Записки научных
семинаров ПОМИ
Том 339, 2006 г.

И. С. Ахундов, В. Б. Невзоров

УСЛОВНЫЕ РАСПРЕДЕЛЕНИЯ РЕКОРДНЫХ ВЕЛИЧИН И ХАРАКТЕРИЗАЦИИ РАСПРЕДЕЛЕНИЙ

Рассмотрим последовательность независимых случайных величин (с.в.) $X_1, X_2, \ldots$ с одинаковой непрерывной функцией распределения (ф.р.) F. Определим рекордные моменты $L(1)$, $L(2), \ldots$ и рекордные величины $X(1), X(2), \ldots$ следующим образом:

$$L(1) = 1,\ L(n+1) = \min\{j : X_j > X_{L(n)}\},\ n = 1, 2, \ldots; \qquad (1)$$

$$X(n) = X_{L(n)},\ n = 1, 2, \ldots \qquad (2)$$

Введенные с.в. являются верхними рекордными моментами и верхними рекордными величинами соответственно. Если в (1) знак неравенства “ $>$” заменить на “ $<$”, то будем иметь дело с нижними рекордными моментами и нижними рекордными величинами. Результаты для нижних рекордов легко получаются из соответствующих результатов для верхних рекордов, если вместо исходных с.в. рассматривать величины $Y_1 = -X_1$, $Y_2 = -X_2, \ldots$. Поэтому ограничимся изучением только верхних рекордных величин $X(n)$, $n = 1, 2, \ldots$, называя в дальнейшем их просто рекордными величинами. Основы теории рекордов приведены в монографиях [1, 2].

Ниже мы рассмотрим ряд регрессионных соотношений для рекордных величин и получим новые характеризации распределений, связанные с этими соотношениями. Некоторые подобные характеризации распределений регрессионными свойствами рекордов ранее были получены в работах [3–7] (см. также §§4.2.2 и 4.2.3 в [1]).

Условные распределения рекордов

Предположим дополнительно, что исходные с. в. $X_1, X_2, \ldots$ имеют плотность распределения f. Тогда (см., например, равенство 15.9 в [2]) совместная плотность распределения рекордных величин $X(1)$, $X(2),\ldots,X(n)$ задается следующим образом:

$$f_{1,2,\ldots,n}(x_1,x_2,\ldots,x_n)=r(x_1)r(x_2)\ldots r(x_n)(1-F(x_n)), \qquad (3)$$

если $x_1<x_2<\cdots<x_n$, где $r(x)=f(x)/(1-F(x))$, и

$$f_{1,2,\ldots,n}(x_1,x_2,\ldots,x_n)=0 \quad \text{иначе.}$$

Из (3) следует, что для любого $1\leqslant k\leqslant n$ совместная плотность распределения с. в. $X(k)$, $X(k+1),\ldots,X(n)$ имеет вид

$$f_{k,\ldots,n}(x_k,\ldots,x_n)=\frac{r(x_k)\ldots r(x_n)(1-F(x_n))(-\ln(1-F(x_k)))^{k-1}}{(k-1)!} \qquad (4)$$

если $x_k<x_{r+1}<\cdots<x_n$. В частности, для любого $n\geqslant 1$ плотность распределения с. в. $X(n)$ задается равенством

$$f_n(x)=\frac{(-\ln(1-F(x))^{n-1}f(x)}{(n-1)!} \qquad (5)$$

Из (4) легко выводится, что для $1\leqslant k<n$ двумерная плотность распределения рекордных величин $X(k)$ и $X(n)$ имеет вид

$$f_{k,n}(u,v)=\frac{(\ln(1-F(u)))^{k-1}(\ln(1-F(u))-\ln(1-F(v)))^{n-k-1}r(u)f(v)}{(k-1)!(n-k-1)!} \qquad (6)$$

если $u<v$.

Из соотношений (4) и (6) следует, что условная плотность распределения рекордов $X(k+1),\ldots,X(n-1)$ при условии, что $X(k)=u$ и $X(n)=v$, задается равенством

$$f_{k+1,\ldots n-1|k,n}(x_{k+1},\ldots,x_{n-1}|u,v)=\frac{(n-k-1)!r(x_{k+1})\ldots r(x_{n-1})}{(\ln(1-F(u))-\ln(1-F(v)))^{n-k-1}}, \qquad (7)$$

где $u<x_{k+1}<\cdots<x_{n-1}<v$.

