МАТЕМАТИКА

УДК 519.2
MSC 62G32

О РЕКОРДНЫХ ВЕЛИЧИНАХ В ПОСЛЕДОВАТЕЛЬНОСТЯХ ВЫБОРОЧНЫХ РАЗМАХОВ

И. В. Бельков, В. Б. Невзоров

Санкт-Петербургский государственный университет,
Российская Федерация, 199034, Санкт-Петербург, Университетская наб., 7–9

При работе с рекордными величинами важную роль играет классическое представление рекордов в последовательностях независимых случайных величин, имеющих стандартное $E(1)$-экспоненциальное распределение, в виде сумм экспоненциально распределенных случайных слагаемых. Предложено некоторое обобщение данного представления. Получен новый аналогичный результат, который позволяет подобным же образом через суммы независимых экспоненциально распределенных случайных величин представлять рекордные значения выборочных размахов. Библиогр. 5 назв.

Ключевые слова: рекордные величины, экспоненциальное распределение, отрицательное экспоненциальное распределение, распределение Лапласа, выборочные размахи.

Введение. Рассмотрим последовательность независимых случайных величин (с. в.) $X_1, X_2, \ldots$ с общей непрерывной функцией распределения (ф. р.) $F(x)$. Этой последовательности соответствуют наборы порядковых статистик

$$X_{1,n} = \min\{X_1, X_2, \ldots, X_n\} \le X_{2,n} \le \ldots \le X_{n,n} = \max\{X_1, X_2, \ldots, X_n\}, \quad n = 1, 2, \ldots$$

По этой же последовательности можно сформировать верхние рекордные моменты $L(n)$, $n = 1, 2, \ldots$, и верхние рекордные величины $X(1) < X(2) < \ldots < X(n) < \ldots$, которые задаются следующим образом:

$$\begin{gathered} L(1) = 1, \quad X(1) = X_1, \\ L(n) = \min\{j : X_j > X_{L(n-1)}\}, \quad X(n) = X_{L(n)}, \quad n = 2, 3, \ldots \end{gathered} \tag{1}$$

Аналогично можно получить последовательности нижних рекордных моментов $l(n)$, $n = 1, 2, \ldots$, и нижних рекордных величин $x(1) > x(2) > \ldots > x(n) > \ldots$, которые определяются соотношениями

$$l(1) = 1, \quad x(1) = X_1,$$
$$l(n) = \min\{j : X_j < x_{l(n-1)}\}, \quad x(n) = x_{l(n)}, \quad n = 2, 3, \ldots \tag{2}$$

В прикладных задачах теории вероятностей и математической статистики часто приходится иметь дело с порядковыми статистиками и рекордными величинами, которые в отличие от исходных $X_1, X_2, \ldots$ теряют важное и удобное во всех отношениях свойство независимости. На примере рекордных величин посмотрим, каким образом и в каких случаях можно применять методы, разработанные для сумм независимых случайных слагаемых. Отметим, что изучению рекордов и их свойств посвящены многие монографии, например, [1–4].

Важную роль при изучении верхних рекордных величин играет семейство экспоненциальных распределений. Остановимся на стандартном представителе этого семейства — экспоненциальном $E(1)$-распределении, представленном плотностью распределения $h_1(x)$ и ф. р. $H_1(x)$, которые имеют вид

$$H_1(x) = \max\{0, 1 - e^{-x}\}, \quad -\infty < x < \infty,$$

и

$$h_1(x) = 0, \quad \text{если} \quad x < 0, \quad \text{и} \quad h_1(x) = e^{-x}, \quad \text{если} \quad x \geq 0.$$

Пусть Z_1, Z_2, … — последовательность независимых с. в., имеющих общую ф. р. $H_1(x)$, а $Z(1) = Z_1 < Z(2) < \ldots . < Z(n) < Z(n+1) < \ldots$ — соответствующая последовательность верхних рекордов. В формулируемом ниже соотношении для рекордных значений $Z(n)$ и далее в тексте статьи будут встречаться независимые с. в. ξ_1, ξ_2, …, тоже имеющие $E(1)$-экспоненциальное распределение. Отметим также, что символ $\stackrel{d}{=}$ в соотношениях вида $X \stackrel{d}{=} Y$ будет обозначать равенство по распределению соответствующих случайных величин или случайных векторов. Справедлив следующий результат.

Представление 1. *При любом $n = 1, 2, \ldots$ для рекордных величин $Z(1) < Z(2) < \ldots < Z(n)$ справедливо соотношение*

$$\{Z(1), Z(2), \ldots, Z(n)\} \stackrel{d}{=} \{\xi_1, \xi_1 + \xi_2, \ldots, \xi_1 + \xi_2 + \ldots + \xi_n\}, \tag{3}$$

где $\xi_1, \xi_2, \ldots$ — независимые $E(1)$-распределенные случайные величины.

