

# ВЫЧИСЛЕНИЕ ВЕРОЯТНОСТЕЙ ОШИБОК ПОСЛЕДОВАТЕЛЬНОГО ТЕСТА ДЛЯ ВРЕМЕННЫХ РЯДОВ С ТРЕНДОМ

Тон Тхат Ту

Последовательный тест, предложенный А. Вальдом [5], интенсивно используется в приложениях [3]. Для модели одинаково распределенных наблюдений его характеристики достаточно хорошо изучены ([4], [6]). В этой статье исследуется конечность теста и метод вычисления вероятностей ошибок для модели временных рядов с трендом.

**1. Математическая модель.** Рассмотрим гипотетическую вероятностную модель временного ряда с трендом [1]:

$$x_t = \theta^T \psi(t) + \xi_t, \quad t = 1, 2, 3, \dots, \quad (1)$$

где  $\psi(t) = (\psi_1(t), \dots, \psi_m(t))^T, t \geq 1$  – базисные функции тренда,  $\theta = (\theta_1, \dots, \theta_m)^T \in \mathbf{R}^m$  – неизвестный вектор параметров,  $\{\xi_t, t \geq 1\}$  – последовательность независимых одинаково распределённых случайных величин,  $\xi_t \sim N(0, \sigma^2)$ .

Рассматриваются две простые гипотезы ( $\theta^0, \theta^1$  – известные векторы):

$$H_0 : \theta = \theta^0, H_1 : \theta = \theta^1. \quad (2)$$

Обозначим статистику

$$\Lambda_n = \Lambda_n(x_1, x_2, \dots, x_n) = \sum_{t=1}^n \lambda_t, \quad (3)$$

где  $\lambda_t = \ln(p_t(x_t, \theta^1) / p_t(x_t, \theta^0))$  – логарифмическое отношение правдоподобия, вычисленное по наблюдению  $x_t$ , и  $p_t(x, \theta)$  – плотность распределения вероятностей случайной величины  $x_t$ . Из (1) следует, что  $\lambda_t = -\frac{2x_t(\theta^0 - \theta^1)^T \psi(t) + [(\theta^1)^T \psi(t)]^2 - [(\theta^0)^T \psi(t)]^2}{2\sigma^2}$ .

В последовательном тесте Вальда [5] при проверке гипотез (2) после  $n$  наблюдений принимается решение

$$d = \mathbf{1}_{[C_+, +\infty)}(\Lambda_n) + 2 \cdot \mathbf{1}_{(C_-, C_+)}(\Lambda_n), \quad (4)$$

где  $\mathbf{1}_D(\cdot)$  означает индикаторную функцию множества  $D$ . Значения  $d = 0$  и  $d = 1$  соответствуют прекращению процесса наблюдения и принятию гипотезы  $H_0$  или  $H_1$  после  $n$  наблюдений. При  $d = 2$  следует сделать  $(n + 1)$ -е наблюдение. Согласно [5] рекомендуется использовать следующие значения порогов:

$$C_- = \ln(\beta_0 / (1 - \alpha_0)), C_+ = \ln((1 - \beta_0) / \alpha_0), \quad (5)$$

где  $\alpha_0, \beta_0$  – заданные предельно допустимые значения вероятностей ошибок первого и второго рода соответственно.

Примем обозначения:  $E^{(k)}(\cdot), D^{(k)}(\cdot)$  – условные математическое ожидание и дисперсия при условии, что справедлива гипотеза  $H_k, k \in \{0,1\}$ ,

$$H_n = \sum_{t=1}^n \psi(t)\psi^T(t), \quad N = \inf\{n \in \mathbf{N} : \Lambda_n \notin (C_-, C_+)\}, \quad \Gamma = (\theta^1 - \theta^0)(\theta^1 - \theta^0)^T;$$

$$\sigma_n^2 = \frac{[(\theta^0 - \theta^1)^T \psi(n)]^2}{\sigma^2}, \quad \mu_n^{(k)} = \frac{(-1)^{k+1} \sigma_n^2}{2}; \quad s_n^2 = \sum_{i=1}^n \sigma_i^2 = \frac{\text{tr}(\Gamma H_n)}{\sigma^2}, \quad m_n^{(k)} = \frac{(-1)^{k+1} s_n^2}{2}.$$

## 2. ОСНОВНЫЕ РЕЗУЛЬТАТЫ

**Теорема 1.** Если  $\text{tr}(\Gamma H_n) \rightarrow +\infty$  при  $n \rightarrow +\infty$ , то с вероятностью 1 одна из гипотез (2) принимается тестом (3), (4) на основе конечного числа наблюдений.

