Санкт-Петербургский государственный университет

Фундаментальная математика и механика

Теория вероятностей и математическая статистика

Обуховская Олеся Олеговна

Рекорды и рекордные величины в последовательности неодинаково распределенных случайных величин

Дипломная работа

Допущена к защите.

Зав. кафедрой, проф.

Я. Ю. Никитин

Научный руководитель:

докт. физ-мат.наук, проф.

В. Б. Невзоров

Рецензент:

к.ф.-м.н., доцент

В.Н. Солев

Санкт-Петербург

2017

SAINT-PETERSBURG STATE UNIVERSITY

Fundamental mathematics and mechanics

Theory of probability and mathematical statistics

Obukhovskaia Olesia Olegovna

Records and record values in a sequence of randomly distributed random variables

Graduation Thesis

Approved:

Head of the chair

Doctor of Physics and Mathematics, Professor

Jakov Nikitin

Scientific supervisor:

Doctor of Physics and Mathematics, Professor

Valerij Nevzorov

Reviewer:

Ph.D., doc.

Valentin Solev

Saint-Petersburg

2017

Содержание

Обозначения и определения

$F(x) = \mathrm{P}\{X \le x\}$ обозначает функцию распределения случайной величины (с.в.) X;
Соответствующая $F(x)$ плотность распределения $f(x)$ обозначается с маленькой буквы;

Будем говорить, что случайная величина X имеет

- *экспоненциальное распределение с параметром* λ (далее будем обозначать $\mathrm{E}(\lambda)$), если плотность распределения X представляется в виде:

  $f(x) = \frac{1}{\lambda}\exp\left(-\frac{x}{\lambda}\right)$, если $x \ge 0$, и $f(x) = 0$, иначе;

- *геометрическое распределение с параметром* p (далее будем обозначать Geo(p)), если распределение X представляется в виде:

  $\mathrm{P}\{X = n\} = p^{n-1}(1-p)$, где $n = 1, 2, \ldots$;

- *распределение Гумбеля*, если её функция распределения представляется в виде:

  $F(x) = \exp(-\exp(-x)),\ -\infty < x < \infty$;

- *стандартное нормальное распределение*, если её плотность представляется в виде:

  $f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right),\ -\infty < x < \infty$;

  $\Phi(x)$ обозначает функцию стандартного нормального распределения.

- *отрицательное биномиальное распределение с параметрами* p, r если распределение X представляется в виде:
  $\mathrm{P}\{X = n\} = C_{n+\mathrm{r}-1}^{n} p^{n-1}(1-p)^r$, где $n = 1, 2, \ldots$.

Буквы n, k, i, j будут использоваться для обозначения целых чисел.

Теория рекордов является в настоящее время активно развивающейся областью теории вероятностей и математической статистики. Существенный интерес представляет рассмотрение и изучение новых рекордных схем. Настоящая выпускная квалификационная работа посвящена рассмотрению некоторых нестационарных рекордных схем. Получен ряд новых результатов для таких рекордных моделей.

1 Введение

1.1 Классическая рекордная схема в случае непрерывных рекордных величин

Пусть $\{X_k,\ k \geq 1\}$ - независимые случайные величины (с.в.). Определим новые последовательности с.в.- верхние рекордные моменты $L(n)$ и верхние рекордные величины $X(n)$ следующим образом:

$$L(0)=0,\ \ L(1)=1, \tag{1.1.1}$$

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

$$X(n)= X_{L(n)},\quad n\geq 1.$$

Нижние рекордные моменты $l(n)$ и нижние рекордные величины $x(n)$ определяются аналогично, путем замены в (1.1.1) второго знака неравенства на противоположный. Далее будут рассматриваться только верхние рекордные моменты и рекордные величины (слово ”верхние” будем опускать), так как, если вместо исходной последовательности $\{X_k,\ k \geq 1\}$ взять $\{Y_k = -X_k,\ k \geq 1\}$, то все результаты для верхних рекордов естественным образом переписываются для нижних рекордов.

