

Я. Ю. НИКИТИН

ОЦЕНКИ СКОРОСТИ СХОДИМОСТИ В НЕКОТОРЫХ ПРЕДЕЛЬНЫХ ТЕОРЕМАХ И СТАТИСТИЧЕСКИХ КРИТЕРИЯХ

(Представлено академиком Ю. В. Линником 5 VII 1971)

1. Рассмотрим k независимых выборок с объемами $n_1, n_2, \dots, n_k$, извлеченных из совокупностей с функциями распределения $F_1, F_2, \dots, F_k$. Пусть $S_{n_1}, S_{n_2}, S_{n_k}$ — соответствующие эмпирические функции распределения,

$n = \sum_{j=1}^k n_j$, G — заданная функция распределения. Для проверки гипотезы о том, что $F_1 = F_2 = \dots = F_k = G$ в математической статистике используются критерии, основанные на изучении статистик

$$T_n^2 = \sup_x \sum_{j=1}^k n_j [S_{n_j}(x) - G(x)]^2,$$

$$W_n^2 = \int_{-\infty}^{\infty} \sum_{j=1}^k n_j [S_{n_j}(x) - G(x)]^2 dG(x).$$

Хорошо известно, что какова бы ни была непрерывная функция распределения G , распределения статистик T_n^2 и W_n^2 не зависят от G и сходятся при $n_1, n_2, \dots, n_k \rightarrow \infty$ к некоторым предельным распределениям. Вид этих распределений найден Кифером ⁽¹⁾. Им показано, что

$$\lim_{n_1, n_2, \dots, n_k \rightarrow \infty} P\{T_n^2 < x\} = A_k(x) =$$

$$= \frac{4}{\Gamma(k/2)(2x)^{k/2}} \sum_{s=1}^{\infty} \frac{(\gamma_{(k-2)/2, s})^{k-2} \exp(-(\gamma_{(k-2)/2, s})^2/(2x))}{[J_{k/2}(\gamma_{(k-2)/2, s})]^2},$$

где $\gamma_{v, s}$ — s -й положительный нуль бесселевой функции $J_v(x)$, и что

$$\lim_{n_1, n_2, \dots, n_k \rightarrow \infty} P\{W_n^2 < x\} = B_k(x) =$$

$$= \frac{2^{(k+1)/2}}{\sqrt{\pi} x^{k/4}} \sum_{s=0}^{\infty} \frac{\Gamma(s + k/2)}{s! \Gamma(k/2)} \exp\left(-2\left(s + \frac{k}{4}\right)^2 \frac{1}{x}\right) \cdot D_{(k-2)/2}\left(\frac{2s + k/4}{\sqrt{x}}\right),$$

где $D_v(x)$ — функция параболического цилиндра.

Представляет интерес оценка скорости этой сходимости. В случае одной выборки объема n оценка

$$\sup_x |P\{T_n^2 < x\} - A_1(x)| = O\left(\frac{1}{\sqrt{n}}\right)$$

вытекает из асимптотического разложения для T_n^2 , полученного Чжан Липянем ⁽²⁾. Изучая скорость сходимости распределения статистики W_n^2 к $B_1(x)$, В. В. Сазонов ⁽³⁾ получил последовательно оценки $O(n^{-1/10+\varepsilon})$ и $O(n^{-1/6+\varepsilon})$, а Розенкранц ⁽⁴⁾ оценку $O(\log^{3/2} n/n^{1/5})$ и высказал предположение о возможности ее улучшения до $O(\log^{3/2} n/n^{1/4})$. В рассматриваемом

случае k выборов удалось получить следующие оценки:

$$\sup_x |P\{T_n^2 < x\} - A_k(x)| \leq C_1(k) \cdot \sum_{i=1}^k \log^{3/4} n_i/n_i^{1/4},$$

$$\sup_x |P\{W_n^2 < x\} - B_k(x)| \leq C_2(k) \cdot \sum_{i=1}^k \log^{5/4} n_i/n_i^{1/4},$$

где $C_1(k)$, $C_2(k)$ — некоторые постоянные, зависящие только от k . Применяемые при доказательстве методы сходны с методами ⁽⁴⁾ и основываются на теореме А. В. Скорохода ⁽⁵⁾ о представлении сумм независимых случайных величин через значения процесса броуновского движения в случайные моменты времени.

