Записки научных
семинаров ПОМИ
Том 260, 1999 г.

Н. М. Халфина

ПРИМЕНЕНИЕ МЕТОДА ШОВЕНЭ ДЛЯ ОБНАРУЖЕНИЯ ГРУБЫХ ОШИБОК В НАБЛЮДЕНИЯХ, СВЯЗАННЫХ В ОДНОРОДНУЮ ЦЕПЬ МАРКОВА

Пусть дана последовательность случайных величин

$$\dots, X_{-2}, X_{-1}, X_0, X_1, \dots, X_n, \dots,$$

связанных в простую однородную цепь Маркова, причем любая пара соседей X_{m-1}, X_m имеет невырожденное двумерное нормальное распределение с параметрами

$$\begin{aligned} &E(X_{m-1}) = E(X_m) = a, \\ &D(X_{m-1}) = D(X_m) = \sigma^2, \\ &R(X_{m-1}, X_m) = \rho, \quad |\rho| \neq 1. \end{aligned}$$

Пусть над случайными величинами $X_1, \dots, X_{n+1}$ произведено $n+1$ последовательных наблюдений $x_1, \dots, x_{n+1}$, при этом параметры a и ρ – неизвестны, а σ^2 – известен.

В предыдущих работах автора [1, 2] были рассмотрены случаи с неизвестными a и σ при известном ρ, неизвестным ρ при известных a и σ.

Как и в указанных выше работах, будем находить вероятность того, что число выброшенных наблюдений, т.е. наблюдений, содержащих грубые ошибки, будет равно m, когда $n \to \infty$. При этом выбрасывание наблюдений будет происходить по методу Шовенэ. Поясним подробнее использование метода Шовенэ в данной задаче, когда наблюдения связаны в цепь Маркова, в то время как сам метод Шовенэ был введен для случая независимых наблюдений.

Определим плотность условного распределения X_i при заданном X_{i-1}

$$f\left(x_i \mid X_{i-1=x_{i-1}}\right) \in N\left[a + \rho(x_{i-1} - a), \sigma\sqrt{1-\rho^2}\right], \quad i = 2, \dots, n+1.$$

Зафиксируем число $\alpha > 0$ и обозначим k_i решение уравнения

$$\int_{k_i}^{\infty} f\left(x_i \mid X_{i-1=x_{i-1}}\right) dx_i = \frac{\alpha}{n}, \quad i = 2, \ldots, n+1.$$

При заданном $\alpha > 0$ будем отбрасывать, т.е. считать содержащими грубые ошибки, те наблюдения x_i, $i \geqslant 2$, которые превышают k_i с вероятностью α/n при условии, что на предыдущем шаге было x_{i-1}.

Итак, для каждого x_i находится своя граница k_i и, если $x_i > k_i$, то x_i – отбрасывается.

Определим случайные величины N_i равенствами

$$N_i = \begin{cases} 1, & \text{если} \quad x_i > k_i, \\ 0, & \text{если} \quad x_i \leqslant k_i, \end{cases}$$

и общее число выбросов $N = \sum_{i=2}^{n+1} N_i$.

Найдем $P\{N = m\}$, вероятность того, что число выброшенных таким образом наблюдений будет равно m, когда $n \to \infty$. Для этого воспользуемся свойствами случайной величины

$$Z_i = \frac{(X_i - a) - \rho(X_{i-1} - a)}{\sigma\sqrt{1-\rho^2}}, \quad i = 2, \ldots, n+1.$$

Очевидно, что Z_i независимы и нормально распределены с параметрами 0 и 1, и Z_i не зависит от X_{i-1}. Обозначим t решение уравнения

$$1 - \Phi(z) = \frac{\alpha}{n}. \tag{1}$$

Тогда

$$P(Z_i > t) = P\left(X_{i>k_i} \mid X_{i-1=x_{i-1}}\right),$$

где

$$k_i = a + \rho(x_{i-1-a}) + t\sigma\sqrt{1-\rho^2}, \quad i = 2, \ldots, n+1 \tag{2}$$

и

$$N_i = \begin{cases} 1, & \text{если} \quad Z_i > t, \quad \text{т.е.} \quad x_i > k_i, \\ 0, & \text{если} \quad Z_i \leqslant t, \quad \text{т.е.} \quad x_i \leqslant k_i, \end{cases}$$

а число выбросов N можно определить, используя величины z_i и границу t. При выборе k_i из равенства (2) находим, что

$$P\left(X_i > k_i \mid X_{i-1} = x_{i-1}\right) = P\left(X_i > k_i\right),$$

т.е. безусловной вероятности.