Далее определим k-рекордные моменты $L(n, k)$ и k-рекордные величины $X(n, k)$:

$$L(0, k) = 0,\quad L(1, k) = k, \tag{1.1.2}$$

$$L(n+1)=min\{j>L(n, k): X_j > X_{j-k, j-1}\},$$

$$X(n, k)=X_{L(n, k)-k+1},\quad n\geq 1,$$

где $X_{k,m}$ означает k-порядковую статистику, построенного по с.в. $X_1,..,X_m$, вариационного ряда:

$$X_{1,m}\leq X_{2,m}\leq \ldots \leq X_{k,m} \leq \ldots \leq X_{m,m}\ .$$

Заметим, что при k =1 определения (1.1.2) и (1.1.1) суть одно и то же.

Определим также межрекордные времена $\Delta(n, k)$:

$$\Delta(n, k)= L(n, k) - L(n-1, k)\ ,\ n\geq 1, \qquad (1.1.3)$$

Также для k-рекордов определим рекордные индикаторы $\zeta_n(k)$ следующим образом:

$$\zeta_n(k) = \begin{cases} 1, & \text{если } X_n > X_{n-k,n-1} \\ 0, & \text{иначе} \end{cases},\ \ n\geq k. \qquad (1.1.4)$$

Другими словами, $\zeta_n(k)$ =1, если найдется такое m, для которого $X_n= X(m, k)$.

Для индикаторов k-рекордов сформулируем следующий важный результат.

Теорема 1.1.1 (Renyi (1962)). Для независимых с.в. $\{X_n,\ n\geq 1\}$, имеющих общую непрерывную функцию распределения F, и для любого натурального k индикаторы $\{\varsigma_n(k),\ n\geq k\}$ независимы и верно равенство:

$$P\{\zeta_n(k) = 1\}= \frac{k}{n},\ \ k\geq 1, \quad n\geq k. \qquad (1.1.5)$$

Из теоремы 1.1 следует, что $L(1, k)$, $L(2, k)$, … образуют цепь Маркова с переходными вероятностями вида:

$$P\{L(n+1, k)=s \mid L(n, k)=m\}=\frac{(m+1-k)\cdot\ldots\cdot(s-k-1)k}{(m+1)\cdot\ldots\cdot(s-1)s}. \qquad (1.1.6)$$

Совместное распределение $L(1, k)$, …, $L(n, k)$ представляется выражением:

$$P\{L(1, k)=k, L(2, k)=m(2), \ldots, L(n, k)= m(n)\}=\frac{k^{n-1}}{C^k_{m(n)}\prod_{i=2}^{n}(m(i)-k)}, \qquad (1.1.7)$$

где $C^k_{m(n)} = \dfrac{m(n)!}{k!(m(n)-k)!}$.

Наряду с рекордными моментами, введем в рассмотрение с.в. $N(n, k)$, обозначающие число рекордов в последовательности $X_1, X_2, \ldots$. Связь между $N(m, k)$ и $L(n, k)$ демонстрируют соотношения:

$$\mathrm{P}\{L(n, k) > m\} = \mathrm{P}\{N(m, k) < n\}, \qquad (1.1.8)$$

$$\mathrm{P}\{L(n, k) = m\} = \mathrm{P}\{N(m\text{-}1, k) = n\text{-}1,\ N(m, k) = n\}. \qquad (1.1.9)$$

Имеет место представление $N(m, k)$ в виде суммы независимых индикаторов:

$$N(m, k) = \zeta_k(k) + \ldots + \zeta_m(k). \qquad (1.1.10)$$

При помощи ЦПТ для независимых и ограниченных с.в. из (1.1.10), применяя стандартную технику, можно получить асимптотическое соотношение следующего вида:

$$\sup_x \left| \mathrm{P}\left\{ N(n,k) - k\log(n) < x\left(k\log(n)\right)^{1/2} \right\} - \Phi(x) \right| \le \frac{C_1(k)}{(\log n)^{1/2}}, \qquad n \to \infty, \qquad (1.1.11)$$

где $C_1(k)$ - зависящая от k константа.