2. Основываясь на представлении Скорохода, удается получить оценки скорости сходимости в некоторых предельных теоремах для функционалов от сумм независимых случайных величин. Пусть $\xi_1, \xi_2, \dots, \xi_n$ — независимые одинаково распределенные случайные величины с нулевыми средними и единичными дисперсиями, $S_n = \sum_{i=1}^n \xi_i$, $0 < \alpha < 1$, $w(t)$ — процесс

броуновского движения. По принципу инвариантности

$$\lim_{n \rightarrow \infty} P\left\{\min_{[an] \leq k \leq n} S_k > x \sqrt{n}\right\} = P\left\{\min_{\alpha \leq t \leq 1} w(t) > x\right\} =$$

$$= \frac{1}{\pi} \int_{x/\sqrt{\alpha}}^{\infty} \exp\left(-\frac{u^2}{2}\right) du \int_0^{(u\sqrt{\alpha}-x)/\sqrt{1-\alpha}} \exp\left(-\frac{t^2}{2}\right) dt,$$

$$\lim_{n \rightarrow \infty} P\{S_1^2 + S_2^2 + \dots + S_n^2 < xn^2\} = P\left\{\int_0^1 |w(t)|^2 dt < x\right\} =$$

$$= \frac{1}{4\pi} \frac{1}{\sqrt{\pi}} \int_0^{x/2} u^{-1/2} du \int_0^{\pi/2} (\cos t)^{-1/2} \theta_1'\left(\frac{t}{2}, \exp\left(-\frac{1}{4u}\right)\right) dt,$$

где $\theta_1'(z, q) = \frac{\partial}{\partial z} \left[2 \sum_{k=0}^{\infty} (-1)^k q^{(2k+1)^2/4} \sin(2k+1)z \right]$.

Впервые эти теоремы были доказаны соответственно в ⁽⁶⁾ и ⁽⁷⁾.

Предположим теперь, что $E|\xi_1|^{2+\delta} < \infty$ для некоторого $\delta \in (0, \infty)$. Тогда могут быть установлены следующие оценки для

$$\Delta_n = \sup_x |P\left\{\min_{[an] \leq k \leq n} S_k > x \sqrt{n}\right\} - P\left\{\min_{\alpha \leq t \leq 1} w(t) > x\right\}|:$$

при $0 < \delta \leq 2$

$$\Delta_n = O(n^{-\delta/2(\delta+3)} (\log n)^{(\delta+2)/2(\delta+3)});$$

при $2 < \delta \leq 3$

$$\Delta_n = O(n^{-(\delta+2)/4(\delta+3)} (\log n)^{(\delta+2)/2(\delta+3)});$$

при $\delta > 3$

$$\Delta_n = O(n^{-1/4} (\log n)^{3/4}).$$

Постоянные в этих оценках зависят только от α и $E|\xi_1|^{2+\delta}$.

Оценки для $\sup_x |P\{S_1^2 + \dots + S_n^2 < xn^2\} - P\left\{\int_0^1 |w(t)|^2 dt < x\right\}|$ отлича-

ются от них только порядком логарифмического множителя, который следует увеличить везде на $1/2$.

3. Оказалось, что метод А. В. Скорохода применим и к многомерным предельным теоремам при наличии некоторых ограничений на распределе-

ний слагаемых. Здесь мы дадим оценку скорости сходимости в многомерном варианте граничной задачи для сумм независимых случайных векторов. Для простоты рассматривается двумерный случай.