Так как Z_i независимы, то

$$P\{N=m\} = C_n^m \left(\frac{\alpha}{n}\right)^m \left(1-\frac{\alpha}{n}\right)^{n-m} \simeq \frac{\alpha^m}{m!} e^{-\alpha} + O\left(\frac{\alpha^2}{n}\right). \qquad (3)$$

Итак, в случае всех известных параметров вероятность того, что число выбросов равно m, определяется формулой (3), при этом граница для определения выбросов пересчитывается на каждом шагу по формуле (2). Если же параметры, один или несколько, неизвестны, то заменяем неизвестный параметр его оценкой максимального правдоподобия или достаточной статистикой, строим новую границу $\widetilde{k}_i$ для выбросов, заменяя в формуле (2) неизвестный параметр его оценкой, и, соответственно, считаем

$$\widetilde{N}_i = \begin{cases} 1, & \text{если} \quad x_i > \widetilde{k}_i, \\ 0, & \text{если} \quad x_i \leqslant \widetilde{k}_i, \end{cases}$$

и $\widetilde{N} = \sum_{i=2}^{n+1}$ – общее число выбросов при такой замене. Будем исследовать асимптотическое поведение при $n \to \infty$ распределения случайной величины $\widetilde{N}$, т.е. находить $P\{\widetilde{N}=m\}$ при $n \to \infty$.

В качестве оценок для ρ и a возьмем их оценки максимального правдоподобия $\widehat{\rho}$ и $\widehat{a}$. Для их нахождения выпишем функцию правдоподобия для реализации $x_1, \ldots, x_n$

$$f(x_1, \ldots, x_n) = (2\pi)^{-n/2} \sigma^n (1-\rho^2)^{-\frac{n-1}{2}} \times$$

$$\times \exp\left\{ -\frac{1}{2\sigma^2(1-\rho^2)})(\mathbf{x}-a)^T B_n^{-1} (\mathbf{x}-a) \right\}$$

где $\mathbf{x}^T = (x_1, \ldots, x_n)$, а

$$B_n^{-1} = \begin{pmatrix} 1 & -\rho & 0 & 0 & \ldots & 0 & 0 \\ -\rho & 1+\rho^2 & -\rho & 0 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & 0 & \ldots & 1+\rho^2 & -\rho \\ 0 & 0 & 0 & 0 & \ldots & -\rho & 1 \end{pmatrix}.$$

Введем следующие обозначения

$$R_n^{(1)} = \frac{\partial \ln f}{\partial a} \quad \text{и} \quad R_n^{(2)} = \frac{\partial \ln f}{\partial \rho} \quad (\text{см. } [3]).$$

Решение системы уравнений

$$\begin{cases} R_n^{(1)} = 0 \\ R_n^{(2)} = 0 \end{cases}$$

дает оценки максимального правдоподобия $\widehat{a}$ и $\widehat{\rho}$. С помощью этих оценок строим границу $\widetilde{k}_i$ для выбросов

$$\widetilde{k}_i = t\sigma\sqrt{1-\widehat{\rho}^2} + \widehat{a} + \widehat{\rho}(x_{i-1} - \widehat{a}), \quad i = 2, \ldots .$$

Определим

$$Y_i = \frac{(X_i - \widehat{a}) - \widehat{\rho}(X_{i-1} - \widehat{a})}{\sigma\sqrt{1-\widehat{\rho}^2}}, \quad i = 2, \ldots .$$