Воспользовавшись соотношением (1.1.8) , из (1.1.11) получаем асимптотическую нормальность случайных величин $T(n) = \log L(n,k)$:

$$\sup_x \left| \mathrm{P}\left\{ k\log L(n,k) - n < x\sqrt{n} \right\} - \Phi(x) \right| \le \frac{C_2(k)}{\sqrt{n}}, \qquad n \to \infty. \qquad (1.1.12)$$

Теперь рассмотрим в сравнении моментные характеристики рекордов и k-рекордов.

В случае обычных рекордов из теоремы 1.1.1 получаем:

$$\mathrm{P}\{L(2) = i\} = \mathrm{P}\{L(2) > i\text{-}1\} - \mathrm{P}\{L(2) > i\} = \frac{1}{i-1} - \frac{1}{i} = \frac{1}{i(i-1)},\ i \ge 1. \qquad (1.1.13)$$

Откуда следует, что $\mathrm{E}L(2) = \infty$ и тем более $\mathrm{E}L(n) = \infty$.

В случае же k-рекордов справедливо следующее соотношение:

$$\mathrm{E}L(n,k) = \frac{k^n}{(k-1)^{n-1}},\ n \ge 1,\ k \ge 2, \qquad (1.1.14)$$

$$\mathrm{D}L(n,k)=\frac{k^n(k-1)}{(k-2)^{n-1}}+\frac{k^n}{(k-1)^{n-1}}-\frac{k^{2n}}{(k-1)^{2n-2}},\ n\geq 1,\ k\geq 3. \qquad (1.1.15)$$

Из (1.1.14) следует, что $\mathrm{E}L(n, k) < \infty$ при $k \geq 2$, а из (1.1.15) - $\mathrm{D}L(n, k) < \infty$ при $k \geq 3$. Это важное свойство k-рекордных моментов дает преимущество их использования в статистических критериях.

Сформулируем несколько результатов, относящихся к распределению рекордных величин, построенных по последовательности независимых случайных величин с общей функцией распределения.

Представление 1. Пусть $\{X_n,\ n\geq 1\}$ - с.в., имеющие общую непрерывную функцию распределения F, $\{Z_n,\ n\geq 1\}$- E(1) - распределенные с.в.. Тогда для любых $k\geq 1$, $n \geq 1$ верно следующее равенство:

$$(X(1, k), \ldots , X(n, k)) \stackrel{d}{=} (H(Z(1, k)), \ldots , H(Z(n, k))) , \qquad (1.1.16)$$

где $H(x)= F^{-1}(1- \exp(-x))$, а F^{-1} - функция, обратная к функции распределения F, которая определяется следующим образом: $F^{-1}(y)=\inf\{x : F(x)\geq y\}$. Знак „ $\stackrel{d}{=}$ " означает равенство по распределению.

Представление 2. Пусть $\{Z_n,\ n\geq 1\}$- E(1)-распределенные с.в.. Тогда для любых $k\geq 1$, $n \geq 1$верно, что

$$(Z(1, k), \ldots , Z(n, k)) \stackrel{d}{=} \left(\frac{\xi_1}{k},\frac{\xi_1+\xi_2}{k},\ldots,\frac{\xi_1+\xi_2+\ldots+\xi_n}{k}\right), \qquad (1.1.17)$$

где $\xi_1,\xi_2,\ldots$ - независимые E(1)-распределенные с.в..

