ДОКЛАДЫ АН ВШ РФ

2011 июль–декабрь № 2 (17)

ТЕХНИЧЕСКИЕ НАУКИ

УДК 62-83: 531.3

ПРОВЕРКА ПРОСТЫХ И СЛОЖНЫХ ГИПОТЕЗ С ИСПОЛЬЗОВАНИЕМ ПОСЛЕДОВАТЕЛЬНОГО КРИТЕРИЯ ВАЛЬДА

С.Н. Постовалов

Новосибирский государственный технический университет

postovalov@ngs.ru

Представлены результаты исследований разработанного автором алгоритма определения критических границ последовательного критерия отношения правдоподобия Вальда с помощью компьютерного моделирования. Представлены результаты численного моделирования границ критерия на примере простой и сложной гипотез о нормальном распределении против конкурирующей гипотезы о логистическом распределении. Проведено сравнение среднего количества испытаний для различения гипотез.

Ключевые слова: последовательный анализ, критерий отношения правдоподобия Вальда, простые и сложные гипотезы.

Введение

Целью первичной обработки экспериментальных наблюдений обычно является выбор закона распределения, наиболее хорошо описывающего случайную величину, выборку которой мы наблюдали. Проверка того, насколько хорошо наблюдаемая выборка описывается тем или иным теоретическим законом, осуществляется с использованием различных критериев [1]. Целью проверки гипотезы о принадлежности выборки конкретному закону является стремление удостовериться в том, что данная модель теоретического закона не противоречит наблюдаемым данным, и использование ее не приведет к существенным ошибкам при вероятностных расчетах.

В отличие от классических методов математической статистики, в которых число производимых экспериментов фиксируется заранее, методы последовательного анализа характеризуются тем, что момент прекращения наблюдений является случайным и определяется наблюдателем в зависимости от значений наблюдаемых данных. Преимущество последовательных методов было продемонстрировано Вальдом [2] на задаче различения двух простых гипотез, установившим, что такие методы дают выигрыш в среднем числе наблюдений по сравнению с любым другим способом с фиксированным объемом выборки и теми же вероятностями ошибок первого и второго рода. Вальд указал и тот последовательный метод, названный им критерием последовательных отношений вероятностей

Работа выполнена в рамках государственного контракта №П2365 по ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009–2013 гг. (проект 421П/27)

(SPRT), который оказался оптимальным в классе всех последовательных методов.

В этой работе представлен алгоритм, который позволяет численно определить критические границы для любой пары заданных простых или сложных гипотез.

1. Постановка задачи

Пусть в результате экспериментов наблюдаются значения случайной величины: $x_1, x_2, \ldots$, которые являются независимыми и одинаково распределенными. Эксперименты продолжаются до тех пор, пока не будет принято решение о виде распределения случайной величины. Объем выборки заранее не известен и определяется в процессе последовательной процедуры.

Пусть основная гипотеза о виде распределения случайной величины имеет вид $H_0 : f(x) = f_0(x, \theta)$, где $f(x)$ – функция плотности, а θ – известный в случае простой гипотезы или неизвестный в случае сложной гипотезы параметр распределения. Конкурирующая гипотеза имеет вид $H_1 : f(x) = f_1(x)$. При проверке сложной гипотезы оценка параметра $\hat{\theta}$ вычисляется по той же самой выборке, по которой проверяется гипотеза.

Последовательный критерий отношения правдоподобия строится следующим образом. Сначала выбирают критические границы c_0 и c_1 ($c_0 < c_1$). После первого наблюдения x_1 вычисляют значение статистики критерия

$$\lambda_1 = \ln\left(\frac{L_1(X_1)}{L_0(X_1)}\right) = \ln\left(\frac{f_1(x_1)}{f_0(x_1)}\right),$$

где L_0 – функция правдоподобия при законе, соответствующем гипотезе H_0, а L_1 – при законе, соответствующем гипотезе H_1.