Замечание 1. Отметим, что равенство (3) часто оказывается полезным при изучении рекордных величин $X(1) < X(2) < \ldots$, построенных по последовательностям $X_1, X_2, \ldots$, с произвольной непрерывной ф. р. $F(x)$. В этом случае преобразование Смирнова позволяет получить для любого $n = 1, 2, \ldots$ соотношение

$$\{X(1), X(2), \ldots, X(n)\} \stackrel{d}{=} \{G(\xi_1), G(\xi_1 + \xi_2), \ldots, G(\xi_1 + \xi_2 + \ldots + \xi_n)\}, \tag{4}$$

где $G(x) = F^{-1}(1 - e^{-x})$ и $F^{-1}(x)$ — функция, обратная ф. р. $F(x)$. Видим, что равенство (4) позволяет и в случае любой непрерывной ф. р. $F(x)$ в какой-то степени использовать классические методы, разработанные для сумм независимых случайных слагаемых.

Соотношение (3) позволяет легко получить аналогичные равенства для нижних рекордов.

Представление 2. *Если независимые с. в. $X_1, X_2, \ldots$ имеют отрицательное экспоненциальное распределение с ф. р. $H_0(x) = e^x$, $x < 0$, то для любого $n = 1, 2, \ldots$ и соответствующих нижних рекордов $x(1) > x(2) > \ldots > x(n)$ справедливо равенство*

$$\{x(1), x(2), \ldots, x(n)\} \stackrel{d}{=} \{-Z(1), -Z(2), \ldots, -Z(n)\} \stackrel{d}{=}$$
$$\stackrel{d}{=} \{-\xi_1, -(\xi_1 + \xi_2), \ldots, -(\xi_1 + \xi_2 + \ldots + \xi_n)\}. \quad (5)$$

Цель данной работы — исследовать распределения рекордных значений в последовательностях выборочных размахов $W_n = X_{n,n} - X_{1,n}$, $n = 1, 2, \ldots$ В условиях представлений 1 и 2, где рассматриваются ограниченные нулем (соответственно, снизу и сверху) случайные величины, вполне естественно в качестве начала отсчета взять вырожденную с. в. $Z_0 = 0$ (в первом случае) и $X_0 = 0$ (во втором случае). Тогда, если и в первом, и во втором случаях рассмотреть рекордные значения $W(1) < W(2) < \ldots$ соответствующих выборочных размахов, то получим для них удобное единообразное представление вида

$$\{W(1), W(2), \ldots, W(n)\} \stackrel{d}{=} \{\xi_1, \xi_1 + \xi_2, \ldots, \xi_1 + \xi_2 + \ldots + \xi_n\}, \quad n = 1, 2, \ldots \quad (6)$$

Ниже мы рассмотрим существенно более общую ситуацию, в которой для рекордных значений выборочных размахов справедливым остается соотношение (6).

Две модели построения выборочных размахов. Опишем вначале рекордную схему, предложенную в работе [5] и рассмотренную подробно в монографии [3]. В последовательности независимых с. в. $X_1, X_2, \ldots$ с общей непрерывной ф. р. $F(x)$ величина X_1 объявляется одновременно нулевым верхним и нулевым нижним рекордами. Очередными рекордными моментами в этой последовательности объявляются моменты m, когда $X_m > X_{m-1,m-1}$ (появление верхнего рекорда) или когда $X_m < X_{1,m-1}$ (появление нижнего рекорда), т. е. по существу рекордным в этой постановке является любой момент, при котором очередной размах выборки $W_m = X_{m,m} - X_{1,m}$ становится больше всех предыдущих размахов. Пусть $1 < T(1) < T(2) < \ldots$ представляют собой такого рода рекордные моменты. С каждым из этих моментов связывают с. в. $S(n) = \max\{X_1, X_2, \ldots, X_{T(n)}\}$ и $R(n) = \min\{X_1, X_2, \ldots, X_{T(n)}\}$, $n = 1, 2, \ldots$, т. е. $S(n)$ и $R(n)$ совпадают, соответственно, с верхним и нижним на данный момент рекордными значениями в последовательности $X_1, X_2, \ldots$ Получены (см., например, § 8.4 в монографии [3]) плотности распределения с. в. $S(n), R(n)$ и соответствующих рекордных выборочных размахов $W(n) = S(n) - R(n)$, $n = 1, 2, \ldots$

В частности, показано, что если с. в. $X_1, X_2, \ldots$ имеют плотность распределения $f(x)$, то плотность распределения $p_n(x)$ размаха $W(n)$ принимает вид

$$p_n(x) = \frac{2^n}{(n-1)!} \int_{-\infty}^{\infty} f(x+u) f(u) (-\log(1 - F(x+u) + F(u)))^{n-1} du. \quad (7)$$

Отметим здесь также следующий результат для с. в. $S(n)$.