**Доказательство.** Основано на рассмотрении вероятности  $P_l(N > n), l \in \{0,1\}$ .

Пусть  $\tilde{\Lambda}_n = \sum_{i=1}^n \tilde{\lambda}_i, \quad \tilde{\lambda}_i = f_{C_- - C_+}^{C_+ - C_-}(\lambda_i); \quad f_u^v(x) = u \cdot \mathbf{1}_{(-\infty, u]}(x) + x \cdot \mathbf{1}_{(u, v)}(x) + v \cdot \mathbf{1}_{[v, +\infty)}(x);$   
 $\tilde{N} = \inf\{n \in \mathbf{N} : \tilde{\Lambda}_n \notin (C_-, C_+)\}$ . Видно, что  $\lambda_t = \tilde{\lambda}_t \Leftrightarrow \lambda_t \in [C_- - C_+, C_+ - C_-]$ , и кроме того,  $P_k(\tilde{\lambda}_i > x) = P_k(\lambda_i > x), \forall x \in [C_- - C_+, C_+ - C_-], k \in \{0,1\}$ .

**Лемма.** В указанных выше обозначениях справедливы представления:

$$(i) \quad P_k(N = i) = P_k(\tilde{N} = i), \forall i \geq 1, k \in \{0,1\},$$

$$(ii) \quad P_0(\tilde{\Lambda}_{\tilde{N}} \geq C_+) = P_0(\Lambda_N \geq C_+), \quad P_1(\tilde{\Lambda}_{\tilde{N}} \leq C_-) = P_1(\Lambda_N \leq C_-).$$

**Доказательство.** Основано на тождественных преобразованиях случайных событий, связанных с  $\tilde{N}, \tilde{\Lambda}_{\tilde{N}}$ .

Рассмотрим случай, когда гипотеза  $H_0$  справедлива и

$$(\theta^0 - \theta^1)^T \psi(t) = a, \forall t \geq 1. \quad (6)$$

Тогда  $\{\lambda_t, t \geq 1\}$  – последовательность независимых одинаково и нормально распределённых случайных величин  $N(\mu, \sigma_0^2)$ , где

$$\mu = -a^2 / (2\sigma^2), \quad \sigma_0^2 = a^2 / \sigma^2. \quad (7)$$

Пусть задано  $K \in \mathbf{N}, h = (C_+ - C_-) / K; \quad \bar{L}_n = \sum_{i=1}^n \bar{z}_i$  и  $\underline{L}_n = \sum_{i=1}^n \underline{z}_i$ , где

$$\underline{z}_n = j, \text{ если } jh \leq \tilde{\lambda}_n < (j+1)h, j \in \{-K, K+1, \dots, K-1, K\},$$

$\bar{z}_n = j$ , если  $(j-1)h < \tilde{\lambda}_n \leq jh, j \in \{-K, -K+1, \dots, K-1, K\}$ .

Распределение случайной величины  $\underline{z}_n$  имеет вид:

$$P(\underline{z}_n = k) = \Phi(((k+1)h - \mu) / \sigma_0) - \Phi((kh - \mu) / \sigma_0), k \in \{-K+1, -K+2, \dots, K-2, K-1\},$$

$$P(\underline{z}_n = -K) = \Phi((-K+1)h - \mu) / \sigma_0, P(\underline{z}_n = K) = 1 - \Phi((Kh - \mu) / \sigma_0).$$

Аналогично, распределение случайной величины  $\bar{z}_n$ :

$$P(\bar{z}_n = k) = \Phi((kh - \mu) / \sigma_0) - \Phi(((k-1)h - \mu) / \sigma_0), k \in \{-K+1, -K+2, \dots, K-2, K-1\},$$

$$P(\bar{z}_n = -K) = \Phi((-Kh - \mu) / \sigma_0), P(\bar{z}_n = K) = 1 - \Phi(((K-1)h - \mu) / \sigma_0).$$

Определим случайные последовательности  $\underline{S}_n$  и  $\bar{S}_n$ :

$$\underline{S}_1 = f_{\gamma_1-1}^{\gamma_2+1}(\underline{L}_1), \underline{S}_n = \begin{cases} \gamma_1 - 1, & \underline{S}_{n-1} = \gamma_1 - 1, \\ \gamma_2 + 1, & \underline{S}_{n-1} = \gamma_2 + 1, (n \geq 2); \\ f_{\gamma_1-1}^{\gamma_2+1}(\underline{L}_n), & \text{иначе,} \end{cases}$$

$$\bar{S}_1 = f_{\gamma_1-1}^{\gamma_2+1}(\bar{L}_1), \bar{S}_n = \begin{cases} \gamma_1 - 1, & \bar{S}_{n-1} = \gamma_1 - 1, \\ \gamma_2 + 1, & \bar{S}_{n-1} = \gamma_2 + 1, (n \geq 2), \\ f_{\gamma_1-1}^{\gamma_2+1}(\bar{L}_n), & \text{иначе,} \end{cases}$$

где  $\gamma_1 = [C_- / h] + 1, \gamma_2 = \begin{cases} C_+ / d - 1, & C_+ / d \in \mathbf{Z}, \\ [C_+ / d], & \text{иначе.} \end{cases}$

В силу построения,  $\underline{S}_n$  и  $\bar{S}_n$  – однородные цепи Маркова с пространством состояний  $V = \{\gamma_1 - 1, \gamma_1, \dots, \gamma_2, \gamma_2 + 1\}$ ;  $\gamma_1 - 1, \gamma_2 + 1$  – поглощающие.