Если $\lambda_1 < c_0$, то принимают гипотезу H_0; если $\lambda_1 > c_1$, принимают H_1; если $c_0 \le \lambda_1 \le c_1$, то производят второе наблюдение x_2 и также исследуют величину

$$\lambda_2 = \ln\left(\frac{L_1(X_2)}{L_0(X_2)}\right) = \ln\left(\frac{f_1(x_1) f_1(x_2)}{f_0(x_1) f_0(x_2)}\right).$$

Известно [3], что с вероятностью, равной единице, процесс оканчивается либо выбором H_0, либо выбором H_1, т.е. на каком-то шаге n статистика

$$\lambda_n = \ln\left(\frac{L_1(X_n)}{L_0(X_n)}\right) = \sum_{i=1}^{n} \ln\left(\frac{f_1(x_i)}{f_0(x_i)}\right)$$

выйдет за интервал $[c_0, c_1]$.

Величины c_0 и c_1 определяются из условия, чтобы вероятности ошибок первого и второго рода имели заданные значения α и β. Обозначим эту зависимость через $c_0(\alpha,\beta)$ и $c_1(\alpha,\beta)$. В случае проверки простой гипотезы известна следующая теорема [3].

Теорема. *Критические значения* c_0 *и* c_1 *критерия Вальда удовлетворяют неравенствам:*

$$c_0 \geq \ln\left(\frac{\beta}{1-\alpha}\right), \quad c_1 \leq \ln\left(\frac{1-\beta}{\alpha}\right). \tag{1}$$

В критерии Вальда критические границы неизвестны, точнее сказать они зависят от конкретных проверяемых гипотез, и поэтому на практике используют оценки этих границ, которые в явном виде зависят от вероятностей ошибок первого и второго рода. Однако использование этих оценок приводит к тому, что фактические вероятности ошибок первого и второго рода будут меньше и это приведет к увеличению количества необходимых испытаний.

Обычно на практике вместо точных критических границ используют оценки (1). Точные границы c_0 и c_1 сложно вычислить аналитически для произвольной пары гипотез, однако это можно сделать с помощью технологии компьютерного моделирования, успешно применяющейся авторами для исследования множества статистик, используемых на практике [5–11].

2. Алгоритм нахождения критических значений

В процессе исследований нахождение критических значений статистики критерия Вальда осуществлялось по следующей схеме.

Этап 1. Моделирование функций вероятности ошибок первого и второго рода. Методом Монте-Карло моделируется зависимость $\alpha(c_0, c_1)$ и $\beta(c_0, c_1)$, путем перебора всех значений c_0 из интервала $[\underline{c_0}, 0]$, c_1 из интервала $[0, \overline{c_1}]$ с малым шагом δ, где в случае простой гипотезы

$$\underline{c_0} = \ln\left(\frac{\beta}{1-\alpha}\right) \text{ и } \overline{c_1} = \ln\left(\frac{1-\beta}{\alpha}\right) \tag{2}$$

определяются условиями (1), а в случае сложной гипотезы область моделирования находится экспериментально.

Для определения вероятности ошибки α по заданным значениям критических границ c_0, c_1 применятся следующий алгоритм.

– моделируется последовательность реализаций случайной величины по закону, соответствующему гипотезе H_0, до тех пор, пока не

произойдет принятие одной из конкурирующих гипотез. Запоминается, какая гипотеза была принята;
- моделирование повторяется N раз;
- эмпирическая вероятность ошибки первого рода будет равна отношению числа случаев, когда ошибочно была принята гипотеза H_1 к числу N.

Согласно центральной предельной теореме при объеме моделирования $N = 100000$, использованном для получения численных значений критических границ в данной работе, эмпирическая вероятность ошибки первого рода будет отклоняться от истинного значения не более чем на 0.004 с доверительной вероятностью 0.99 [4].

Этап 2. Построение линий равного уровня.

