Записки научных
семинаров ПОМИ
Том 320, 2004 г.

А. И. Мартикайнен

О ПРЕДЕЛЬНОМ ПОВЕДЕНИИ ЭКСТРЕМАЛЬНЫХ ПОРЯДКОВЫХ СТАТИСТИК

Пусть $X, X_1, X_2, \ldots$ -последовательность независимых одинаково распределенных случайных величин, $X_{(1)} \leqslant \cdots \leqslant X_{(n)}$ – порядковые статистики, построенные по выборке объема n. В дальнейшем предполагается, что $\operatorname{ess\,sup} X = \infty$.

Исследование асимптотического поведения максимальной порядковой статистики $X_{(n)}$ восходит к работе Гнеденко (1943), предложившего следующую версию закона больших чисел.

Теорема А. *Для существования постоянных v_n, таких, что*

$$X_{(n)} - v_n \xrightarrow{P} 0,$$

необходимо и достаточно выполнение при каждом $\varepsilon > 0$ условия

$$P(X \geqslant x+\varepsilon)/P(X \geqslant x) \to 0, \quad x \to \infty. \tag{1}$$

Нетрудно видеть, что этот эффект устойчивости не может наблюдаться у решетчатых распределений (достаточно выбрать ε меньшим, чем шаг решетки). Андерсон (1970) предложил для этой ситуации иную форму устойчивости.

Теорема В. *Предположим, что распределение X сосредоточено на множестве натуральных чисел и $P(X = k) > 0$ при всех достаточно больших натуральных k. Для существования постоянных v_n, таких, что*

$$P(X_{(n)} = v_n \quad \text{или} \quad v_n + 1) \to 1,$$

необходимо и достаточно выполнение условия (1) *с $\varepsilon = 1$.*

Отметим, что теорема А вытекает из этого результата. Действительно, применим теорему В к решетчатому распределению с шагом $h > 0$ случайной величины $Y = h[X/h]$ и устремим h к 0. Здесь $[\ldots]$ – целая часть.

Работа выполнена при поддержке грантов Е02-1.0-56 Минобразования РФ, 02-01-00779 РФФИ и НШ-2258.2003.1.

Athreya и Sethuraman (2001) изучили аналогичное асимптотическое поведение “экстремальных” порядковых статистик $X_{(n-k)}$ при фиксированных $k = 1, 2, \ldots$. Sethuraman (2003) получил следующий результат.

Теорема C. *Предположим, что распределение X сосредоточено на множестве натуральных чисел, $k, l \in \{1, 2, \ldots\}$. Для существования постоянных v_n, таких, что*

$$P(X_{(n-k)} \in \{v_n, \ldots v_n + l\}) \to 1,$$

необходимо и достаточно выполнение условия (1) *с $\varepsilon = l$.*

В настоящей заметке предлагается следующее описание эффекта устойчивости, включающее в качестве частных случаев как закон Гнеденко для нерешетчатых распределений, так и упомянутые результаты для решетчатых распределений.

Теорема. *Пусть $f : R^+ \to R^+$ – положительная непрерывная неубывающая функция. Эквивалентны утверждения:*

A) *существуют постоянные v_n, такие, что*

$$P(v_n \leqslant X_{(n)} \leqslant v_n + f(v_n)) \to 1, \tag{2}$$

B) $P(X_1 > x + f(x))/P(X_1 \geqslant x) \to 0$ *при* $x \to \infty$.

Эти утверждения остаются эквивалентными и после замены в A) *максимальных порядковых статистик $X_{(n)}$ на экстремальные $X_{(n-k)}$ с фиксированным $k = 1, 2, \ldots$.*

Частные случаи:

1) при $f = \varepsilon$ получаем закон Гнеденко;
2) если X_1 принимает целые положительные значения и $f(x) = 1.5$, получаем результат Андерсона;
3) если X_1 принимает целые положительные значения и $f(x) = l + 0.5$, получаем результат Сетурамана;
4) при $f(x) = \sqrt{x}$ получаем иные асимптотики.

Доказательство теоремы. Сразу заметим, что даже при $f(x) =$ Const. не удается использовать для доказательства теорему C и прием дискретизации распределений, описанный выше – ведь утверждение доказываемой теоремы становится неверным, скажем, при замене в B) строгого неравенства на нестрогое. Ниже предлагается простое и прямое доказательство.