Для удобства вычисления перенумеруем пространство состояний  $V$  следующим образом ( $\gamma = \gamma_2 - \gamma_1 + 1$ ):

$$\{\gamma_1 - 1, \gamma_1, \dots, \gamma_2, \gamma_2 + 1\} \rightarrow \{\{\gamma_1 - 1\}, \{\gamma_2 + 1\}, \{\gamma_1\}, \dots, \{\gamma_2\}\} \rightarrow \{\{0\}, \{\gamma + 1\}, \{1\}, \{2\}, \dots, \{\gamma\}\}.$$

Матрицы  $\underline{P}$  и  $\bar{P}$  одношаговых переходных вероятностей имеют вид:

$$\underline{P} = \begin{pmatrix} \mathbf{I}_2 & \mathbf{O}_{2 \times 2} \\ \underline{R} & \underline{Q} \end{pmatrix}, \bar{P} = \begin{pmatrix} \mathbf{I}_2 & \mathbf{O}_{2 \times 2} \\ \bar{R} & \bar{Q} \end{pmatrix},$$

где  $\underline{Q}, \bar{Q}$  – квадратные матрицы размера  $\gamma$ ,  $\underline{R}, \bar{R}$  – матрицы размера  $\gamma \times 2$ .

Примем обозначения:  $\underline{\pi} = (\underline{\pi}_0, \underline{\pi}_{\gamma+1}, \underline{\pi}')$ ,  $\bar{\pi} = (\bar{\pi}_0, \bar{\pi}_{\gamma+1}, \bar{\pi}')$  – векторы начальных вероятностей цепей Маркова  $\underline{S}_n, \bar{S}_n$  соответственно, причём  $\underline{\pi}' = (\underline{\pi}_1, \underline{\pi}_2, \dots, \underline{\pi}_\gamma)$ ,  $\bar{\pi}' = (\bar{\pi}_1, \bar{\pi}_2, \dots, \bar{\pi}_\gamma)$ ;  $\underline{R}_2, \bar{R}_2$  – вторые столбцы матриц  $\underline{R}, \bar{R}$  соответственно;  $\underline{\alpha}, \bar{\alpha}$  – вероятности поглощения цепей Маркова  $\underline{S}_n, \bar{S}_n$  в состоянии  $\gamma + 1$ .

**Теорема 2.** При справедливости гипотезы  $H_i, i \in \{0, 1\}$ , тесты на основе  $\underline{L}_n, \bar{L}_n$  завершаются конечно с вероятностью 1.

**Доказательство.** Этот результат получен из леммы Штейна [4].

**Теорема 3.** *Справедливы следующие выражения:*

$$\underline{\alpha} = \underline{\pi}_{\gamma+1} + \underline{\pi}'(\mathbf{I} - \underline{Q})^{-1} \underline{R}_2, \quad \bar{\alpha} = \bar{\pi}_{\gamma+1} + \bar{\pi}'(\mathbf{I} - \bar{Q})^{-1} \bar{R}_2.$$

**Доказательство.** Основано на теории конечных цепей Маркова с поглощающими состояниями (см. [2]).

**Замечание.** В силу  $\underline{S}_n h \leq \tilde{\Lambda}_n \leq \bar{S}_n h$  и Теоремы 6.2 (см. [3]) имеем  $\underline{\alpha} \leq \alpha \leq \bar{\alpha}$ . На практике можно вычислять  $\alpha$  приближённо следующим образом:

$$\alpha \approx \frac{1}{2}(\underline{\alpha} + \bar{\alpha}) = \frac{1}{2}(\underline{\pi}_{\gamma+1} + \bar{\pi}_{\gamma+1} + \underline{\pi}'(\mathbf{I} - \underline{Q})^{-1} \underline{R}_2 + \bar{\pi}'(\mathbf{I} - \bar{Q})^{-1} \bar{R}_2) \quad (8)$$

### 3. РЕЗУЛЬТАТЫ КОМПЬЮТЕРНЫХ ЭКСПЕРИМЕНТОВ

Рассмотрим вероятностную модель (1) и гипотезы (2) при следующих значениях параметров:

$$m = 4, \sigma = 20, \theta^0 = (1, 2, 1, 2)^T, \theta^1 = (1, 1, 1, 1)^T, \psi(t) = (1, t/10, t^2/100, 10 - t/10)^T.$$