Напомним, что $\widetilde{N}_i = 1$, когда $x_i > \widetilde{k}_i$, т.е. когда $Y_i > t$. Так как Y_i зависимы, то найти непосредственно $\{\widetilde{N} = m\}$ нельзя. Как и в предыдущих работах [1, 2], будем оценивать $P\{\widetilde{N} = m\}$ с помощью очевидного неравенства

$$\Big|P\{N = m\} - P\{\widetilde{N} = m\}\Big| \leqslant P\{N \neq \widetilde{N}\} \leqslant \sum_{i=2}^{n+1} P\{N_i \neq \widetilde{N}_i\}. \qquad (4)$$

Запишем Y_i в виде

$$Y_i = Z_i + Z_i\left(\frac{\sqrt{1-\rho^2}}{1-\widehat{\rho}^2} - 1\right) + \frac{a - \widehat{a}}{\sigma\sqrt{1-\widehat{\rho}^2}} +$$

$$+ \frac{(X_{i-1} - a)(\rho - \widehat{\rho})}{\sigma\sqrt{1-\widehat{\rho}^2}} - \frac{\widehat{\rho}(a - \widehat{a})}{\sigma\sqrt{1-\widehat{\rho}^2}}, \quad i = 2, \ldots, n+1$$

или $Y_i = Z_i + R_i$, где

$$R_i = Z_i\left(\frac{\sqrt{1-\rho^2}}{1-\widehat{\rho}^2} - 1\right) + \frac{a - \widehat{a}}{\sigma\sqrt{1-\widehat{\rho}^2}} +$$

$$+ \frac{(X_{i-1} - a)(\rho - \widehat{\rho})}{\sigma\sqrt{1-\widehat{\rho}^2}} - \frac{\widehat{\rho}(a - \widehat{a})}{\sigma\sqrt{1-\widehat{\rho}^2}}, \quad i = 2, \ldots, n+1.$$

Напомним, что $N_i = 1$, когда $Z_i > t$, а $\widetilde{N}_i = 1$, когда $Y_i > t$, где t – решение уравнения (1). Тогда для любого $h > 0$ выполняется неравенство

$$P\{N_i \neq \widetilde{N}_i\} \leqslant P\{|Z_i - t| \leqslant h\} + P\{|R_i| > h\}. \qquad (5)$$

Из (1) видно, что $t = O(\sqrt{\ln n/\alpha})$, поэтому, если выбрать h так, чтобы $h \underset{n\to\infty}{\longrightarrow} 0$, то получим

$$P\{|Z_i - t| \leqslant h\} = \Phi(t+h) - \Phi(t-h) = O(h\Phi'(t)) = O\left(\frac{h\sqrt{\ln n}}{n}\right). \quad (6)$$

Теперь оценим $P\{|R_i| > h\}$. Для первого слагаемого в R_i имеем

$$P\left\{\left|Z_i\left(\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^{\,2}}} - 1\right)\right| > \frac{h}{4}\right\} \leqslant$$

$$\leqslant P\{|Z_i| > H\} + P\left\{\left|\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^{\,2}}} - 1\right| > \frac{h}{4H}\right\}$$

для любого $H > 0$.

Так как $Z_i \in N(0,1)$, то

$$P\{|Z_i| > H\} = O\left(\frac{1}{H}e^{-H^2/2}\right) \quad (7)$$

при $H \to \infty$.

Далее,

$$P\left\{\left|\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^{\,2}}} - 1\right| > \frac{h}{4H}\right\} \leqslant$$

$$\leqslant P\left\{\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^{\,2}}} > 1 + \frac{h}{4H}\right\} + P\left\{\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^{\,2}}} < 1 - \frac{h}{4H}\right\}. \quad (8)$$

