УДК 519.2

А.Ю. ХАРИН, С.Ю. ЧЕРНОВ

ОЦЕНИВАНИЕ ВЕРОЯТНОСТЕЙ ОШИБОК ПОСЛЕДОВАТЕЛЬНОГО КРИТЕРИЯ ОТНОШЕНИЯ ВЕРОЯТНОСТЕЙ

The problem of the error probabilities evaluation for the sequential probability ratio test is considered. Lower and upper estimates for the error probabilities are constructed and the accuracy of the approximation is analysed. Numerical results are given to illustrate the theory.

Последовательный анализ [1] используется для статистического решения многих практических задач, в которых имеется необходимость проверки гипотетических предположений о параметрах исследуемого процесса. Такой подход позволяет затратить в среднем наименьшее число наблюдений [2] для принятия решений среди всех возможных статистических критериев с такими же значениями вероятностей ошибок I и II рода. Однако вычисление точных значений вероятностей ошибочных решений последовательного критерия представляет собой задачу, решенную лишь для некоторых частных случаев [3]. В [4] эта задача решена для семейства дискретных распределений вероятностей. В [5] предложен подход для оценивания указанных вероятностей сверху и снизу. С целью развития этого подхода в данной работе построены две специальные цепи Маркова, позволяющие вычислять верхнюю и нижнюю границы в явном виде, а также проведен асимптотический анализ полученных оценок.

Математическая модель

Пусть на измеримом пространстве $(\Omega, \Im)$ наблюдается последовательность независимых одинаково распределенных случайных величин $x_1, x_2, \ldots \in \mathbf{R}$, имеющих плотность распределения вероятностей $f(x; \theta)$ с параметром $\theta \in \Theta = \{\theta_0, \theta_1\}$, истинное значение которого неизвестно. Обозначим функцию распределения вероятностей наблюдений x_i, $i = 1, 2, \ldots$, через $F(x; \theta)$; $\Lambda_n = \Lambda_n(x_1, \ldots, x_n) = \sum_{t=1}^{n} \lambda_t$, где

$$\lambda_t = \lambda(x_t) = \ln(f(x_t, \theta_1) / f(x_t, \theta_0)) - \tag{1}$$

логарифм статистики отношения правдоподобия, вычисленной по одному наблюдению x_t.

Относительно параметра θ имеются две простые гипотезы:

$$H_0: \ \theta = \theta_0, \ \ H_1: \ \theta = \theta_1. \tag{2}$$

Для проверки гипотез (2) используется последовательный критерий отношения вероятностей (ПКОВ) [1]:

$$N = \min\{n \in \mathrm{N} : \Lambda_n \notin (C_-, C_+)\}, \tag{3}$$

$$d = \mathbf{1}_{[C_+, +\infty)}(\Lambda_N), \tag{4}$$

где N – случайный момент остановки, после которого принимается решение d в соответствии с (4). В (3), (4) C_-, $C_+ \in \mathbf{R}$, $C_- < C_+$ – заданные параметры критерия, называемые порогами. На практике для их задания пользуются соотношениями [1] $C_- = \ln\left(\beta_0 / (1 - \alpha_0)\right)$, $C_+ = \ln\left((1 - \beta_0) / \alpha_0\right)$, где α_0, $\beta_0 \in (0,1)$ – величины, близкие к приемлемым значениям вероятностей ошибок I и II рода. В [6] показано, что фактические значения вероятностей ошибок I и II рода α, β последовательного критерия (3), (4) могут значительно отличаться от α_0, β_0.

Сформулируем следующие предположения: П1) функция $f(x; \theta)$ имеет конечные производные 1-го и 2-го порядка по переменной x, а также $f(x; \theta) \neq 0$, $\theta \in \Theta$; П2) функция $\lambda(x)$, определенная (1), строго монотонна по переменной x, а также имеет отличную от нуля производную 1-го порядка.

Данным предположениям удовлетворяют, например, представители экспоненциального семейства распределений вероятностей, у которых плотность распределения имеет вид $f(x, \theta) = a(x) b(\theta) \exp\{c(x) d(\theta)\}$, где 1) $a(x)$, $c(x)$ дважды дифференцируемы, а также $a(x) \neq 0$, $x \in \mathrm{D}$ и $b(\theta) \neq 0$, $\theta \in \Theta$; 2) $\operatorname{sign} c'(x) = \mathrm{const}$, $d(\theta_0) \neq d(\theta_1)$.