Тогда  $a = (\theta^0 - \theta^1)^T \psi(t) = 10, \forall t \geq 1$ . Этот факт означает, что  $AB = CD$  при любых положениях точек В и D (рисунок 1). Обозначим оценку  $\alpha$  методом Монте-Карло через  $\hat{\alpha}$ . Результаты, полученные с помощью Теоремы 3, приведены в таблице 1. Из таблицы 1 видно, что при увлечении значения  $K$ , значения  $\underline{\alpha}$  и  $\bar{\alpha}$  сближаются. В результате  $(\underline{\alpha} + \bar{\alpha})/2$  сходится к соответствующей оценке  $\hat{\alpha}$  по методу Монте-Карло.

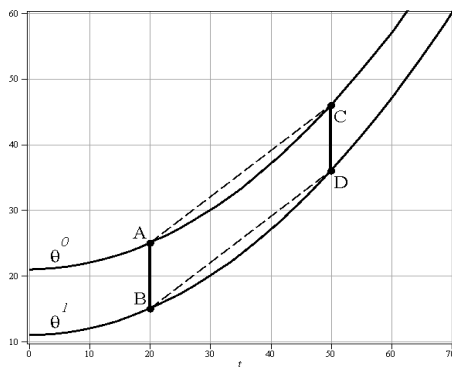


Рисунок 1. Графики функций тренда

Таблица 1

$\alpha_0$	$\beta_0$	$K$	$\underline{\alpha}$	$\bar{\alpha}$	$(\underline{\alpha} + \bar{\alpha})/2$	$\hat{\alpha}$
0.1	0.1	50	0.031	0.161	0.096	0.078
		100	0.049	0.112	0.081	
		500	0.070	0.083	0.077	
0.05	0.1	50	0.010	0.118	0.064	0.040
		100	0.020	0.069	0.045	
		500	0.034	0.043	0.039	
0.01	0.1	50	0.001	0.076	0.039	0.007
		100	0.002	0.025	0.014	
		500	0.006	0.010	0.008	

Ошибки первого рода

### Литература

1. Андерсон, Т. Статистический анализ временных рядов / Т. Андерсон. — М.: Мир, 1976. — 758 с.

2. *Кемени, Дж.* Конечные цепи Маркова. / Дж. Кемени, Дж. Снелл. – М.: Наука, 1976. – 276 с.
3. *Харин, А.Ю.* Робастность байесовских и последовательных статистических решающих правил / А.Ю. Харин. – Минск: БГУ, 2013. – 207 с.
4. *Tartakovsky, A.* Sequential analysis: Hypothesis testing and changepoint detection / A. Tartakovsky, I. Nikiforov, M. Basseville. – New York: CRC Press, 2015. – 574 p.
5. *Wald, A.* Sequential analysis / A. Wald. – New York: John Wiley and Sons, 1947. – 212 p.
6. *Kharin, A.Y.* An approach to performance analysis of the sequential probability ratio test for the simple hypotheses testing / A.Y. Kharin // Proceedings of the Belarusian State University. – 2002. – Vol. 1. – P. 92–96.

## **МОДЕЛИРОВАНИЕ ГИДРАВЛИЧЕСКОГО РАЗРЫВА ПЛАСТАНА ОСНОВЕ МЕТОДА ДИСКРЕТНЫХ ЭЛЕМЕНТОВ**

**В. В. Чайко**

### **1. ОПИСАНИЕ ПРОБЛЕМЫ И ЦЕЛЬ РАБОТЫ**

Гидравлический разрыв пласта (ГРП) – процесс нагнетания жидкости в подземный пласт под давлением, достаточно высоким для того, чтобы вызвать разрыв этого пласта. Данный технологический процесс имеет высокую промышленную применимость в сфере добычи полезных ископаемых для интенсификации работы нефтяных и газовых скважин [1].

При проектировании ГРП важную роль занимает прогнозирование геометрии трещины в зависимости от гидромеханических характеристик материалов, темпа нагнетания жидкости, локальных напряжений в пласте. При этом необходимо принимать в расчет целый набор физических явлений: упругую и пластичную деформацию пласта, распространение трещины, динамику жидкости в области трещины, фильтрацию жидкости в пласте, естественную трещиноватость и прочее.

Цель данной работы спроектировать математическую модель, описывающую вышеназванные процессы и реализовать вычислительно эффективную систему для симуляции гидроразрыва.

### **2. ОПИСАНИЕ РАЗРАБОТАННОЙ МОДЕЛИ**

Для обеспечения возможности моделирования всего набора физических процессов, происходящих при гидроразрыве необходима разработка метода, состоящего из нескольких численных схем для решения