Обозначим $\frac{h}{4H} = \varepsilon \simeq \frac{h}{H} \underset{n\to\infty}{\longrightarrow} 0$ и оценим первое слагаемое в правой части (8)

$$P\left\{\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^{\,2}}} > 1+\varepsilon\right\} = P\left\{\frac{\sqrt{1-\widehat{\rho}^{\,2}}}{\sqrt{1-\rho^2}} < \frac{1}{1+\varepsilon}\right\} =$$

$$= P\left\{1 - \frac{\sqrt{1-\widehat{\rho}^{\,2}}}{\sqrt{1-\rho^2}} > 1 - \frac{1}{1+\varepsilon}\right\} = P\left\{\frac{\sqrt{1-\rho^2} - \sqrt{1-\widehat{\rho}^{\,2}}}{\sqrt{1-\rho^2}} > \frac{\varepsilon}{1+\varepsilon}\right\} <$$

$$< P\left\{\frac{|\widehat{\rho}^{\,2} - \rho^2|}{\sqrt{1-\rho^2}(\sqrt{1-\rho^2} + \sqrt{1-\widehat{\rho}^{\,2}})} > \frac{\varepsilon}{1+\varepsilon}\right\} <$$

$$< P\left\{\frac{|\widehat{\rho} - \rho|}{1-\rho^2} > \frac{\varepsilon}{2(1+\varepsilon)}\right\} = P\left\{\frac{|\widehat{\rho} - \rho|}{1-\rho^2} > \varepsilon_1\right\},$$

где $\varepsilon_1 = \frac{\varepsilon}{2(1+\varepsilon)} \simeq \frac{h}{H} \underset{n\to\infty}{\longrightarrow} 0$.

В главе 3 книги [3] показано, что оценки максимального правдоподобия $\widehat{\rho}$ и $\widehat{a}$ при известном σ^2 равномерно в интервале $-\infty < a < \infty$, $-1 < \rho < 1$ имеют асимптотически нормальное распределение, т.е.

$$\frac{(\widehat{\rho}-\rho)\sqrt{(n-1)(1+\rho^2)}}{1-\rho^2} \in N(0,1) \quad \text{при} \quad n \to \infty,$$

$$\frac{(a-\widehat{a})\sqrt{2+(n-2)(1-\rho)}}{\sigma\sqrt{1+\rho}} \in N(0,1) \quad \text{при} \quad n \to \infty.$$

Тогда

$$P\left\{\frac{|\widehat{\rho}-\rho|}{1-\rho^2} > \varepsilon_1\right\} = O\left(e^{-\varepsilon_1^2(n-1)(1+\rho^2)/2}\right) \tag{9}$$

при условии $\varepsilon_1\sqrt{n-1} \underset{n\to\infty}{\longrightarrow} \infty$, т.е. $\frac{h\sqrt{n}}{H} \underset{n\to\infty}{\longrightarrow} \infty$.

Аналогичную оценку получаем и для второго слагаемого в правой части (8). В результате имеем

$$P\left\{\left|\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^2}} - 1\right| > \frac{h}{4H}\right\} = O\left(e^{-\varepsilon_1^2(n-1)}\right) \tag{10}$$

при условии

$$\frac{h\sqrt{n}}{H} \underset{n\to\infty}{\longrightarrow} \infty. \tag{11}$$

Таким образом, для I слагаемого в R_i из (7) и (10) имеем оценку

$$P\left\{\left|Z_i\left(\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^2}} - 1\right)\right| > \frac{h}{4}\right\} \leqslant O\left(\frac{1}{H}e^{-\frac{H^2}{2}} + e^{-\varepsilon_1^2(n-1)}\right), \tag{12}$$

причем

$$H \underset{n\to\infty}{\longrightarrow} \infty, \quad h \underset{n\to\infty}{\longrightarrow} 0, \quad \frac{h\sqrt{n}}{H} \underset{n\to\infty}{\longrightarrow} \infty. \tag{13}$$

Теперь оценим второе слагаемое в R_i

$$P\left\{\left|\frac{a-\widehat{a}}{\sigma\sqrt{1-\widehat{\rho}^2}}\right| > \frac{h}{4}\right\} \leqslant$$

$$\leqslant P\left\{\left|\frac{a-\widehat{a}}{\sigma\sqrt{1-\rho^2}}\right| > \frac{h}{4\sqrt{2}}\right\} + P\left\{\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^2}} > \sqrt{2}\right\}.$$

