

Я. Ю. НИКИТИН

О ГРАНИЧНОЙ ЗАДАЧЕ ДЛЯ ЭМПИРИЧЕСКОГО ПРОЦЕССА

(Представлено академиком Ю. В. Линником 27 I 1972)

1. Пусть $\xi_1, \xi_2, \dots, \xi_n$ — выборка объема n из равномерного распределения на $[0, 1]$, $F_n(t)$ — эмпирическая функция распределения, построенная по этой выборке. Процесс $\xi_n(t) = \sqrt{n}(F_n(t) - t)$ будем называть в дальнейшем эмпирическим. Нас будет интересовать асимптотическое поведение при $n \rightarrow \infty$ вероятности

$$Q_n(a, b) \equiv P\{g_1(t) < \xi_n(t) < g_2(t), a \leq t \leq b\},$$

где $0 \leq a < b \leq 1$, а функции $g_i(t)$, $i = 1, 2$, удовлетворяют при $t \geq 0$ условиям

$$g_1(t) < g_2(t), \quad g_1(0) < 0 < g_2(0), \\ |g_i(t+h) - g_i(t)| \leq Kh, \quad h > 0,$$

для некоторой постоянной $K > 0$.

Точная формула для $Q_n(a, b)$ в виде определителя n -го порядка, удобная при малых объемах выборок, получена в (1).

В дальнейшем $w(t)$ — стандартный винеровский процесс, $\bar{w}(t) \equiv w(t) - tw(1)$ — так называемый броуновский мост,

$$\varepsilon_n = K \sqrt{\frac{\ln n}{n}} + O\left(\sqrt[4]{\frac{\ln^3 n}{n}}\right).$$

Здесь и ниже постоянная в символе O будет предполагаться абсолютной, а в символах O_a и $O_{a,b}$ — зависящей лишь от a или от a и b , $0 < a < b < 1$.

Теорема 1. *Справедливы следующие оценки:*

- а) $|Q_n(0, 1) - P\{g_1(t) - \varepsilon_n < \bar{w}(t) < g_2(t) + \varepsilon_n, 0 \leq t \leq 1\}| = O\left(\frac{\ln n}{\sqrt{n}}\right);$
- б) $\left|Q_n(a, 1) - P\left\{g_1(t) - \varepsilon_n < \bar{w}(t) < g_2(t) + \varepsilon_n, a + O_a\left(\frac{\sqrt{\ln n}}{n}\right) \leq t \leq 1\right\}\right| = O_a\left(\frac{\ln n}{\sqrt{n}}\right);$
- в) $\left|Q_n(0, b) - P\left\{g_1(t) - \varepsilon_n < \bar{w}(t) < g_2(t) + \varepsilon_n, 0 \leq t \leq b + O_b\left(\sqrt{\frac{\ln n}{n}}\right)\right\}\right| = O_b\left(\frac{\ln n}{\sqrt{n}}\right);$
- г) $\left|Q_n(a, b) - P\left\{g_1(t) - \varepsilon_n < \bar{w}(t) < g_2(t) + \varepsilon_n, a + O_a\left(\sqrt{\frac{\ln n}{n}}\right) \leq t \leq b + O_b\left(\sqrt{\frac{\ln n}{n}}\right)\right\}\right| = O_{a,b}\left(\frac{\ln n}{\sqrt{n}}\right).$

Рассуждения, используемые при доказательстве, сходны с рассуждениями в (2). От эмпирического процесса делается переход к процессу $S_n(t) - tS_n(1)$, где $S_n(t)$ — случайная ломаная, построенная на отрезке $[0, 1]$ по суммам независимых одинаково распределенных случайных величин, имеющих показательное распределение с параметром 1. Гра-

ничная же задача для процесса $S_n(t) - tS_n(1)$ может быть исследована с помощью метода А. В. Скорохода ⁽³⁾, основанного на случайной замене времени.

Теорема 1 может иметь различные приложения. В приводимых ниже примерах получены оценки скорости сходимости к пределу распределений некоторых непараметрических статистик.

Пример 1 (статистики А. Реньи ⁽⁴⁾). Пусть $g_1(t) = -g_2(t) = -z(1-t)$, $z > 0$; $a = 0$. Тогда

$$Q_n(0, b) = P \left\{ \sup_{0 \leq t \leq b} \frac{|\xi_n(t)|}{1-t} < z \right\};$$

используя оценку в) доказанной теоремы, можно получить

$$\sup_z \left| P \left\{ \sup_{0 \leq t \leq b} \frac{|\xi_n(t)|}{1-t} < z \right\} - P \left\{ \sup_{0 \leq t \leq \frac{1}{1-b}} |w(t)| < z \right\} \right| = O_b \left(\sqrt[4]{\frac{\ln^3 n}{n}} \right).$$

Заметим, что эта оценка улучшает результат А. И. Орлова, анонсировавшего в ⁽⁵⁾ оценку $O_{b,e}(n^{-1/5+\varepsilon})$.