Для заданных значений вероятностей ошибок первого рода $\alpha_i = 0.15, 0.1,\ 0.05,\ 0.01$ и второго рода $\beta_j =\ 0.15, 0.1,\ 0.05,\ 0.01$ строятся линии равного уровня $\alpha(c_0, c_1) = \alpha_i$ и $\beta(c_0, c_1) = \beta_j$.

Этап 3. Вычисление критических границ.

Пересечение линий равного уровня дает искомые критические значения для заданных вероятностей ошибок первого и второго уровня (α_i, β_j).

3. Проверяемые гипотезы

Нахождение критических значений будет рассмотрено на примере проверки простой и сложной гипотез о нормальном распределении против конкурирующей гипотезы о логистическом распределении:

$$H_0 : f(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{(x-\theta)^2}{2} \right\},$$

$$H_1 : f(x) = \frac{\pi}{\sqrt{3}} \exp\left\{ -\frac{\pi x}{\sqrt{3}} \right\} \Bigg/ \left[1 + \exp\left\{ -\frac{\pi x}{\sqrt{3}} \right\} \right]^2,$$

где параметр сдвига θ равен нулю в случае простой гипотезы. Эти два закона близки и трудно различаемы с помощью критериев согласия. В случае сложной гипотезы H_0 параметр сдвига оценивается по методу максимального правдоподобия.

4. Определение критических значений для проверки простой гипотезы о нормальном распределении против простой конкурирующей гипотезы о логистическом законе

Методами статистического моделирования были получены функции $\alpha(c_0, c_1)$ и $\beta(c_0, c_1)$, на основании которых были построены линии равного уровня (рис. 1). Выбранный интервал моделирования определялся наименьшими рассматриваемыми значениями вероятностей ошибок первого и второго рода. При $\alpha = 0.01$ и $\beta = 0.01$ границы, соответствующие (2), равны: $\underline{c}_0 = -4.6$ и $\overline{c}_1 = 4.6$. Критические значения, полученные в результате пересечения линий равного уровня функций $\alpha(c_0, c_1)$ и $\beta(c_0, c_1)$, приведены в табл. 1.

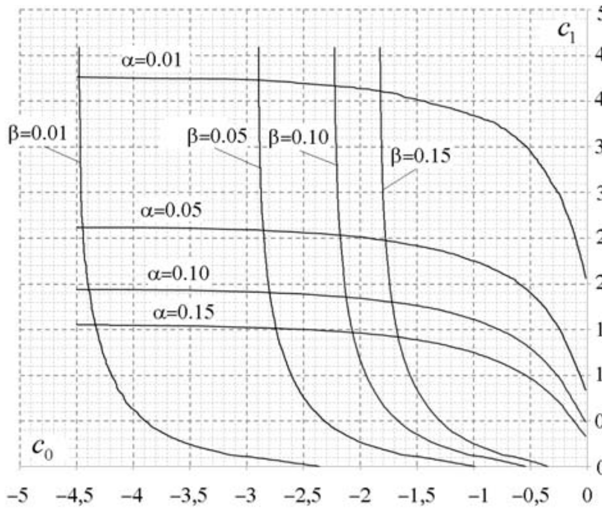


Рис. 1. Линии равного уровня функции $\alpha(c_0, c_1)$ и $\beta(c_0, c_1)$ при проверке простой гипотезы о нормальном законе против простой конкурирующей гипотезы о логистическом законе

Таблица 1

Критические значения c_0, c_1 при проверке простой гипотезы о нормальном распределении против простой конкурирующей гипотезы о логистическом законе

α \ β	0.15	0.10	0.05	0.01
0.15	–1.67, 1.42	–2.07, 1.47	–2.74, 1.52	–4.33, 1.56
0.10	–1.72, 1.81	–2.12, 1.86	–2.80, 1.91	–4.39, 1.95
0.05	–1.78, 2.48	–2.18, 2.54	–2.85, 2.59	–4.45, 2.63
0.01	–1.82, 4.11	–2.22, 4.17	–2.89, 4.23	–4.48, 4.26