Без ограничения общности будем считать, что истинной гипотезой является H_0 (случай H_1 рассматривается аналогично). Таким образом, задача состоит в вычислении вероятности ошибки I рода для ПКОВ (3), (4).

Теоретические результаты

Разобьем интервал (C_-, C_+) на m промежутков длиной $h = (C_+ - C_-)/m$, где $m \in \mathrm{N}$ – параметр разбиения (аппроксимации). Введем случайные последовательности

$$\Lambda_n^- = \sum_{t=1}^{n} \lambda_t^-, \ \ \Lambda_n^+ = \sum_{t=1}^{n} \lambda_t^+;$$

$$\lambda_1^- = C_- + \left[\frac{\lambda_1 - C_-}{h}\right] h, \ \ \lambda_t^- = \left[\frac{\lambda_t}{h}\right] h, \ \ t \geq 2; \ \ \lambda_i^+ = \lambda_i^- + h, \ \ i \in \mathbf{N}.$$

Последовательности Λ_n^-, Λ_n^+ – однородные цепи Маркова, так как имеют независимые приращения, поскольку являются суммами борелевских функций от независимых одинаково распределенных случайных величин.

Построим поглощающие цепи Маркова L_n^-, L_n^+ со множеством значений $\{0, 1, \ldots, m, m+1\}$ и двумя поглощающими состояниями 0 и $m + 1$:

$$L_n^- = \begin{cases} 0, & \Lambda_n^- \in (-\infty, C_- - h], \\ i, & \Lambda_n^- = C_- + (i-1)h, \quad i = \overline{1,m}, \\ m+1, & \Lambda_n^- \in [C_+, \infty), \end{cases} \quad L_n^+ = \begin{cases} 0, & \Lambda_n^+ \in (-\infty, C_-], \\ i, & \Lambda_n^+ = C_- + ih, \quad i = \overline{1,m}, \\ m+1, & \Lambda_n^+ \in [C_+ + h, \infty). \end{cases} \qquad (5)$$

Теорема 1. *Если для рассмотренной модели наблюдений выполнены предположения П1 и П2, то векторы вероятностей начальных состояний и матрицы вероятностей переходов цепей Маркова L_n^-, L_n^+ вычисляются поэлементно следующим образом:*

$$\pi_i^\pm = F(\lambda^{-1}(C_- + ih)) - F(\lambda^{-1}(C_- + (i-1)h)),$$
$$\pi_0^\pm = F(\lambda^{-1}(C_-)), \quad \pi_{m+1}^\pm = 1 - F(\lambda^{-1}(C_+));$$
$$p_{i,j}^- = p_{j-i}^- = F(\lambda^{-1}((j-i+1)h)) - F(\lambda^{-1}((j-i)h)),$$
$$p_{i,0}^- = p_{-i}^- = F(\lambda^{-1}((1-i)h)), \quad p_{i,m+1}^- = p_{m+1-i}^- = 1 - F(\lambda^{-1}((m-i+1)h)), \quad i, j = \overline{1,m};$$
$$p_{k,l}^+ = p_{l-k}^+ = p_{l-k-1}^-, \quad k = \overline{1,m}, \quad l = 0,\ 1,\ \ldots,\ m,\ m+1,$$

где $\lambda^{-1}(\cdot)$ – функция, обратная функции $\lambda(\cdot)$, заданной (1).

Доказательство состоит в вычислении вероятностей соответствующих случайных событий с использованием свойства независимости приращений анализируемых случайных последовательностей. ■

Теорема 1 позволяет с использованием соответствия (5) вычислять в явном виде вероятности выхода цепей Маркова Λ_n^-, Λ_n^+ за верхние границы C_+ и $C_+ + h$ соответственно.