Как следствие (1.1.17) получаем, что в условиях представления 1 верно равенство:

$$(X(1, k), \ldots , X(n, k)) \stackrel{d}{=} \left(H\left(\frac{\xi_1}{k}\right),H\left(\frac{\xi_1+\xi_2}{k},\right)\ldots,H\left(\frac{\xi_1+\xi_2+\ldots+\xi_n}{k}\right)\right). \qquad (1.1.18)$$

Применяя данное представление, можно доказать, что рекордные величины $X(1, k)$, $X(2, k)$, … образуют цепь Маркова и верно равенство:

$$\mathrm{P}\{\ X(n+1,k)\mid X(n,k)=y\}=\left(\frac{1-F(x)}{1-F(y)}\right)^k . \qquad (1.1.19)$$

Из (1.1.18) также следует выражение, связывающее обычные и k-рекордные величины:

$$(X(1,k),\ \ldots,\ X(n,k)) \overset{d}{=} (Y_1(1),\ \ldots,\ Y_n(n)), \quad k\geq 1,\ n\geq 1,$$

где с.в. Y_1, Y_2,… определяются следующим образом:

$$Y_1=min(X_1,\ \ldots,\ X_k),\ Y_2=min(X_{k+1},\ \ldots,\ X_{2k}),\ldots \quad .$$

В дальнейшем нам понадобится также следующее представление для порядковых статистик $Z_{1,n} \leq \ldots \leq Z_{n,n}$, построенных по независимым E(1)-распределенным с.в. $Z_1, \ldots, Z_n$.

Представление 3.

$$(Z_{1,n},\ \ldots,\ Z_{n,n}) \overset{d}{=} \left(\frac{\xi_1}{n}, \frac{\xi_1}{n}+\frac{\xi_2}{n-1}, \ldots, \frac{\xi_1}{n}+\frac{\xi_2}{n-1}+\xi_n\right), \qquad n\geq 1 . \qquad (1.1.20)$$

Приведем также некоторые важные результаты для рекордных величин в случае дискретных распределений. Не умаляя общности, будем считать, что распределение задано на множестве натуральных чисел. Также предположим, что распределение не имеет такой точки x, для которой выполнялось бы следующее:

$$\mathrm{P}\{X<x\}<\mathrm{P}\{X\leq x\}=1. \qquad (1.1.21)$$

Условие (1.1.21) необходимо для того, чтобы с вероятностью единица для любого n были определены $X(n)$.

Определим рекордные индикаторы η_n, принимающие значение равное единице, если существует такое m, для которого $X(m)=X_n$. Сформулируем следующий результат.

Теорема 1.1.2(Shorrock(1972)). Рекордные индикаторы $\{\eta_n,\ n\geq 1\}$ независимы и справедливо равенство:

$$\mathrm{P}\{\eta_n=1\}=|\ \mathrm{P}\{X=n\mid X\geq n\}=\frac{\mathrm{P}\{X=n\}}{\mathrm{P}\{X\geq n\}} . \qquad (1.1.22)$$

Эта теорема позволяет представить распределение $X(n)$ в виде суммы независимых индикаторов и, применяя ЦПТ, получить асимптотические соотношения для $X(n)$.
Ограничимся здесь ещё одним представлением.

Представление (Shorrock). Пусть $\{Y_n,\ n \geq 1\}$- Geo(1)-распределенные с.в.. Тогда для любых $k\geq 1$, $n \geq 1$верно, что

$$Y(n) \overset{d}{=} \vartheta_1 + \vartheta_2 + \ldots + \vartheta_n,$$

где $\vartheta_n,\ n \geq 1$ - Geo(1)-распределенные независимые с.в..

Рассмотрим также предельное поведение рекордных величин. Сформулируем следующий важный результат.

Теорема 1.1.3. Предельная функция T распределения надлежащим образом центрированных и нормированных рекордных величин $X(n)$ имеет следующий вид:

$$T(x)= \Phi(g_k(x)), \qquad (1.1.23)$$

где k=1,2,3 и $g_k(x)$ одна из функций следующих трех типов:

$g_1(x)=x$,

$g_2(x)=\gamma \log x$, $\gamma>0$, если $x>0$, и $g_2(x)=-\infty$, если $x<0$,

$g_3(x)=-\gamma \log(-x)$, $\gamma>0$, если $x<0$, и $g_3(x)=\infty$, если $x>0$.