Так как $\frac{(a-\widehat{a})\sqrt{2+(n-2)(1-\rho)}}{\sigma\sqrt{1+\rho}}$ имеет асимптотически нормальное распределение (см. [3], глава 3), то

$$P\left\{\frac{a-\widehat{a}}{\sigma\sqrt{1-\rho^2}} > \frac{h}{4\sqrt{2}}\right\} =$$

$$= P\left\{\frac{|a-\widehat{a}|\sqrt{2+(n-2)(1-\rho)}}{\sigma\sqrt{1-\rho^2}\sqrt{1+\rho}} > \frac{h}{4\sqrt{2}}\frac{\sqrt{2+(n-2)(1-\rho)}}{\sqrt{1+\rho}}\right\} =$$

$$= O\left(e^{-\frac{h^2[2+(n-2)(1-\rho)](1-\rho^2)}{64(1+\rho)}}\right) = O\left(e^{-\frac{h^2 n(1-\rho)^2}{64}}\right). \qquad (14)$$

При этом должно выполняться условие

$$h\sqrt{n}(1-\rho) \underset{n\to\infty}{\longrightarrow} \infty,$$

из которого следует, что ρ не должно быть очень близко к 1.

Теперь будем оценивать слагаемое $P\left\{\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^{\,2}}} > \sqrt{2}\right\}$.

$$P\left\{\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^{\,2}}} > \sqrt{2}\right\} = P\left\{\frac{\sqrt{1-\widehat{\rho}^{\,2}}}{\sqrt{1-\rho^2}} < \frac{1}{\sqrt{2}}\right\} =$$

$$= P\left\{1-\frac{\sqrt{1-\widehat{\rho}^{\,2}}}{\sqrt{1-\rho^2}} > 1-\frac{1}{\sqrt{2}}\right\} = P\left\{\frac{\sqrt{1-\rho^2}-\sqrt{1-\widehat{\rho}^{\,2}}}{\sqrt{1-\rho^2}} > \frac{\sqrt{2}-1}{\sqrt{2}}\right\} <$$

$$< P\left\{\frac{|\widehat{\rho}^{\,2}-\rho^2|}{\sqrt{1-\rho^2}(\sqrt{1-\rho^2}+\sqrt{1-\widehat{\rho}^{\,2}})} > \frac{0,4}{\sqrt{2}}\right\} <$$

$$< P\left\{\frac{|\widehat{\rho}-\rho}{1-\rho^2} > \frac{0,4}{2\sqrt{2}}\right\},$$

и так как

$$(\widehat{\rho}-\rho)\sqrt{(n-1)(1+\rho^2)} \in N(0,1) \quad \text{при} \quad n\to\infty,$$

то

$$P\left\{\frac{|\widehat{\rho}-\rho|}{1-\rho^2} > \frac{0,4}{2\sqrt{2}}\right\} = O\left(e^{-\frac{(n-1)(1+\rho^2)0,16}{16}}\right) = O\left(e^{-\frac{n-1}{100}}\right). \qquad (15)$$

Итак, из (14) и (15) имеем оценку второго слагаемого в R_i:

$$P\left\{\left|\frac{a-\widehat{a}}{\sigma\sqrt{1-\widehat{\rho}^{\,2}}}\right| > \frac{h}{4}\right\} \leqslant O\left(e^{-\frac{h^2 n(1-\rho)^2}{64}} + e^{-0,01(n-1)}\right). \qquad (16)$$

При этом должно выполняться условие

$$h\sqrt{n}(1-\rho) \underset{n\to\infty}{\longrightarrow} \infty. \tag{17}$$

Напомним, что для оценки (12) первого слагаемого в R_i должны выполняться условия (13).