Теорема 2. *Для каждого $\omega \in \Omega$ траектории цепей Маркова Λ_n^-, Λ_n и Λ_n^+, $n = 1, 2, \ldots$, удовлетворяют соотношению $\Lambda_n^- \le \Lambda_n \le \Lambda_n^+$.*

Доказательство. Утверждение теоремы следует из свойств целой части действительного числа и способа построения цепей Маркова Λ_n^- и Λ_n^+. ■

В силу теоремы 2 построенные цепи Маркова Λ_n^- и Λ_n^+ будем называть нижней и верхней *граничными цепями* для последовательного критерия (3), (4).

Пусть α_m^-, α_m^+ – вероятности поглощения цепей Маркова L_n^-, L_n^+ в состоянии $(m + 1)$.

Теорема 3. *Для рассмотренной модели наблюдений вероятности α_m^-, α и α_m^+ удовлетворяют соотношению $\alpha_m^- \le \alpha \le \alpha_m^+$.*

Доказательство проводится по схеме, близкой к доказательству теоремы 2 в [5].

Введем блочную структуру матриц вероятностей переходов: матрицы $P^\pm \in \mathrm{R}^{m\times m}$ и $R^\pm \in \mathrm{R}^{m\times 2}$ такие, что $(P^\pm)_{i,j} = p_{ij}^\pm$, $(R^\pm)_{i,1} = p_{i,0}^\pm$, $(R^\pm)_{i,2} = p_{i,m+1}^\pm$, $i, j = \overline{1,m}$. По теореме 1 полученные матрицы $P^\pm$ – теплицевы [7], и для работы с ними (например, для обращения) можно использовать специальные алгоритмы, позволяющие существенно ускорить вычисления.

Пусть $O_{p\times q}(h^k)$ – матрица размерности $p \times q$, каждый элемент которой может быть записан в виде $O(h^k)$, $h \to 0$.

Лемма 1. *Если для рассмотренной модели наблюдений выполнены предположения П1 и П2, то в асимптотике $h \to 0$ справедливы соотношения*

$$P^\pm = O_{m\times m}(h), \quad R^\pm = O_{m\times 2}(1), \quad \pi^\pm = O_{1\times m}(h);$$
$$P^+ - P^- = O_{m\times m}(h^2), \quad R^+ - R^- = O_{m\times 2}(h).$$

Доказательство. Для краткости обозначим $f(y) = f(y;\theta_0)$. По теореме 1

$$p_{ij}^+ = F\left(\lambda^{-1}((j-i)h)\right) - F\left(\lambda^{-1}((j-i-1)h)\right) = \int_{\lambda^{-1}((j-i-1)h)}^{\lambda^{-1}((j-i)h)} f(y)\, dy, \quad i, j = \overline{1,m}.$$

В силу предположений П1, П2 получаем

$$p_{ij}^+ = f\left(\lambda^{-1}((j-i)h) - \frac{A}{2}\right)A + O(A^3), \quad A = \lambda^{-1}((j-i)h) - \lambda^{-1}((j-i-1)h).$$

Учитывая, что $\lambda^{-1}((j-i-1)h)=\lambda^{-1}((j-i)h)-(\lambda^{-1})'((j-i)h)h+O(h^2),$ имеем

$$A=(\lambda^{-1})'((j-i)h)h+O(h^2),\ \ p_{ij}^{+}=f\left(\lambda^{-1}((j-i)h)\right)(\lambda^{-1})'((j-i)h)h+O(h^2).$$

Аналогично $p_{ij}^{-}=f\left(\lambda^{-1}((j-i)h)\right)(\lambda^{-1})'((j-i)h)h+O(h^2).$ Таким образом, $P^{\pm}=O_{m\times m}(h)$ и $P^{+}-P^{-}=O_{m\times m}(h^2).$

Рассмотрим элементы матрицы $R^{+}-R^{-}$:

$$p_{i,0}^{+}-p_{i,0}^{-}=f\left(\frac{\lambda^{-1}((1-i)h)+\lambda^{-1}(-ih)}{2}\right)(\lambda^{-1}((1-i)h)-\lambda^{-1}(-ih))+O(h^3).$$

С учетом $\lambda^{-1}((1-i)h)-\lambda^{-1}(-ih)=O(h)$ получаем $p_{i,0}^{+}-p_{i,0}^{-}=O(h).$ Аналогично $p_{i,m+1}^{+}-p_{i,m+1}^{-}=O(h).$ Следовательно, $R^{+}-R^{-}=O_{m\times 2}(h).$

Соотношение $\pi^{\pm}=O_{1\times m}(h)$ доказывается по такой же схеме. ■

Обозначим через I единичную матрицу размерности m.