Из соотношений (3) и (5) вытекает, что условная плотность распределения рекордов $X(1),\ldots,X(k-1)$, $X(k+1),\ldots,X(n)$ при условии, что $X(k)=u$, имеет вид

$$\begin{aligned} &f_{1,\ldots,k-1,k+1,\ldots,n|k}(x_1,\ldots,x_{k-1},x_{k+1},\ldots,x_n|u)\\ &\quad=T(x_1,\ldots,x_{k-1},u)S(u,x_{k+1},\ldots,x_n), \end{aligned} \qquad (8)$$

где

$$T(x_1,\ldots,x_{k-1},u)=\frac{(k-1)!r(x_1)\ldots r(x_{k-1})}{(\ln(1-F(u))^{k-1}}, \qquad (9)$$

$$S(u, x_{k+1}, \ldots, x_n) = \frac{r(x_{k+1}) \ldots r(x_n)(1 - F(x_n))}{(1 - F(u))} \tag{10}$$

и

$$x_1 < x_2 < \cdots < x_{k-1} < u < x_{k+1} < \cdots < x_n.$$

Отсюда получаем, что условная совместная плотность рекордов $X(k+1), \ldots, X(n)$ при фиксированном значении $X(k) = u$ задается равенством

$$f_{k+1,\ldots,n|k}(x_{k+1}, \ldots, x_n | u) = \frac{r(x_{k+1}) \ldots r(x_n)(1 - F(x_n))}{(1 - F(u))},$$
$$u < x_{k+1} < \cdots < x_n. \tag{11}$$

Из приведенных соотношений вытекает ряд интересных следствий.

Следствие 1. *Для любого* $1 < k < n$ *случайные векторы* $(X(1), \ldots, X(k-1))$ *и* $(X(k+1), \ldots, X(n))$ *условно независимы при любом фиксированном значении* $X(k) = u$. *При этом условное распределение рекордов* $X(1)$, $X(2), \ldots, X(k-1)$ *совпадает с безусловным распределением порядковых статистик*

$$Y_{1,k-1} \leqslant Y_{2,k-1} \leqslant \cdots \leqslant Y_{k-1,k-1},$$

построенных по независимым с. в. Y_1, $Y_2, \ldots, Y_{k-1}$, *имеющим ф. р.*

$$H_u(x) = \frac{\ln(1 - F(x))}{\ln(1 - F(u))}, \quad x \leqslant u.$$

Условное же распределение рекордов $X(k+1), \ldots, X(n)$ *совпадает с безусловным распределением рекордных величин*

$$V(1) < V(2) < \cdots < V(n-k),$$

построенных по последовательности независимых с. в. V_1, $V_2, \ldots$ *с ф. р.*

$$V_u(x) = \frac{(F(x) - F(u))}{(1 - F(u))} = P\{X < x | X > u\}, \quad x \geqslant u.$$

Замечание 1. Тот факт, что условное распределение рекордов $X(1)$, $X(2), \ldots, X(k-1)$ при фиксированном значении $X(k) = u$ совпадает с распределением порядковых статистик, соответствующих ф. р. H_u, был получен в 1978 г. в работе [4].

Следствие 2. *Из соотношения (7) получаем, что условное распределение рекордов $X(k+1), \ldots, X(n-1)$ при условии, что $X(k) = u$ и $X(n) = v$, совпадает с безусловным распределением порядковых статистик*

$$W_{1,n-k-1} \leqslant W_{2,n-k-1} \leqslant \cdots \leqslant W_{n-k-1,n-k-1},$$

построенных по независимым с. в. $W_1, \ldots, W_{n-k-1}$, имеющим ф. р.

$$W^{(u,v)}(x) = \frac{(\ln(1-F(x)) - \ln(1-F(u)))}{(\ln(1-F(v)) - \ln(1-F(u)))}, u \leqslant x \leqslant v.$$