Теперь займемся оценкой третьего члена в R_i.

$$P\left\{\left|\frac{(X_{i-1}-a)(\rho-\widehat{\rho})}{\sigma\sqrt{1-\widehat{\rho}^{\,2}}}\right| > \frac{h}{4}\right\} <$$

$$< P\left\{\left|\frac{X_{i-1}-a}{\sigma}\right| > H\right\} + P\left\{\left|\frac{\rho-\widehat{\rho}}{\sqrt{1-\widehat{\rho}^{\,2}}}\right| > \frac{h}{4H}\right\} =$$

$$= O\left(\frac{1}{H}e^{-\frac{H^2}{2}}\right) + P\left(\left|\frac{\rho-\widehat{\rho}}{\sqrt{1-\widehat{\rho}^{\,2}}}\right| > \frac{h}{4H}\right).$$

Далее,

$$P\left\{\left|\frac{\rho-\widehat{\rho}}{\sqrt{1-\widehat{\rho}^{\,2}}}\right| > \frac{h}{4H}\right\} <$$

$$< P\left\{\left|\frac{\rho-\widehat{\rho}}{1-\widehat{\rho}^{\,2}}\right| > \frac{h}{4H\sqrt{2}}\right\} + P\left\{\frac{1-\rho^2}{\sqrt{1-\widehat{\rho}^{\,2}}} > \sqrt{2}\right\}.$$

В свою очередь,

$$P\left\{\frac{1-\rho^2}{\sqrt{1-\widehat{\rho}^{\,2}}} > \sqrt{2}\right\} < P\left\{\frac{\sqrt{1-\rho^2}}{\sqrt{1-\widehat{\rho}^{\,2}}} > \sqrt{2}\right\} =$$

$$= P\left\{\frac{\sqrt{1-\widehat{\rho}^{\,2}}}{\sqrt{1-\rho^2}} < \frac{1}{\sqrt{2}}\right\} = P\left\{1-\frac{\sqrt{1-\widehat{\rho}^{\,2}}}{\sqrt{1-\rho^2}} > 1-\frac{1}{\sqrt{2}}\right\} =$$

$$= P\left\{\frac{\sqrt{1-\rho^2}-\sqrt{1-\widehat{\rho}^{\,2}}}{\sqrt{1-\rho^2}} > \frac{\sqrt{2}-1}{\sqrt{2}}\right\} <$$

$$< P\left\{\frac{|\widehat{\rho}^{\,2}-\rho^2|}{\sqrt{1-\rho^2}(\sqrt{1-\rho^2}+\sqrt{1-\widehat{\rho}^{\,2}})} > \frac{\sqrt{2}-1}{\sqrt{2}}\right\} <$$

$$P\left\{\frac{|\widehat{\rho}-\rho|}{1-\rho^2} > \frac{\sqrt{2}-1}{2\sqrt{2}}\right\} =$$

$$= O\left(e^{-\frac{0{,}16(n-1)(1+\rho^2)}{16}}\right) = O\left(e^{-\frac{n-1}{100}}\right).$$

И

$$P\left\{\left|\frac{\rho-\widehat{\rho}}{1-\rho^2}\right|>\frac{h}{4H\sqrt{2}}\right\}=O\left(e^{-\frac{h^2(n-1)(1+\rho^2)}{64H^2}}\right)=O\left(e^{-\frac{h^2(n-1)}{64H^2}}\right),$$

так как

$$\frac{(\widehat{\rho}-\rho)\sqrt{(n-1)(1+\rho^2)}}{1-\rho^2}\in N(0,1)\quad\text{при}\quad n\to\infty.$$

Возникающее при этом условие $\frac{h\sqrt{n}}{H}\underset{n\to\infty}{\longrightarrow}\infty$ было уже учтено в (13).

Итак, получена оценка для 3-го члена в R_i

$$P\left\{\left|\frac{(X_{i-1}-a)(\rho-\widehat{\rho})}{\sigma\sqrt{1-\widehat{\rho}^{\,2}}}\right|>\frac{h}{4}\right\}<$$

$$<O\left(\frac{1}{H}e^{-\frac{H^2}{2}}+e^{-\frac{n-1}{100}}+e^{-\frac{h^2(n-1)}{64H^2}}\right).\tag{18}$$

Осталось оценить последний, четвертый член в R_i.

$$P\left\{\left|\frac{(a-\widehat{a})\widehat{\rho}}{\sigma\sqrt{1-\widehat{\rho}^{\,2}}}\right|>\frac{h}{4}\right\}<P\left\{\left|\frac{\widehat{\rho}}{\rho}\right|>\sqrt{2}\right\}+P\left\{\left|\frac{(a-\widehat{a})\rho}{\sigma\sqrt{1-\widehat{\rho}^{\,2}}}\right|>\frac{h}{4\sqrt{2}}\right\}.$$