Лемма 2. *В условиях леммы 1 имеет место асимптотическое соотношение*

$$(I-P^{+})^{-1}-(I-P^{-})^{-1}=O_{m\times m}(h^2),\ \ h\to 0. \tag{6}$$

Доказательство. Так как все характеристические числа матриц $P^{\pm}$ по модулю меньше единицы [8], справедливо следующее соотношение:

$$(I-P^{\pm})^{-1}=I+O_{m\times m}(h). \tag{7}$$

Докажем равенство

$$(I-P^{+})^{-1}-(I-P^{-})^{-1}=(I-P^{+})^{-1}(P^{+}-P^{-})(I-P^{-})^{-1}. \tag{8}$$

Правую и левую части (8) домножим на матрицу $(I-P^{-})$ справа и на $(I-P^{+})$ – слева. Получим тождество $I-P^{-}-(I-P^{+})=P^{+}-P^{-}$.

Преобразуем (6) с учетом (7), (8) и результата леммы 1:

$$(I-P^{+})^{-1}-(I-P^{-})^{-1}=(I+O_{m\times m}(h))O_{m\times m}(h^2)(I+O_{m\times m}(h))=O_{m\times m}(h^2).\ ■$$

Теорема 4. *Если для рассмотренной модели наблюдений выполнены предположения П1 и П2, то выполняется асимптотическое соотношение:*

$$\alpha_m^{+}-\alpha_m^{-}=O(h),\ \ h\to 0.$$

Доказательство использует теорию поглощающих цепей Маркова [9] и основано на результатах лемм 1, 2 и определениях вероятностей α_m^{-}, α_m^{+}. ■

Таким образом, при $m\to\infty$ ($h\to 0$) вероятность α_m^{+} стремится к фактическому значению вероятности ошибки первого рода α сверху, а α_m^{-} – снизу. В качестве точечного приближения значения α выберем

$$\alpha_m=\left(\alpha_m^{+}+\alpha_m^{-}\right)\big/2. \tag{9}$$

Следствие. При $m\to\infty$ ($h\to 0$) приближение (9) сходится к фактическому значению вероятности ошибки I рода ПКОВ (3), (4) со скоростью $O(h),$ причем абсолютное уклонение от этого значения не превосходит половины длины отрезка $[\alpha_m^{-},\alpha_m^{+}]$:

$$|\alpha-\alpha_m|\le\frac{1}{2}\left(\alpha_m^{+}-\alpha_m^{-}\right). \tag{10}$$

Доказательство. Преобразуем левую часть неравенства (10):

$$|\alpha-\alpha_m|=\left|\alpha-\frac{\alpha_m^{+}+\alpha_m^{-}}{2}\right|\le\frac{1}{2}\left(|\alpha-\alpha_m^{+}|+|\alpha-\alpha_m^{-}|\right).$$

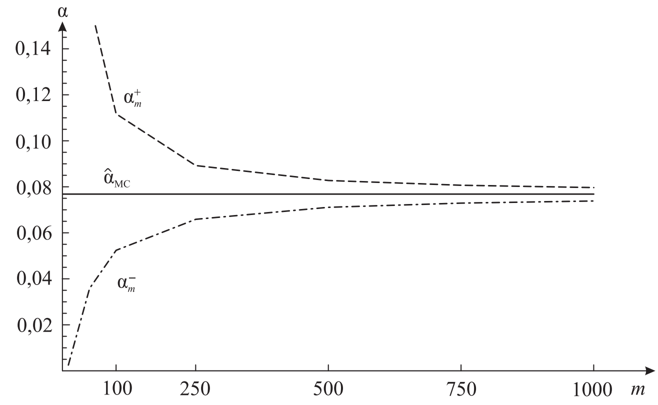
В силу $\alpha_m^{-}\le\alpha\le\alpha_m^{+}$ получим $|\alpha-\alpha_m|\le\frac{1}{2}\left(\alpha_m^{+}-\alpha+\alpha-\alpha_m^{-}\right)=\frac{1}{2}\left(\alpha_m^{+}-\alpha_m^{-}\right)=O(h).$ ■

Таким образом, доказана не только сходимость приближения (9) к истинному значению вероятности ошибки первого рода ПКОВ, но и указана максимальная погрешность такого приближения, а также ее порядок малости.