Замечание 2. Результаты, сформулированные в следствиях 1 и 2, формально получены в предположении, что исходные с. в. X_1, $X_2, \ldots$ имеют абсолютно непрерывное распределение. Тем не менее они остаются справедливыми и при более общем предположении непрерывности ф. р. F. Дело в том, что если F – произвольная непрерывная ф. р., то рассматривая обратную к ней функцию G, можно воспользоваться равенством

$$(X(1), \ldots, X(n)) \stackrel{d}{=} (G(U(1)), \ldots, G(U(n))), \tag{12}$$

где $U(1) < U(2) < \ldots$ – рекордные величины для равномерного на интервале $[0, 1]$ распределения. Для рекордов $U(1)$, $U(2), \ldots$ результаты, приведенные в следствиях 1 и 2, справедливы. Далее, воспользовавшись равенством (12), нетрудно показать, что эти результаты остаются справедливыми и для рекордных величин $X(1)$, $X(2), \ldots$.

Регрессионные соотношения, связывающие рекорды

Различного рода регрессионные соотношения, связывающие две соседние рекордные величины, рассматривались в работах [3, 4, 7]. По три рекордных величины фигурировали в регрессионных равенствах, изучавшихся в работах [5] (соотношения вида $E((X(n-1) + X(n+1))|X(n)) = 2X(n))$ и [6] (равенства типа $E(\psi(X(n))|X(n-1), X(n+1)) = \varphi(X(n-1), X(n+1)))$.

Мы рассмотрим регрессионные соотношения для большего числа рекордов.

Лемма 1. *Пусть F – непрерывная функция распределения. Тогда для любого $n = 2, 3, \dots$ и любой функции h такой, что*

$$\int\limits_{-\infty}^{\infty} \frac{|h(x)|}{(1 - F(x))} dF(x) < \infty,$$

выполняется равенство

$$E((h(X(1)) + \cdots + h(X(n-1)))|X(n) = u)$$

$$= -\frac{n-1}{\ln(1 - F(u))} \int\limits_{-\infty}^{u} \frac{h(x)}{(1 - F(x))} dF(x), \qquad (13)$$

справедливое для всех u, являющихся точками роста ф. р. F.

Доказательство. Равенство (13) получается из результата, приведенного в следствии 1, если учесть замечание 2. Действительно, при фиксированном значении $X(n) = u$ условное распределение с. в. $X(1)$, $X(2), \dots, X(n-1)$ совпадает с безусловным распределением порядковых статистик

$$Y_{1,n-1}, Y_{2,n-1}, \dots, Y_{n-1,n-1},$$

построенных по независимым с. в. Y_1, $Y_2, \dots, Y_{n-1}$ с ф. р.

$$H_u(x) = \frac{\ln(1 - F(x))}{\ln(1 - F(u))}, \quad x \leqslant u.$$

Тогда

$$E((h(X(1)) + \cdots + h(X(n-1)))|X(n) = u)$$

$$= E(h(Y_{1,n-1}) + \cdots + h(Y_{n-1,n-1}))$$

$$= E(h(Y_1) + \cdots + h(Y_{n-1})) = (n-1)Eh(Y_1)$$

$$= (n-1) \int\limits_{-\infty}^{u} h(x) dH_u(x) = -\frac{n-1}{\ln(1 - F(u))} \int\limits_{-\infty}^{u} \frac{h(x)}{(1 - F(x))} dF(x).$$

Лемма 2. *Пусть F – непрерывная функция распределения. Тогда для $1 \leqslant k \leqslant n-2$, любой функции h и любых u и v $(u < v)$ – точек роста функции F справедливы соотношения*

$$E((h(X(k+1)) + \cdots + h(X(n-1)))|X(k) = u, X(n) = v)$$

$$= \frac{n-k-1}{\ln(1-F(u)) - \ln(1-F(v))} \int\limits_u^v \frac{h(x)}{(1-F(x))} dF(x)$$

$$= E((h(X(2)) + \cdots + h(X(n-k)))|X(1) = u, X(n-k+1) = v). \quad (14)$$

Доказательство. Вспоминая утверждение следствия 2 и определенные там порядковые статистики

$$W_{1,n-k-1} \leqslant W_{2,n-k-1} \leqslant \cdots \leqslant W_{n-k-1,n-k-1},$$