Учитывая (16), где h заменено на $\frac{h}{\sqrt{2}\rho}$, получим

$$P\left\{\left|\frac{(a-\widehat{a})\rho}{\sigma\sqrt{1-\widehat{\rho}^{\,2}}}\right|>\frac{h}{4\sqrt{2}}\right\}\leqslant P\left\{\left|\frac{a-\widehat{a}}{\sigma\sqrt{1-\widehat{\rho}^{\,2}}}\right|>\frac{h}{4\sqrt{2}\rho}\right\}\leqslant$$

$$\leqslant O\left(e^{-\frac{h^2n(1-\rho)^2}{128\rho^2}}+e^{-\frac{n-1}{100}}\right)\tag{19}$$

при условии

$$\frac{h\sqrt{n}(1-\rho)}{\rho}\underset{n\to\infty}{\longrightarrow}\infty.\tag{20}$$

Наконец,

$$P\left\{\left|\frac{\widehat{\rho}}{\rho}\right|>\sqrt{2}\right\}\leqslant P\{\widehat{\rho}^{\,2}>2\rho^2\}=P\{\widehat{\rho}^{\,2}-\rho^2>\rho^2\}=$$

$$=P\{(\widehat{\rho}-\rho)(\widehat{\rho}+\rho)>\rho^2\}\leqslant P\{|\widehat{\rho}-\rho|>\rho^2/2\}=$$

$$=P\left\{\frac{|\widehat{\rho}-\rho|\sqrt{(n-1)(1+\rho^2)}}{1-\rho^2}>\frac{\rho^2\sqrt{(n-1)(1+\rho^2)}}{2(1-\rho^2)}\right\}=$$

$$= O\left(e^{-\frac{\rho^4(n-1)(1+\rho^2)}{8(1-\rho^2)^2}}\right). \tag{21}$$

При этом должно выполняться условие $\rho\sqrt[4]{n} \underset{n\to\infty}{\longrightarrow} \infty$, т.е. ρ не должно быть очень близко к 0.

Итак, суммируя (19) и (21), получаем оценку последнего члена в R_i

$$P\left\{\left|\frac{(a-\widehat{a})\widehat{\rho}}{\sigma\sqrt{1-\widehat{\rho}^2}}\right| > \frac{h}{4}\right\} \leqslant$$

$$\leqslant O\left(e^{-\frac{h^2n(1-\rho)^2}{128\rho^2}} + e^{-\frac{n-1}{100}} + e^{-\frac{\rho^4(n-1)(1+\rho^2)}{8(1-\rho^2)^2}}\right) \tag{22}$$

при условии

$$\rho\sqrt[4]{n} \underset{n\to\infty}{\longrightarrow} \infty. \tag{23}$$

Из (4), (5), (6) для любого $h > 0$ получаем

$$\left|P\{N = m\} - P\{\widetilde{N} = m\}\right| \leqslant O(h\sqrt{\ln n}) + nP\{|R_i| > h\},$$

и, используя (12), (16), (18) и (22) для оценки $P\{|R_i| > h\}$, получим

$$\left|P\{N = m\} - P\{\widetilde{N} = m\}\right| \leqslant$$

$$\leqslant O\left(h\sqrt{\ln n}\right) + nO\left(\frac{1}{H}e^{-H^2} + e^{-\varepsilon_1^2(n-1)} + e^{-\frac{h^2n(1-\rho)^2}{64}} + e^{-\frac{(n-1)}{100}} + \right.$$

$$\left. + e^{-\frac{h^2(n-1)}{64H^2}} + e^{-\frac{h^2n(1-\rho)^2}{128\rho^2}} + e^{-\frac{\rho^4(n-1)(1+\rho^2)}{8(1-\rho^2)^2}}\right). \tag{24}$$