Пусть $X_i = (\xi_i, \eta_i)$, $i = 1, \dots, n$, — последовательность независимых одинаково распределенных случайных векторов со средним 0 и матрицей

ковариаций B , $S_n = \sum_{i=1}^n X_i$, $\xi(t)$ — винеровский процесс, определяемый характеристической функцией

$$E e^{i(z, \xi(t))} = \exp\left(-\frac{t}{2}(Bz, z)\right), \quad z \in R^2,$$

Γ — кусочно-гладкая поверхность в $R^2 \times [0, 1]$, ограничивающая некоторую область $\mathcal{D}$, сечение которой уровнем $t \in [0, 1]$ мы обозначим $\mathcal{D}_t$. Как доказано, например, в (8),

$$\lim_{n \rightarrow \infty} P\left\{\frac{1}{\sqrt{n}} S_k \in \mathcal{D}_{k/n}, k = 1, \dots, n\right\} = P\{\xi(t) \in \mathcal{D}_t, 0 \leq t \leq 1\}.$$

Наложим на поверхность Γ следующие ограничения:

- а) $(0, 0) \in \mathcal{D}_0 \setminus \partial \mathcal{D}_0$;
- б) все сечения $\mathcal{D}_t$ — выпуклые множества;
- в) кривые, образующиеся при пересечении Γ с плоскостями, проходящими через ось t , удовлетворяют условию Липшица с константой K .

От случайных векторов X_i мы потребуем, чтобы регрессия одной компоненты на другую была линейной, т. е. чтобы п.н. $E(\xi_i | \eta_i) = a\eta_i$, где a — некоторая постоянная.

Тогда имеет место следующая

Теорема. Пусть $E|\xi_i|^{2+\delta} = b_1 < \infty$, $E|\eta_i|^{2+\delta} = b_2 < \infty$ для некоторого $\delta \in (0, \infty)$ и выполнены все условия, наложенные нами на Γ и распределение X_i .

Тогда

$$\left| P\left\{\frac{1}{\sqrt{n}} S_k \in \mathcal{D}_{k/n}, k = 1, \dots, n\right\} - P\{\xi(t) \in \mathcal{D}_t, 0 \leq t \leq 1\} \right| \leq C_3 \varphi_n,$$

где C_3 — постоянная, зависящая от $K, \delta, \mathcal{D}_0, b_1$ и b_2 , а порядок убывания по n величины φ_n точно такой же, как и для рассмотренной выше величины Δ_n .

Если же длины случайных векторов X_i ограничены с вероятностью 1 константой C , то

$$\left| P\left\{\frac{1}{\sqrt{n}} S_k \in \mathcal{D}_{k/n}, k = 1, \dots, n\right\} - P\{\xi(t) \in \mathcal{D}_t, 0 \leq t \leq 1\} \right| \leq C_4 \sqrt{\frac{\log n}{n}},$$

где C_4 — постоянная, зависящая от $K, \delta, \mathcal{D}_0, C$.

Аналогичная оценка может быть получена и в случае большего числа измерений. При этом константы C_3 и C_4 будут, разумеется, зависеть и от размерности пространства.

Ленинградский государственный университет
им. А. А. Жданова

Поступило
16 VI 1971

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

- ¹ J. Kiefer, Ann. Math. Stat., 30, № 2, 420 (1959). ² Чжап Ли-цяннь, Математика, 4, № 2, 121 (1960). ³ В. В. Сазонов, Теория вероятн. и ее применения, 14, в. 4, 667 (1969). ⁴ W. Rosenkrantz, Trans. Am. Math. Soc., 139, 329 (1969). ⁵ А. В. Скороход, Исследования по теории случайных процессов, Киев. 1961. ⁶ A. Mark, Bull. Am. Math. Soc., 55, № 10, 885 (1949). ⁷ P. Erdős, M. Kac, Bull. Am. Math. Soc., 52, № 4, 292 (1946). ⁸ А. В. Скороход, Н. П. Слободенюк, Пределные теоремы для случайных блужданий, Киев. 1970.