получаем, что

$$E((h(X(k+1)) + \cdots + h(X(n-1)))|X(k) = u, X(n) = v)$$

$$= E(h(W_{1,n-k-1}) + \cdots + h(W_{n-k-1,n-k-1})) = E(h(W_1) + \ldots$$

$$+h(W_{n-k-1})) = (n-k-1)Eh(W_1) = (n-k-1) \int\limits_u^v h(x) dW^{(u,v)}(x)$$

$$= \frac{n-k-1}{\ln(1-F(u)) - \ln(1-F(v))} \int\limits_u^v \frac{h(x)}{(1-F(x))} dF(x). \quad (15)$$

Остается заметить, что финальное выражение в соотношении (15) зависит от k и n через их разность. Отсюда следует, что для любого $k = 1, 2, \ldots$ справедливо равенство

$$E((h(X(2)) + \cdots + h(X(n-k)))|X(1) = u, X(n-k+1) = v)$$

$$= E((h(X(k+1)) + \cdots + h(X(n-1)))|X(k) = u, X(n) = v).$$

Характеризации распределений регрессионными свойствами сумм рекордных величин

1). Рассмотрим соотношение

$$E((X(1)+\cdots+X(n))|X(n+1)=u)=cu+d \tag{16}$$

и найдем среди непрерывных ф. р. F такие, для которых (16) справедливо для всех точек роста u функции F. Из утверждения леммы 1 следует, что (16) равносильно равенству

$$-\frac{n}{\ln(1-F(u))}\int\limits_{-\infty}^{u} xd(-\ln(1-F(x)))=cu+d.$$

Обозначим

$$T(x)=-\ln(1-F(x)),\ \alpha=\frac{c}{n},\ \beta=\frac{d}{n}.$$

Приходим к равенству

$$\int\limits_{-\infty}^{u} xdT(x)=(\alpha u+\beta)T(u). \tag{17}$$

Оно сводится к достаточно простому дифференциальному уравнению

$$\alpha T(u)=T'(u)(u(1-\alpha)-\beta), \tag{18}$$

при решении которого надо рассматривать три варианта: $\alpha<1$, $\alpha=1$ и $\alpha>1$. В результате выясняется, что регрессионное соотношение (16) имеет место, когда ф. р. F принадлежит одному из трех типов функций, приведенных ниже в теореме 1. Но можно пойти по еще более простому пути доказательства нужного факта и обратиться для этого к результату, полученному в работе [4]. Там было показано, что если выполняется равенство

$$E(X(n)|X(n+1)=u)=cu+d, \tag{19}$$

то с точностью до сдвига и изменения масштаба соответствующая функция распределения F имеет вид:

$$F(x)=1-\exp(-x^{\theta}), x>0, \quad \text{если} \quad 0<c<1;$$

$$F(x)=1-\exp(-\exp(x)), \quad \text{если} \quad c=1;$$

$$F(x) = 1 - \exp(-(-x)^{\theta}), x < 0, \quad \text{если} \quad c > 1,$$

где

$$\theta = \frac{c}{n(1-c)}.$$

Видим, что при $n = 1$ соотношения (16) и (19) совпадают. Это позволяет воспользоваться результатом работы [4], в котором остается в качестве c взять $\frac{c}{n}$, и получить следующую теорему.

Теорема 1. *Пусть F – непрерывная функция распределения. Для того, чтобы при некоторых $n = 1, 2, \ldots, c > 0$, $-\infty < d < \infty$ выполнялось равенство*

$$E((X(1) + \cdots + X(n))|X(n+1) = u) = cu + d \quad \textit{п. н.},$$

необходимо и достаточно, чтобы с точностью до параметров масштаба и сдвига F имела следующий вид:

$$a) \quad F(x) = 1 - \exp(-x^{\frac{c}{(n-c)}}), x > 0, \quad \textit{и} \quad F(x) = 0, x \leqslant 0,$$

$$\textit{если} \quad 0 < c < n;$$

$$b) \quad F(x) = 1 - \exp(-\exp(x)), -\infty < x < \infty, \quad \textit{если} \quad c = n;$$

$$c) \quad F(x) = 1 - \exp(-(-x)^{\frac{c}{(n-c)}}), x < 0, \quad \textit{и} \quad F(x) = 1, x \geqslant 0,$$

$$\textit{если} \quad c > n.$$

2). Рассмотрим теперь еще одну ситуацию. Предположим, что при некоторых α, β, γ и $1 \leqslant k \leqslant n-2$ справедливо соотношение

$$E((X(k+1) + \cdots + X(n-1))|X(k) = u, X(n) = v) = \alpha u + \beta v + \gamma \quad \text{п. н.} \tag{20}$$