5. Определение критических значений для проверки сложной гипотезы о нормальном распределении против простой конкурирующей гипотезы о логистическом законе

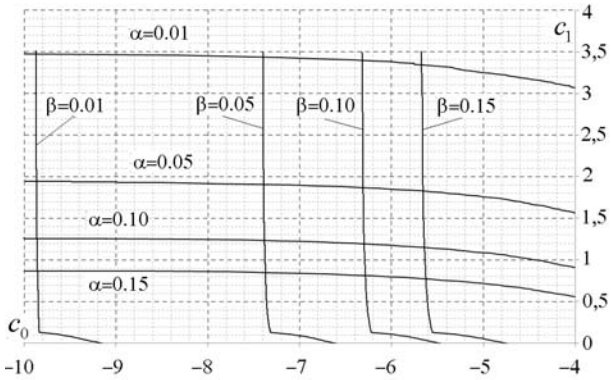
Применение критерия Вальда для проверки сложных гипотез затруднено тем, что теорема об оценках границ (1) в данном случае не имеет места. Поэтому перед применением алгоритма нахождения критических границ необходимо найти область, на которой следует осуществлять моделирование. Для этого выполняется предварительное моделирование на заведомо большей области с большим шагом и малым количеством экспериментов. На основании предварительного шага было выявлено, что функции $\alpha(c_0,c_1)$ и $\beta(c_0,c_1)$ определены (т.е. отличны от 0 и 1 при заданном объеме моделирования) при $c_0 \in [-30,2]$ и $c_1 \in [-1,5]$.

Для проверки зависимости критических границ от истинного значения оцениваемого параметра было рассмотрено два варианта моделирования:

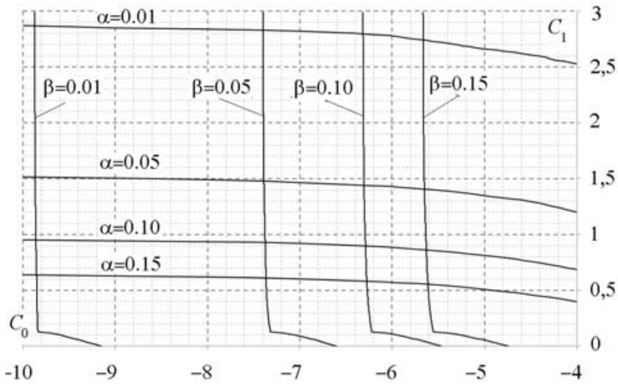
а) при верной основной гипотезе выборка моделировалась по стандартному нормальному закону;

б) при верной основной гипотезе выборка моделировалась по нормальному закону с параметром сдвига $\theta = 0{,}1$.

В результате проведенных исследований были получены функции $\alpha(c_0,c_1)$ и $\beta(c_0,c_1)$, на основании которых были построены линии равного уровня (рис. 2).



а)



б)

Рис. 2. Линии равного уровня функции $\alpha(c_0,c_1)$ и $\beta(c_0,c_1)$ при проверке сложной гипотезы о нормальном законе против простой конкурирующей гипотезы о логистическом законе, при моделировании: а) стандартного нормального распределения; б) нормального распределения со сдвигом 0.1

Хорошо видно, что по сравнению с рис. 1 линии равного уровня для $\alpha(c_0,c_1)$ и $\beta(c_0,c_1)$ сместились вниз и влево, в результате чего значения точек их пересечения стали существенно меньше. Данный результат вполне логичен, так как при оценивании параметров закона, соответствующего гипотезе H_0, знаменатель отношения правдоподобия увеличивается, а статистика отношения правдоподобия уменьшается.

Также была выявлена зависимость критических границ от истинного значения оцениваемого параметра, причем при увеличении расстояния между основной и альтернативной гипотезой критические границы c_0 и c_1 становятся ближе.

Критические значения, полученные в результате пересечения линий равного уровня функций $\alpha(c_0, c_1)$ и $\beta(c_0, c_1)$, приведены в табл. 2.