Если положить, например, $g_1(t) = -\infty$, $g_2(t) = z(1-t)$, $z > 0$, то из г) получим соотношение

$$\sup_z \left| P \left\{ \sup_{a \leq t \leq b} \frac{\xi_n(t)}{1-t} < z \right\} - P \left\{ \sup_{\frac{a}{1-a} \leq t \leq \frac{b}{1-b}} w(t) < z \right\} \right| = O_{a,b} \left(\sqrt[4]{\frac{\ln^3 n}{n}} \right).$$

Пример 2 (статистика Г. М. Манья ⁽⁶⁾). Положим $g_1(t) = -\infty$, $g_2(t) = z > 0$. Тогда получится оценка скорости сходимости в одном из вариантов критерия Г. М. Манья:

$$\begin{aligned} \sup_z \left| P \left\{ \sup_{a \leq t \leq b} \xi_n(t) < z \right\} - P \left\{ \sup_{\frac{a}{1-a} \leq t \leq \frac{b}{1-b}} (w(t) - z(t+1)) \leq 0 \right\} \right| = \\ = O_{a,b} \left(\sqrt[4]{\frac{\ln^3 n}{n}} \right). \end{aligned}$$

Точно такие же оценки скорости сходимости к пределу можно получить и для распределений некоторых других статистик, тесно связанных со статистиками Реньи и Манья. Примерами могут служить

$$\sup_{F_n(t) \geq a} \frac{|\xi_n(t)|}{F_n(t)}, \quad \sup_{a \leq F_n(t) \leq b} |\xi_n(t)|, \quad \sup_{F_n(t) \leq b} \frac{\xi_n(t)}{1-F_n(t)}, \quad \dots$$

Точные распределения статистик такого типа изучались в ⁽⁷⁾.

2. Методы, сходные с использованными в п. 1, позволяют получить оценку скорости сходимости к пределу статистики

$$U_n^2 = \int_0^1 \left(\xi_n(t) - \int_0^1 \xi_n(t) dt \right)^2 dt,$$

предложенной Ватсоном ⁽⁸⁾. Статистика U_n^2 обладает рядом преимуществ по сравнению с классической статистикой ω_n^2 при классификации распределений на окружности и имеет простое предельное распределение. Если

обозначить $K(z) = \sum_{s=-\infty}^{\infty} (-1)^s \exp(-2s^2 z^2)$, то справедлива следующая оценка:

$$\sup_z |P\{U_n^2 < z\} - K(\pi \sqrt{z})| = O \left(\sqrt[4]{\frac{\ln^3 n}{n}} \right).$$

3. В этом пункте исследуется граничная задача для эмпирического процесса, построенного по выборке случайного объема. Пусть $\xi_1, \xi_2, \dots, \xi_v$ — выборка объема v из равномерного распределения на $[0, 1]$, где v — случайная величина, не зависящая от ξ_i и распределенная по закону Пуассона с параметром λ ; $F_v(t)$ — соответствующая эмпирическая функция распределения.

Рассмотрим, следуя М. Кацу ⁽⁹⁾, эмпирический процесс вида

$$\eta_\lambda(t) = \sqrt{\lambda} \left(\frac{v}{\lambda} F_v(t) - t \right).$$

Пусть $g_i(t)$, $i = 1, 2$, — такие же кривые, как и в п. 1,

$$R_\lambda = P\{g_1(t) < \eta_\lambda(t) < g_2(t), 0 \leq t \leq 1\}.$$

В этих обозначениях справедлива

Теорема 2.

$$\left| R_\lambda - P \left\{ g_1(t) + O_K \left(\frac{\ln \lambda}{\sqrt{\lambda}} \right) < w(t) < g_2(t) + \right. \right. \\ \left. \left. + O_K \left(\frac{\ln \lambda}{\sqrt{\lambda}} \right), 0 \leq t \leq 1 \right\} \right| = O_K \left(\frac{1}{\sqrt{\lambda}} \right).$$

При доказательстве совершается переход от граничной задачи для процесса $\eta_\lambda(t)$ к граничной задаче для сумм независимых случайных величин, имеющих показательное распределение. Далее используется, например, оценка С. В. Нагаева из ⁽¹⁰⁾. В качестве примера использования теоремы 2 получим оценку скорости сходимости к пределу распределения статистики М. Каца ⁽⁹⁾. Положим $-g_1(t) = g_2(t) = z > 0$. Тогда

$$\sup_z |P\{\sup_{0 \leq t \leq 1} |\eta_\lambda(t)| < z\} - P\{\sup_{0 \leq t \leq 1} |w(t)| < z\}| = O_K \left(\frac{\ln \lambda}{\sqrt{\lambda}} \right).$$

Другими примерами могут служить рассмотренные в ⁽¹¹⁾ статистики

$$\sup_{0 \leq t \leq b} \eta_\lambda(t), \quad \sup_{0 \leq t \leq b} \frac{\eta_\lambda(t)}{1-t}, \dots$$

Автор выражает глубокую благодарность акад. Ю. В. Линнику, чье предложение получить оценки скорости сходимости в критериях Реньи и Манья послужило толчком к этой работе, а также И. А. Ибрагимову за постоянное внимание и поддержку.

Ленинградский государственный университет
им. А. А. Жданова

Поступило
21 I 1972

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

- ¹ G. Steck, Ann. Math. Stat., 42, 1, 1 (1971). ² W. Rosenkrantz, Trans. Am. Math. Soc., 139, 329 (1969). ³ А. В. Скороход, Исследования по теории случайных процессов, Киев, 1961. ⁴ A. Renyi, Acta Math. Acad. Sci. Hung., 4, 191 (1953). ⁵ А. И. Орлов, Теория вероятн. и ее примен., 16, 3, 583 (1971). ⁶ Г. М. Манья, ДАН, 69, № 4, 495 (1949). ⁷ Z. Birnbaum, B. Lientz, Zastos. mat., 10, 179 (1969). ⁸ G. Watson, Biometrika, 48, 109 (1961). ⁹ М. Кас, Proc. Nat. Acad. Sci. U.S.A., 35, 252 (1949). ¹⁰ С. В. Нагаев, Теория вероятн. и ее примен., 15, 2, 178 (1970). ¹¹ M. Csörgö, M. Alvo, Ann. Inst. St. Math., 22, 2, 257 (1970).