и найдем вид непрерывных ф. р. F, для которых (20) справедливо. Нетрудно убедиться, что в правой части этого соотношения γ обязано равняться нулю, а $\alpha + \beta = n - k - 1$. Для этого достаточно в (20) устремить u к v и получить равенство

$$(n-k-1)v = (\alpha + \beta)v + \gamma,$$

справедливое для всех точек роста v непрерывнвой ф. р. F. Таким образом, мы будем иметь дело с соотношением

$$E((X(k+1) + \cdots + X(n-1))|X(k) = u, X(n) = v) = \alpha u + \beta v, \tag{21}$$

где $\alpha+\beta=n-k-1$. Если воспользоваться утверждением леммы 2, из (21) получим равенство

$$\frac{n-k-1}{\ln(1-F(v))-\ln(1-F(u))}\int\limits_u^v xd(\ln(1-F(x)))=\alpha u+\beta v. \qquad (22)$$

Обозначим

$$T(x)=-\ln(1-F(x))$$

и перепишем (22) в виде

$$\int\limits_u^v xdT(x)=(\delta u+\gamma v)(T(v)-T(u)), \qquad (23)$$

где $\delta+\gamma=1$. Из (23) дифференцированием по v получаем равенство

$$vT'(v)=\gamma(T(v)-T(u))+(\delta u+\gamma v)T'(v),$$

из которого следует, что

$$\delta uT'(v)-\gamma T(u)=C(v), \qquad (24)$$

где

$$C(v)=vT'(v)-\gamma T(v)-\gamma vT'(v).$$

Из (24) выводим, что $T(u)$ должна быть линейной функцией, т. е.

$$T(u)=cu+d$$

при некоторых значениях c и d, а тогда

$$F(x)=1-\exp(-(cx+d)).$$

Поскольку $F(x)$ – функция распределения, то $c>0$. Подставляя в (23) $T(x)=cx+d$, приходим к равенству

$$v+u=2(\delta u+\gamma v),$$

из которого уже следует, что $\delta=\gamma=1/2$. Таким образом, справедлив следующий результат.

Теорема 2. *Если при некоторых α, β, γ и $1\leqslant k\leqslant n-2$ выполнено соотношение* (20), *то $\gamma=0$, $\alpha=\beta=(n-k-1)/2$ и с точностью до параметров масштаба и сдвига*

$$F(x)=1-\exp(-x), x\geqslant 0.$$

Работа частично поддержана грантами РФФИ 06-01-00763 и 06-01-00179.

Литература

1. B. C. Arnold, N. Balakrishnan, H. N. Nagaraja, *Records.* Wiley, NY (1998).
2. В. Б. Невзоров, *Рекорды. Математическая теория.* Фазис. М. (2000).
3. H. N. Nagaraja, *On a characterization based on record values.* — Australian J. Statist. **19** (1977), 70–73.
4. H. N. Nagaraja, *On the expected values of record values.* — Australian J. Statist. **20** (1978), 176–182.
5. M. Ahsanullah, V. B. Nevzorov, *Characterizations of distributions by regressional properties of records.* — J. Appl. Statist. Science **13**, No. 1 (2004), 33–39.
6. I. Bairamov, M. Ahsanullah, A. Pakes, *A characterization of continuous distributions via regression on pairs of record values.* — Austr. N. Z. J. Statist. **47** (2005), 543–547.
7. M. Ahsanullah, R. C. Gupta, *Some characterization results based on the conditional expectation of a function of non-adjacent order statistics (record values).* — Ann. Inst. Statist. Math. (2004), 721–732.

Akhundov I. S., Nevzorov V. B. Conditional distributions of record values and characterizations of distributions.

Some new characterizations of distributions by regressional properties of sums of record values are obtained.

McMaster universily,
Hamilton, Canada

Поступило 17 ноября 2006 г.

С.-Петербургский
государственный университет

E-mail: valery.nevzorov@pobox.spbu.ru