6. Вычисление среднего количества экспериментов, необходимого для различения конкурирующих гипотез

Так как мощность последовательных критериев определяется выбором критических значений c_0 и c_1, то для сравнения различных критериев логично использовать среднее количество экспериментов, необходимых для различения пары гипотез.

Сравнение критерия Вальда с использованием приближенных критических значений по формуле (2) и с использованием полученных в данной работе критических границ, приведенных в табл. 1, было проведено методами компьютерного моделирования на той же паре конкурирующих гипотез «нормальный закон – логистический».

Таблица 2

Критические значения c_0, c_1 при проверке сложной гипотезы о нормальном распределении против простой конкурирующей гипотезы о логистическом законе, при моделирование в случае верной основной гипотезы стандартного нормального распределения

α \ β	0.15	0.10	0.05	0.01
0.15	–5.63, 0.81	–6.30, 0.85	–7.37, 0.88	–9.93, 0.92
0.10	–5.64, 1.18	–6.31, 1.23	–7.38, 1.27	–9.93, 1.31
0.05	–5.66, 1.86	–6.32, 1.90	–7.39, 1.93	–9.93, 1.99
0.01	–5.67, 3.37	–6.33, 3.43	–7.40, 3.46	–9.94, 3.49

В табл. 3–7 приведены оценки среднего количества экспериментов, необходимых для различения двух рассматриваемых конкурирующих гипотез, полученные в результате исследования методами статистического моделирования. При сравнении табл. 3 и 5 (а также 4 и 6) хорошо виден выигрыш по среднему количеству экспериментов в выборке при использовании найденных критических значений, по сравнению с использованием оценок по формуле (2). Сокращение среднего количества экспериментов составляет от 3% до 11%, когда верна основная гипотеза; и от 7% до 17%, когда верна конкурирующая гипотеза.

Таблица 3

Среднее количество экспериментов необходимое для различения двух простых гипотез с использованием полученных критических значений (при верной гипотезе H_0)

α \ β	0.15	0.10	0.05	0.01
0.15	118	150	203	332
0.01	135	168	226	362
0.05	155	191	251	395
0.01	174	212	275	425

Таблица 4

Среднее количество экспериментов необходимое для различения двух простых гипотез с использованием полученных критических значений (при верной гипотезе H_1)

α \ β	0.15	0.10	0.05	0.01
0.15	102	116	132	148
0.10	126	142	160	177
0.05	168	187	207	227
0.01	268	293	319	343

Таблица 5

Среднее количество экспериментов необходимое для различения двух простых гипотез с использованием приближенных критических значений (при верной основной гипотезе)

α \ β	0.15	0.10	0.05	0.01
0.15	133	167	226	362
0.10	149	184	245	387
0.05	166	203	266	413
0.01	182	220	286	437

Таблица 6

Среднее количество экспериментов необходимое для различения двух простых гипотез с использованием приближенных критических значений (при верной конкурирующей гипотезе)

α \ β	0.15	0.10	0.05	0.01
0.15	123	138	156	173
0.10	149	166	185	203
0.05	193	212	234	253
0.01	293	318	344	368

Таблица 7

Среднее количество экспериментов необходимое для различения двух гипотез (основная – сложная, конкурирующая – простая) при верной основной гипотезе, при моделировании в случае верной основной гипотезы стандартного нормального закона

α \ β	0.15	0.10	0.05	0.01
0.15	296	360	461	691
0.10	321	389	496	736
0.05	352	422	533	786
0.01	377	451	567	827

При сравнении результатов в табл. 3 и 7 видно, что для проверки сложной гипотезы в среднем требуется существенно большее число экспериментов.