При этом должны выполняться следующие условия при соответствующем выборе h, H (см. (13), (17), (20) и (23)):

$$h \underset{n\to\infty}{\longrightarrow} 0, \quad H \underset{n\to\infty}{\longrightarrow} \infty, \quad \frac{h}{H} \underset{n\to\infty}{\longrightarrow} 0, \quad \frac{h\sqrt{n}}{H} \underset{n\to\infty}{\longrightarrow} 0,$$

$$h\sqrt{n}(1-\rho) \underset{n\to\infty}{\longrightarrow} \infty, \quad \frac{h\sqrt{n}(1-\rho)}{\rho} \underset{n\to\infty}{\longrightarrow} \infty, \quad \rho\sqrt[4]{n} \underset{n\to\infty}{\longrightarrow} \infty. \tag{25}$$

Выберем $h = n^{-\delta}$, $H = n^l$, где $0 < \delta < \frac{1}{2}$, $0 < l < \frac{1}{2}$, $\delta + l < \frac{1}{2}$. При таком выборе h и H первые четыре условия выполняются автоматически, а из последних трех условий вытекает некоторое ограничение на поведение ρ у концов отрезка $[0, 1]$.

При выбранном $h = n^{-\delta}$, $0 < \delta < \frac{1}{2}$ и $\rho \to 0$ условия $h\sqrt{n}(1-\rho) \underset{n\to\infty}{\longrightarrow} \infty$ и $\frac{h\sqrt{n}(1-\rho)}{\rho} \underset{n\to\infty}{\longrightarrow} \infty$ выполняются, так как $n^{-\delta+1/2} \underset{n\to\infty}{\longrightarrow} \infty$. А при $h = n^{-\delta}$, $0 < \delta < \frac{1}{2}$ и $\rho \to 1$ эти условия выполняются, если ограничить стремление ρ к 1, т.е. положить

$$1 - \rho = n^{-T}, \quad \text{где} \quad T < \frac{1}{2} - \delta. \tag{26}$$

Тогда $n^{-\delta+1/2-T} \underset{n\to\infty}{\longrightarrow} \infty$, т.е. условия выполняются.

Из последнего условия $\rho\sqrt[4]{n} \underset{n\to\infty}{\longrightarrow} \infty$ следует ограничение на стремление ρ к 0, т.е. ρ должно стремиться к 0 медленнее, чем $n^{-1/4}$.

Положим

$$\rho = n^{-k}, \quad \text{где} \quad k < \frac{1}{4}. \tag{27}$$

Тогда

$$\rho\sqrt[4]{n} = n^{-k+1/4} \underset{n\to\infty}{\longrightarrow} \infty.$$

Итак, все условия (25) выполнены при

$$h = n^{-\delta}, \quad 0 < \delta < \frac{1}{2}; \quad H = n^{l}, \quad 0 < l < \frac{1}{2}, \quad \delta + l < \frac{1}{2}.$$

В результате при выбранных таким образом h и H

$$\left|P\{N = m\} - P\{\widetilde{N} = m\}\right| \leqslant O\left(\frac{\sqrt{\ln n}}{n^{\delta}}\right),$$

где $0 < \delta < \frac{1}{2}$. При этом ρ должно быть отделено от 0 и 1 условиями (26) и (27).

Литература

1. Н. М. Халфина, *Обнаружение грубых ошибок методом Шовенэ в наблюдениях, связанных в однородную цепь Маркова.* — Зап. научн. семин. ЛОМИ **177** (1989), 163–169.
2. Н. М. Халфина, *Обнаружение грубых ошибок методом Шовенэ в наблюдениях, связанных в однородную цепь Маркова.* — Зап. научн. семин. ПОМИ **207** (1993), 137–142.

3. М. Арато, *Линейные стохастические системы с постоянными коэффициентами*. Наука, М. (1989).

Halfina N. M. Application of the Chauvenet test to detection of overliers in observations, connected in homogeneous Markov chain.

Let $x_1, \ldots, x_n$ be random values, connected in homogeneous Markov chain. Asymptotic behavior of the distribution of random value N: number of overliers is investigated for unknown parameters a and ρ.

С.-Петербургское отделение
Математического института
им.В. А. Стеклова РАН

Поступило 12 ноября 1999 г.