1.2 Неклассические рекордные схемы

Выше мы рассмотрели классическую схему, в рамках которой изучаются рекордные значения в последовательностях независимых и одинаково распределенных с.в., и изложили некоторые важные результаты для неё. Теперь перейдем к рассмотрению неклассических схем.

Первая нестационарная рекордная модель была предложена Йенгом (Yang, (1975)). В своей модели Йенг рассматривал рекорды среди максимальных значений в группах независимых одинаково распределенных с.в.. Т.е. в последовательности случайных величин, определенных следующим образом:

$$Y_k = \max\{X_{k,1}, \ldots, X_{k,n(k)}\} ,$$

где для любых $k \geq 1$, $n(k) \geq 1$ с.в. $X_{k,1}, \ldots, X_{k,n(k)}$ независимы и имеют общую непрерывную функцию распределения F.

Перейдем к обобщению модели Йенга - так называемой F^{α} - схеме.

Определение F^{α}- схемы. Последовательность независимых с.в. $\{X_n,\ n \geq 1\}$ образует F^{α}- схему, если существует непрерывная функция распределения F и положительные константы $\alpha(1)$, $\alpha(2)$,…, такие что для любого $k \geq 1$функция распределения с.в. X_k представляется в виде $F_k(x) = F^{\alpha(k)}(x)$.

Заметим, что в модели Йенга с.в. $\{Y_n,\ n \geq 1\}$ образуют F^{α}-схему, в которой $\alpha(k)= n(k)$ целые числа.

Так же, как и в классической схеме, рассмотрим рекордные индикаторы ζ_1, ζ_2, ... , принимающие значение равное единице, если существует такое m, для которого $X(m)=X_n$. Сформулируем результат, обобщающий теорему 1.1.1.

Теорема 1.2.1. В F^{α}- схеме рекордные индикаторы ζ_1, ζ_2, ... независимы и верно равенство:

$$p_n = P\{ \zeta_n=1\}=\frac{\alpha(n)}{S(n)} \text{ , где } S(n)= \alpha(1)+\ldots+ \alpha(n) \text{ , } n\geq 1, \qquad (1.2.1)$$

Из теоремы 1.2.1 следует, что, как и в случае классической схемы, число рекордов $N(n)$ представляется в виде суммы независимых индикаторов:

$$N(n)= \zeta_1+\ldots+ \zeta_n \text{ , } n\geq 1, \qquad (1.2.2)$$

Обозначим

$$A(n)= EN(n)= \sum_{j=1}^{n} p_j . \qquad (1.2.3)$$

Из курса анализа известно, что последовательности $N(n)$ *и* $A(n)$ при $n \to \infty$ одновременно имеют или не имеют конечный предел. Из этого факта, а также из равенства, связывающего рекордные моменты и число рекордов:

$$\mathrm{P}\{\, L(n) \le \, m\}=\mathrm{P}\{\, N(n)\ge n\},$$

следует, что любой рекордный момент существует тогда и только тогда, когда $S(n)\to\infty$. Поэтому естественно в F^{α}- схеме полагать ограничение на коэффициенты $\alpha(n)$, при котором $S(n)\to\infty$.

В случае классической схемы $S(n)= n$, поэтому это условие обеспечивается автоматически.