7. Сравнение последовательного критерия отношения правдоподобия с критериями согласия

В работах [12–13] для пары конкурирующих гипотез (8)–(9) приведены результаты сравнительного анализа мощности критериев согласия Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга, χ^2 Пирсона, Рао–Робсона–Никулина. В случае простой гипотезы наибольшей мощностью среди рассмотренных критериев [12–13] обладает критерий χ^2 Пирсона при асимптотически оптимальном группировании [12,14]. При заданных $\alpha = 0.1$ и $\beta = 0.078$ для различения конкурирующих гипотез по критерию χ^2 Пирсона требуется порядка 1000 экспериментов. В то время как последовательный критерий отношения правдоподобия Вальда при $\alpha = 0.1$ и $\beta = 0.05$ позволяет различать эти гипотезы в среднем за 226 экспериментов при использовании найденных границ (см. табл. 3) и 245 экспериментов при использовании приближенных границ (см. табл. 5), т.е. примерно в четыре раза быстрее, причем использование полученных автором критических границ позволяет уменьшить число экспериментов на 19 (примерно на 8%).

В случае сложной проверяемой гипотезы наибольшей мощностью среди рассмотренных критериев согласия обладает критерий Андерсона–Дарлинга [13]. При 1000 экспериментов и заданной вероятности $\alpha = 0.1$ для рассматриваемых конкурирующих гипотез критерий Андерсона–Дарлинга обеспечивает вероятность ошибки 2-го рода β не более 0.019. В то время как последовательному критерию отношения правдоподобия Вальда для различения этих гипотез при $\alpha = 0.1$ и $\beta = 0.01$ потребуется в среднем 736 экспериментов (см. табл. 7).

8. Вычислительные затраты для определения точных критических границ

Важным вопросом для использования точных критических границ на практике является время, которое необходимо для их вычисления. В табл. 8 приведено время вычисления оценок вероятностей ошибок первого рода для разных видов гипотез, рассмотренных в данной работе, при объеме моделирования 100 000 серий на компьютере с процессором Intel Core 2 Quad Q6600 с тактовой частотой 2.4 ГГц. Наименьшее время потребовалось при простой гипотезе, а наибольшее, около трех суток, – при сложной гипотезе, когда при моделировании использовался параметр сдвига, равный нулю.

Следует отметить, что данная вычислительная задача допускает распараллеливание, т.к. серии экспериментов можно моделировать параллельно, а затем складывать полученные матрицы, в которых хранятся количества ошибок первого и второго рода.

Таблица 8

Время моделирования вероятностей ошибок первого и второго рода

Гипотеза	Область	Шаг	Время моделирования α, мин.	Время моделирования β, мин.	Общее время моделирования, мин.
Простая, $\theta = 0$	[–4,6; 0]x[0; 4,6]	0,01	675	592	1267
Сложная, $\theta = 0$	[–10;0]x[–1;4]	0,01	2763	1348	4111
Сложная, $\theta = 0{,}1$	[–10;0]x[–1;3]	0,01	1445	857	2302

Заключение

В результате проведенных исследований можно подвести следующие итоги.

Показано, что методами компьютерного моделирования могут быть получены критические значения для последовательного критерия отношения правдоподобия Вальда, обеспечивающие при заданных вероятностях ошибок 1-го и 2-го рода минимальное количество требуемых экспериментов при проверке простой и сложной гипотез относительно заданной конкурирующей гипотезы. Построены таблицы критических значений для проверки гипотезы о нормальном распределении против простой конкурирующей гипотезы о логистическом законе. Данный подход можно распространить и на другие вероятностные модели.

Использование полученных критических значений в последовательном критерии отношения правдоподобия Вальда позволяет быстрее различить две простые гипотезы. Выигрыш составляет от 3% до 17%, что в случае проведения дорогостоящих экспериментов является существенным фактором.

В случае проверки сложной гипотезы по критерию Вальда использование оценок (1) критических значений является неправомерным. Критические границы критерия Вальда при проверке сложной гипотезы зависят от истинных значений параметров наблюдаемого закона распределения. При проверке сложной гипотезы требуемое для принятия решения число экспериментов существенно больше, чем при проверке простой гипотезы.