Результаты вычислительных экспериментов

Вычислительные эксперименты проводились для случая $f(x;\theta) = (1/\sqrt{2\pi})\exp\{-(x-\theta)^2/2\}$, $\alpha_0 = \beta_0 = 0{,}1$, $\Delta = \theta_1 - \theta_0$, $\theta_0 = 0$. Истинной гипотезой во всех экспериментах была гипотеза H_0.

Обозначим через $\widehat{\alpha}_{\text{MC}}$ оценку вероятности ошибки I рода ПКОВ (3), (4), полученную методом Монте-Карло усреднением по 5 000 000 реализаций.



Зависимость от m границ для вероятности ошибки I рода ПКОВ

Результаты вычислительных экспериментов представлены в таблице, в которой при различных значениях Δ показана зависимость от m вероятностей α_m^- и α_m^+, а также величин α_m, $(\alpha_m^+ - \alpha_m^-)/2$. Согласно экспериментам с увеличением m оценка α_m стремится к α – значению вероятности ошибки первого рода для ПКОВ, а также, что α лежит между α_m^- и α_m^+.

Для Δ = 0,5 сходимость величин α_m^+ и α_m^- к $\widehat{\alpha}_{\text{MC}}$ представлена на рисунке.

Зависимость вероятностей ошибок первого рода от *m*

Δ	m	α_m^-	α_m	$\widehat{\alpha}_{\text{MC}}$	α_m^+	$(\alpha_m^+ - \alpha_m^-)/2$
1,5	100	0,04352	0,04505	0,04474	0,04657	0,001525
	500	0,04470	0,04501		0,04531	0,000305
	1000	0,04485	0,04500		0,04515	0,00015
1,0	500	0,05762	0,05869	0,05843	0,05976	0,00107
	1000	0,05814	0,05868		0,05922	0,00054
	2000	0,05841	0,05868		0,05895	0,00027
0,5	1000	0,07384	0,07675	0,07667	0,07966	0,00291
	2000	0,07525	0,07671		0,07817	0,00146
	3000	0,07573	0,07670		0,07767	0,00097
0,3	1000	0,07692	0,08576	0,08535	0,09460	0,00884
	2000	0,08104	0,08545		0,08986	0,00441
	3000	0,08245	0,08539		0,08833	0,00294
0,2	2000	0,08022	0,09053	0,08993	0,10084	0,01031
	3000	0,08336	0,09023		0,09710	0,00687
	4000	0,08498	0,09013		0,09527	0,00515

1. Вальд А. Последовательный анализ. М., 1960.
2. Siegmund D. Sequential analysis. Tests and confidence intervals. New York, 1985.
3. Lai T. // Statistica Sinica. 2001. Vol. 11. P. 303.
4. Kharin A. // Austrian Jornal of Statistics. 2002. Vol. 31 (4). P. 267.
5. Kharin A., Kishylau D. // Austrian Jornal of Statistics. 2005. Vol. 34 (2). P. 153.
6. Харин А.Ю. // Вестн. БГУ. Сер. 1. 2002. № 1. С. 92.
7. Воеводин В.В., Тыртышников Е.Е. Вычислительные процессы с теплицевыми матрицами. М., 1987.
8. Гантмахер Ф.Р. Теория матриц. М., 2004.
9. Кемени Дж., Снелл Дж. Конечные цепи Маркова. М., 1970.

Поступила в редакцию 07.10.10.

Алексей Юрьевич Харин – кандидат физико-математических наук, доцент кафедры теории вероятностей и математической статистики.

Сергей Юрьевич Чернов – аспирант кафедры теории вероятностей и математической статистики. Научный руководитель – А.Ю. Харин.