Независимость рекордных индикаторов позволяет обобщить на F^{α}- схему многие результаты, характерные для классической схемы. Так получаем, что и в случае F^{α}- схемы рекордные моменты образуют цепь Маркова и верно следующее:

$$\mathrm{P}\{L(n) =j \mid L(n\text{-}1) =i\,\}=S(i)\left(\frac{1}{S(j-1)}-\frac{1}{S(j)}\right), \tag{1.2.4}$$

$$\mathrm{P}\{\, L(n) >j \mid L(n\text{-}1) =i\,\}=\frac{S(i)}{S(j)}\ . \tag{1.2.5}$$

Также из независимости рекордных индикаторов следует формула для совместного распределения рекордных моментов:

$$\mathrm{P}\{L(1) =1,\, L(2) =k(2),\ldots\, L(n) =k(n)\}=\frac{\prod_{i=2}^{n}\frac{\alpha(k(i))}{S(k(i)-1)}}{S(k(n))}\quad . \tag{1.2.6}$$

Отметим важное свойство F^{α}- схемы.

Теорема1. 2.2 (Ballerini and Resnick, 1987) В F^{α}- схеме рекордные индикаторы ζ_1 , ... , ζ_n и $M(n)= \max\{\, X_1, \ldots, X_n\}$ независимы для любого $n\ge1$.

Используя этот факт, с помощью формулы полной вероятности можно получить соотношение, связывающее рекордные моменты и рекордные величины:

$$\mathrm{P}\{X(n) < x\}=\mathrm{E}(F(x))^{S(L(n))}\ . \tag{1.2.7}$$

Приведем также результаты, касающиеся асимптотического распределения рекордных величин. Оказывается, что для F^{α}- схемы при нежестких ограничениях на $\alpha(k)$ предельные функции распределения исчерпываются теми же функциями, что и в случае классической модели.

Сформулируем здесь несколько важных результатов.

Теорема 1.2.3 Пусть в F^{α}- схеме при $n \to \infty$ выполнены условия:

1. $S(n) \to \infty$
2. $p_n = \alpha(n)/S(n) \to 0$,
3. $\dfrac{D(n)}{\sqrt{A(n)}} \to 0$

Тогда предельная функция распределения T для надлежащим образом центрированных и нормированных рекордных величин имеет вид, определенный в (1.1.23).

Сформулируем также предельную теорему для межрекордных времен.

Теорема 1.2.4 (Ballerini and Resnick, 1987). Если в F^{α}- схеме

$$\lim_{n\to\infty} \frac{\alpha(n)}{S(n)} = p, \quad 0 < p < 1,$$

то при $n \to \infty$ $\Delta(n)$ имеет геометрическое Geo(p)-распределение.

Рассмотрим ещё одну нестационарную модель – так называемую схему с линейным сносом.
Пусть $\{X_n,\ n \geq 1\}$ независимые одинаково распределенные с.в.. Будем рассматривать рекорды в последовательности $\{Y_n = X_n + c(n),\ n \geq 1\}$, где $c(n)$ - константы, зависящие от n. Логично предполагать, что $c \geq 0$, и наиболее интересной в этом случае является ситуация, когда $c(n) = cn$.

В данной модели для рекордных индикаторов, вероятности которых определяются соотношением:

$$p_n = \mathrm{P}\{\xi_n=1\} = 1 - \mathrm{P}\{\xi_n=0\} = \int_{-\infty}^{\infty} \prod_{j=1}^{n-1} F(x+cj)\,dF(x), \qquad (1.2.8)$$

не характерно свойство независимости, за исключением того случая, когда исходная последовательность с.в. имеет распределение Гумбеля (в этом частном случае $\{Y_n,\ n \geq 1\}$ образуют F^{α}- схему с показателем $\alpha = \exp(cn)$).

Т.к. из (1.2.8)следует, что последовательность p_n убывает, то существует предел $p = \lim_{n\to\infty} p_n$. Вычислению этого предела для определенных классов распределений были посвящены исследования Ballerini and Resnick (1985), Nagaraja (1994).