Сравнение последовательного критерия отношения правдоподобия с критериями (согласия) с фиксированным объемом выборки показало, что критерий Вальда позволяет различать заданную пару гипотез при существенно меньшем объеме выборки. Однако при этом надо учитывать, что критерий Вальда рассматривает конкретную конкурирующую гипотезу, а в критериях согласия задана только проверяемая гипотеза.

ЛИТЕРАТУРА

[1] **Леман Э.** Проверка статистических гипотез. М.: Наука, 1979. – 408 с

[2] **Wald, A.** Sequential Tests of Statistical Hypotheses // *The Annals of Mathematical Statistics*. – 1945. – V. 16 (2). – P. 117–186.

[3] **Вальд А.** Последовательный анализ: Пер. с англ./ Под ред. Б.А.Севастьянова. – М.: Физматгиз, 1960. – 328 с.

[4] **Лемешко Б.Ю., Постовалов С.Н.** Компьютерные технологии анализа данных и исследования статистических закономерностей: Учебное пособие. – Новосибирск: Изд-во НГТУ, 2004. – 119 с.

[5] **Лемешко Б.Ю., Постовалов С.Н.** О распределениях статистик непараметрических критериев согласия при оценивании по выборкам параметров наблюдаемых законов // *Заводская лаборатория*. – 1998. – Т. 64. – № 3. – С. 61–72.

[6] **Лемешко Б.Ю., Чимитова Е.В.** Максимизация мощности критериев типа χ^2 // *Доклады СО АН ВШ*. – 2000. – № 2. – С. 53–61.

[7] **Лемешко Б.Ю., Постовалов С.Н.** Применение непараметрических критериев согласия при проверке сложных гипотез // *Автометрия*. – 2001. – № 2. – С. 88–102.

[8] **Лемешко Б.Ю., Постовалов С.Н., Чимитова Е.В.** О распределениях статистики и мощности критерия типа $\chi2$ Никулина // *Заводская лаборатория. Диагностика материалов*. – 2001. – Т. 67. – № 3. – С. 52–58.

[9] **Лемешко Б.Ю., Постовалов С.Н.** О зависимости распределений статистик непараметрических критериев и их мощности от метода оценивания параметров // *Заводская лаборатория. Диагностика материалов*. – 2001. – Т. 67. – № 7. – С. 62–71.

[10] **Лемешко Б.Ю., Постовалов С.Н.** Непараметрические критерии при проверке сложных гипотез о согласии с распределениями Джонсона // *Доклады СО АН ВШ*. – 2002. – № 1(5). – С. 65–74.

[11] **Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н.** Мощность критериев согласия при близких альтернативах // *Измерительная техника*. – 2007. – № 2. – С.22–27.

[12] **Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н.** Сравнительный анализ мощности критериев согласия при близких конкурирующих гипотезах. I. Проверка простых гипотез // *Сиб. журнал индустр. математики*. – 2008. – Т. 11. – № 2(34). – С.96–111.

[13] **Лемешко Б.Ю., Лемешко С.Б., Постовалов С.Н.** Сравнительный анализ мощности критериев согласия при близких альтернативах. II. Проверка сложных гипотез // *Сиб. журнал индустр. математики*. – 2008. – Т. 11. – № 4(36). – С.78–93.

[14] **Лемешко Б.Ю.** Асимптотически оптимальное группирование наблюдений в критериях согласия // *Заводская лаборатория*. – 1998. – Т. 64. – № 1. – С.56–64.

Postovalov S.N.

SIMPLE AND COMPOSITE HYPOTHESES TESTING BY SEQUENTIAL WALD'S TEST

The results of computer simulation algorithm for determining critical boundaries of sequential likelihood ratio test are presented. The results of numerical simulation of critical boundaries for simple and composite hypothesis of normal distribution versus the alternative of the logistic distribution are presented. A comparison of the average sample number of tests for testing hypotheses is shown.

Keywords: Sequential analysis, SPRT, Simple and composite hypotheses.

Статья поступила 8 сентября 2010 г.,
после доработки 10 января 2011 г.