Представление 3. *Если независимые с. в. $X_1, X_2, \ldots$ имеют общую ф. р. $H_1(x) = \max\{0, 1 - e^{-x}\}, -\infty < x < \infty$; то при каждом фиксированном $n = 1, 2, \ldots$ справедливо равенство*

$$S(n) \stackrel{d}{=} \xi_1 + \frac{1}{2} \sum_{k=2}^{n-1} \xi_k, \quad (8)$$

где, как и выше, $\xi_1, \xi_2, \ldots$ — независимые $E(1)$-распределенные случайные величины.

Видим, что данный результат дает некоторое представление о рекордах в последовательности размахов $W_1, W_2, \ldots$

Мы рассмотрим несколько иной подход к изучению рекордных размахов $W(n)$. Предлагаемая нами схема позволяет получить результат, частными случаями которого являются равенства (3) и (5).

Приведенное выше представление 1 получено для исходных случайных величин, имеющих общую ф. р.

$$H_1(x) = \max\{0, 1 - e^{-x}\}, \qquad -\infty < x < \infty.$$

Представление 2 работает для X_n, $n = 1, 2, \ldots$, с ф. р.

$$H_0(x) = \min\{1, e^x\}, \qquad -\infty < x < \infty.$$

Объединяя в одно семейство эти две функции распределения, введем ф. р. $H_p(x)$, $0 \le p \le 1$, которая представляется как сумма

$$H_p(x) = (1-p)H_0(x) + pH_1(x), \qquad -\infty < x < \infty, \tag{9}$$

с весами $1-p$ и p исходных ф. р. $H_0(x)$ и $H_1(x)$. В частности, при $p = 0$ эта ф. р. совпадает с ф. р. $H_0(x)$, а при $p = 1$ получаем ф. р. $H_1(x)$.

Функция $H_p(x)$ имеет следующий вид:

$$H_p(x) = (1-p)e^x, \quad \text{если} \quad x < 0, \quad \text{и} \quad H_p(x) = 1 - p + pe^{-x}, \quad \text{если} \quad x \ge 0.$$

Отметим также, что при $p = 1/2$ имеем дело с распределением Лапласа с плотностью

$$h_{1/2}(x) = \frac{e^{-|x|}}{2}, \qquad -\infty < x < \infty.$$

Рассмотрим теперь последовательность независимых случайных величин $X_1, X_2, \ldots$ с общей ф. р. $H_p(x)$. Пусть также $X_0 = 0$ представляет собой начальный элемент в этой последовательности. В последовательности размахов $W_n = \max\{0, X_{n,n}\} - \min\{0, X_{1,n}\}$ выделим рекордные значения $W(1) = |X_1| < W(2) < \ldots$ Для удобства изложения положим $W(0) = 0$. Справедлив следующий результат.

Теорема. *Если независимые с. в. $X_1, X_2, \ldots$ имеют общую ф. р. $H_p(x)$, то приращения рекордов $W(1) - W(0), W(2) - W(1), \ldots$ также независимы, и каждое из этих приращений имеет экспоненциальное $E(1)$-распределение.*

Этот результат можно переформулировать следующим образом.

Представление 4. *В рассматриваемой схеме при любом $n = 1, 2, \ldots$ справедливо соотношение*

$$\{W(1), W(2), \ldots, W(n)\} \stackrel{d}{=} \{\xi_1, \xi_1 + \xi_2, \ldots, \xi_1 + \xi_2 + \ldots + \xi_{n-1} + \xi_n\}.$$

Замечание 2. Видим, что в данной схеме распределения рекордных величин в последовательности выборочных размахов одинаковы для любого значения $0 \le p \le 1$. При $p = 1$ из доказываемого результата следует как частный случай соотношение (3). При $p = 0$ получаем результат, указанный в представлении 2. Отметим также, что если зафиксировать $n = 1, 2, \ldots$, то математическое ожидание числа рекордных значений среди размахов $W_1, W_2, \ldots, W_n$ равно сумме

$$\sum_{k=1}^{n} \frac{2 - p^k - (1-p)^k}{k}.$$

Доказательство теоремы. Каждую с. в. X_k с ф. р. $H_p(x)$ можно представить в виде

$$X_k = \eta_k Z_k, \qquad k = 1, 2, \ldots,$$

где $Z_1, Z_2, \ldots$ — независимые с. в., имеющие общую ф. р. $H_1(x) = 1 - e^{-x}, x \geq 0$, а $\eta_1, \eta_2, \ldots$ — независимые между собой и не зависящие от $Z_1, Z_2, \ldots$ случайные величины, принимающие значения -1 и 1 с вероятностями, равными, соответственно, $1-p$ и p. Зафиксируем номер $N(n)$, которому отвечает момент появления рекордного значения $W(n)$. Среди величин $X_1, X_2, \ldots, X_{N(n)}$ будет какое-то случайное число $M(n)$ положительных членов и $R(n) = N(n) - M(n)$ отрицательных членов этой последовательности. Понятно, что если $N(n) = m, m = n, n+1, \ldots$, то $M(n)$ имеет биномиальное распределение с вероятностями

$$P\{M(n) = r\} = P\{R(n) = m - r\} = C_m^r p^r (1-p)^{m-r}, \quad r = 0, 1, 2, \ldots, m.$$

