. Development and application of the proposed model in this paper interval estimation criterion-oriented test results provide a transition from an empirical point estimates often-sti to the estimated probability of correct actions, which have quantified CI. The latter provides for the use of these models to solve the problem of automation of process control criterion-based testing to ensure its efficiency by stopping the procedure after receiving the submission of TK assessment given credibility.

Keywords: interval estimation, confidence intervals, binomial distribution.

REFERENCES

1. Avanesova, T. P., A. N. Pechnikov, and A. N. Shikov. *Jelektronnoe obuchenie: uchebnoe posobie*. SPb.: VAS, 2014.
2. Evsigneev, A. E. *Koncepcija razvitija testovoj tehnologii kontrolja urovnja obuchennosti studentov v sisteme professionalnogo obrazovanija Rossii*. Web. 7 Nov. 2016 <http://www.vevivi.ru/best/downloads.html&req=download&code=confirm_download&id=156184>.

3. Naumov, V. N. *Teoreticheskie osnovy predstavlenija soderzhanija professionalnoj podgotovki specialistov VMF v intellektualnyh obuchajushhijh sistemah*. Dr. diss. Petrodvorec: VVMIRJe im. A. S. Popova, 1994.
4. Pereverzev, V. Ju. *Tehnologija razrabotki testovyh zadaniy: spravocnoe rukovodstvo*. M.: E-Media, 2005.
5. Pechnikov, A. N., and D. A. Pechnikov. "Relevant findings criterion-oriented testing purpose of this pedagogical control." *Fundamentalnye i prikladnye issledovanija v sovremenom mire* 10-5 (2015): 74–85.
6. Avanesov, V. S. *Kompozicija testovyh zadaniy. Uchebnaja kniga dlja prepodavatelej vuzov, uchitelej shkol, aspirantov i studentov pedvuzov*. 4-th ed. M.: Centr testirovanija, 2002.
7. Kim, V. S. *Testirovanie uchebnyh dostizhenij: Monografija*. Ussurijsk: Izdatelstvo UGPI, 2007.
8. Pechnikov, A. N., and D. A. Pechnikov. "Reshenie zadach tekushhego pedagogicheskogo kontrolja na osnove analiza rezultatov kriterialno-orientirovannogo testirovanija." *Obrazovatelnye tehnologii i obshhestvo* 18.2 (2015): 489–513.
9. Ivanov, V. N. *Metod avtomatizirovannoj ocenki kachestva professionalnoj dejatel'nosti v korabel'nyh avtomatizirovannyh sistemah*. PhD diss. Petrodvorec: VMIRJe, 2009.
10. Pechnikov, A. N., and Ju. A. Vetrov. *Proektirovanie i primenenie kompjuternyh tehnologij obuchenija. Ch.1. Konceptcija sistem avtomatizirovannogo obuchenija i modelirovanie processov dejatel'nosti. Kn.1*. SPb: BGTU, 2002.
11. Van der Varden, B. L. *Matematicheskaja statistika*. Translated. M.: Izd-vo Inostrannoj literatury, 1960.
12. Pechnikov, A. N. *Kompleks metodov i modelej ocenki kachestva i jeffektivnosti deja-telnosti operatorov vooruzhenija i voennoj tehniki v trenazherno-obuchajushhijh sistemah VMF*. Dr. diss. (Tech.). SPb.: SPbGJeTU, 2011.
13. Pechnikov, A. N., A. M. Struchkov, E. V. Hekert, et al. *Generator ocenki: The computer program*. Certificate № 2009613362. Publ. 26.06.2009.
14. Platon, A. N. *Avtomatizacija upravljenja kontrolem podgotovki na trenazherah VMF*. Phd diss. SPb.: I CNII MO RF, 2000.
15. Puzyrev, V. Ju. *Diagnostika dejatel'nosti operatorov korabelnyh jelektromehanicheskijh sistem v specializirovannyh trenazherah VMF*. PhD diss. Novorossijsk: NGMA, 2005.
16. Chelyshkova, M. B. *Adaptivnoe testirovanie v obrazovanii (teorija, metodologija, teh-nologija)*. M.: Issledovatel'skij centr problem kachestva podgotovki specialistov, 2001.
17. Chelyshkova, M. B. *Teorija i praktika konstruirovaniya pedagogicheskijh testov: Uchebnoe posobie*. M.: Logos, 2002.

ИНФОРМАЦИЯ ОБ АВТОРЕ

Печников Денис Андреевич —
кандидат технических наук, доцент.
Военно-морская академия имени Адмирала Флота
Советского Союза Н. Г. Кузнецова
19pda72@bk.ru

INFORMATION ABOUT THE AUTHOR

Pechnikov Denis Andreevich —
PhD, associate professor.
Naval Academy named after Admiral
of the Fleet of the Soviet Union N.G. Kuznetsova
19pda72@bk.ru

Статья поступила в редакцию 10 ноября 2016 г.