Если γ_n – последовательность постоянных и $\gamma_n \to 1$, то, очевидно,

$$\gamma_n^n \to 0 \Longleftrightarrow n(\gamma_n - 1) \to \infty,$$

$$\gamma_n^n \to 1 \Longleftrightarrow n(\gamma_n - 1) \to 0.$$

Соотношение (2) эквивалентно тому, что

$$P(X_{(n)} < v_n) = P(X < v_n)^n \to 0,$$

$$P(X_{(n)} \leqslant v_n + f(v_n)) = P(X \leqslant v_n + f(v_n))^n \to 1,$$

что, в свою очередь, эквивалентно соотношениям

$$nP(X \geqslant v_n) \to \infty, \quad nP(X > v_n + f(v_n)) \to 0. \tag{3}$$

Предположим, что имеет место A). Тогда $v_n \to \infty$ и выполнено (3). Пусть $x > v_1$. Выберем n так, что $v_n < x \leqslant v_{n+1}$. Учитывая, что f не убывает, получаем при $x \to \infty$

$$\frac{P(X > x + f(x))}{P(X \geqslant x)} \leqslant \frac{P(X > v_n + f(v_n))}{P(X \geqslant v_{n+1})} \sim \frac{nP(X > v_n + f(v_n))}{(n+1)P(X \geqslant v_{n+1})} \to 0,$$

то есть выполнено B).

Предположим теперь, что имеет место B). Зафиксируем $j \in \{1, 2, \ldots\}$ и при $n > j$ положим $x_{n,j} = \sup\{x : nP(X \geqslant x) \geqslant j\}$. Функция $P(X \geqslant x)$ непрерывна слева, так что $nP(X \geqslant x_{n,j}) \geqslant j$. Ясно, что $x_{n,j} \nearrow \infty$. Если $y_{n,j} > x_{n,j}$, то $nP(X \geqslant y_{n,j}) < j$, и из B) вытекает, что

$$nP(X \geqslant y_{n,j} + f(y_{n,j})) \to 0.$$

Функция $P(X > y)$ непрерывна справа, а поскольку f непрерывна и не убывает, то и функция $P(X > y + f(y))$ непрерывна справа. Постоянные $y_{n,j} > x_{n,j}$ можно выбрать так, что

$$P(X > x_{n,j} + f(x_{n,j})) \leqslant P(X > y_{n,j} + f(y_{n,j})) + n^{-2}.$$

Окончательно имеем

$$nP(X \geqslant x_{n,j}) \geqslant j \quad \text{при} \;\; n > j,$$

$$nP(X > x_{n,j} + f(x_{n,j})) \leqslant 1/j \quad \text{при} \; n > n_j.$$

Стандартное применение диагонального метода по n и j приводит к существованию постоянных v_n, для которых имеет место (3), а следовательно, и к A).

Теорема доказана для случая $k = 0$. Следующее утверждение завершает доказательство. Пусть v_n – последовательность постоянных, $I = \langle a, b\rangle$ – промежуток. При фиксированных k соотношения

$$P(X_{(n-k)} - v_n \in I) \to 1$$

эквивалентны между собой. Для открытого I это доказано в [4]. Общий случай доказывается аналогично.

Литература

1. B. V. Gnedenko, *Sur la distribution limite du terme maximum d'une série aléatoire.* — Ann. Math. **44** (1943), 423–453.
2. C. Anderson, *Extreme value theory for a class of discrete distributions with applications to some stochastic processes.* — J. Appl. Prob. **7** (1970), 99–113.
3. Atthreya, S. Sethuraman, *On the asymptotics of discrete order statistics.* — Stat. Probab. Lett. **54** (2001), 243–249.
4. S. Sethuraman, *A clustering law for some discrete order statistics.* — J. Appl. Prob. **40** (2003), 226–241.

Martikainen A. I. A note on limit behaviour of extremal order statistics.

Anderson (1970) and Sethuraman (2003) described a class of distributions concentrated on positive integers for which sample maximum and other extremes asymptotically cluster on k values. We prove a similar result for distributions of general structure.

С.-Петербургский
государственный университет

Поступило 12 ноября 2004 г.