Наличие r верхних рекордов в последовательности $X_1, X_2, \ldots, X_{N(n)}$ равносильно наличию r рекордов в наборе $E(1)$-распределенных с. в. $Z_1, Z_2, \ldots, Z_{T(n)}$. Каждое такое событие (появление очередного рекордного значения размаха, обусловленного появлением еще одного верхнего рекорда в последовательности $X_1, X_2, \ldots$) означает, что к предыдущему рекордному значению размаха нужно добавить очередное $E(1)$-распределенное слагаемое, которое не зависит от всех предыдущих рекордных величин размаха. Аналогично можно рассмотреть поведение нижних рекордов в последовательности, состоящей из отрицательных X_n. В этом случае так же появление очередного нижнего рекорда в последовательности этих X_n приводит к необходимости добавить к имеющемуся на этот момент рекордному значению выборочного размаха не зависящее от его величины $E(1)$-распределенное случайное слагаемое. Получаем, что каждый раз, когда в исходной последовательности $X_1, X_2, \ldots$ встречается верхнее или нижнее рекордное значение, очередное рекордное значение размаха выборки увеличивается по сравнению с предыдущим на слагаемое, имеющее экспоненциальное $E(1)$-распределение, и размер этого увеличения не зависит от процесса роста размаха до данного рекордного момента. Все это доказывает утверждение теоремы.

Литература

1. *Ahsanullah M., Yanev G. P.* Records and Branching Processes. New-York: Nova Science Publishers, 2008. 175 p.

2. *Ahsanullah M.* Record Values — Theory and Applications. University Press of America, 2004. 314 p.

3. *Arnold B. C., Balakrishnan N., Nagaraja H. N.* Records. New-York: John Wiley & Sons, 1998. 314 p.

4. *Невзоров В. Б.* Рекорды. Математическая теория. М.: Фазис, 2000. 244 с.

5. *Houchens R. L.* Record Value Theory and Inference. Ph.D. Dissertation, University of California, Riverside, California. 1984.

Статья поступила в редакцию 1 июля 2016 г.; рекомендована в печать 22 июня 2017 г.

Сведения об авторах

Бельков Игорь Владимирович — аспирант; igor.belkov@gmail.com

Невзоров Валерий Борисович — доктор физико-математических наук, профессор; vanev@mail.ru

ON RECORD VALUES IN SEQUENCES OF SAMPLE RANGES

Igor V. Belkov, Valery B. Nevzorov

St. Petersburg State University, Universitetskaya nab., 7–9, St. Petersburg, 199034, Russian Federation; igor.belkov@gmail.com, vanev@mail.ru

The classical presentation of record values in sequences of independent exponentially distributed random variables via sums of independent summands is very important in the mathematical theory of records. Some generalization of this presentation is suggested. A new analogical result which allows us to express the record values of the sample ranges via sums of independent exponentially distributed random variables is obtained. Refs 5.

Keywords: record values, exponential distribution, negative exponential distribution, Laplace distribution, sample ranges.

References

1. Ahsanullah M., Yanev G. P., *Records and Branching Processes* (Nova Science Publishers, New York, 2008, 175 p.).

2. Ahsanullah M., *Record Values — Theory and Applications* (University Press of America, 2004, 314 p.).

3. Arnold B. C., Balakrishnan N., Nagaraja H. N., *Records* (John Wiley & Sons, New York, 1998, 314 p.).

4. Nevzorov V. B., *Records. Mathematical Theory.* In Ser. *Translations of Mathematical Monographs* (American Math. Society, **194**, 2001).

5. Houchens R. L., *Record Value Theory and Inference* (Ph.D. Dissertation, University of California, Riverside, California, 1984).

Для цитирования: *Бельков И. В., Невзоров В. Б.* О рекордных величинах в последовательностях выборочных размахов // Вестник СПбГУ. Математика. Механика. Астрономия. 2017. Т. 4 (62). Вып. 4. С. 535–540. https://doi.org/10.21638/11701/spbu01.2017.401

For citation: Belkov I. V., Nevzorov V. B. On record values in sequences of sample ranges. *Vestnik SPbSU. Mathematics. Mechanics. Astronomy*, 2017, vol. 4 (62), issue 4, pp. 535–540. https://doi.org/10.21638/11701/spbu01.2017.401