Также существенные различия с классической моделью наблюдаются в предельном поведении с.в. $N(n)$ – числа рекордов. Для классической модели был характерен логарифмический рост E $N(n)$, а в рамках данной модели верны следующие асимптотические соотношения (Ballerini and Resnick (1985)):

$$\mathrm{E}(N(n)/n) \to p \text{ and } E(N(n)/n - p)^2 \to 0, \quad n \to \infty. \tag{1.2.9}$$

2 Схема рекордных величин Пфайфера

2.1 Определения и свойства рекордных величин

В этом параграфе пойдет речь ещё об одной нестационарной модели, которую предложил рассматривать Пфайфер (Pfeifer, (1982, 1984)) .
Пусть имеется последовательность серий независимых с.в. $\{X_k^{(n)},\ k \ge 1,\ n \ge 1\}$, в которой для любого $n \ge 1$ с.в. $\{X_k^{(n)},\ k \ge 1\}$, принадлежащие одной серии, имеют общую функцию распределения F_n. Суть схемы состоит в том, что каждый раз при достижении нового рекорда осуществляется переход к новой серии случайных величин, имеющих другое распределение.

Введем в рассмотрение межрекордные времена $\Delta(n) = L(n) - L(n-1)$, определив их следующим образом:

$$\Delta(1)=1,\ \Delta(n+1)=\min\{j > L(n) \mid X_j^{(n+1)} > X_{\Delta(n)}^{(n)}\}, \quad n \ge 1, \tag{2.1.1}$$

т.е. $\Delta(n)$ обозначает число случайных величин в n-ой серии, приводящих к появлению нового рекорда.
Рекордные моменты $L(n)$ и рекордные величины $X(n)$ тогда выражаются следующим образом:

$$L(0) = 0,\ L(n) = \Delta(1) + \ldots + \Delta(n), \quad n \ge 1, \tag{2.1.2}$$
$$X(0) = 0,\ X(n) = X_{\Delta(n)}^{(n)}.$$

Из определения рекордной модели следует, что межрекордные времена $\Delta(1), \ldots, \Delta(n)$ условно независимы при фиксированных значениях $X(n)$, …, $X(n\text{-}1)$. При этом верно равенство:

$$\mathrm{P}\{\Delta(1)=1,\ \Delta(2)=m(2),\ \ldots,\ \Delta(n)=m(n)\ \mid X(1)=x_1,\ \ldots,\ X(n\text{-}1)=x_{n-1}\}= \qquad (2.1.3)$$
$$=\prod_{i=2}^{n}\left(1-F_i\left(x_{i-1}\right)\right)\left(F_i\left(x_{i-1}\right)\right)^{m(i)-1}\quad m(j)\geq 1,\ j=2,\ \ldots,\ n.$$

В данной модели, в отличие от классической схемы, рекордные моменты $\{L(n),\ n\geq 1\}$ не образуют цепь Маркова. Ряд полезных свойств, характерных для классической схемы, например, таких как теорема Реньи, не выполняется.

Однако, цепь Маркова образуют рекордные величины $\{X(n),\ n\geq 1\}$ и при этом формула переходной вероятности имеет следующий вид:

$$\mathrm{P}\{\ X(n)>x_n \mid X(n\text{-}1)=x_{n-1}\}=\frac{1-F_n\left(x_n\right)}{1-F_n\left(x_{n-1}\right)},\quad x_n>x_{n-1}\ . \qquad (2.1.4)$$

Пфайфером было показано, что последовательность векторов $(\Delta(n),\ X(n))$ также обладает Марковским свойством. При этом верно, что

$$\mathrm{P}\{\ \Delta(n)=m_n,\ X(n)>x_n \mid \Delta(n-1)=m_{n-1}\ ,\ X(n\text{-}1)=x_{n-1}\}= \qquad (2.1.5)$$
$$\left(1-F_n\left(x_n\right)\right)\left(F_n\left(x_{n-1}\right)\right)^{m_n-1},\quad x_n>x_{n-1}.$$

Вектора $(L(n),\ X(n))$ также образуют цепь Маркова.

Приведем здесь один результат, доказанный Шорроком для классической рекордной схемы.

Лемма 2.1 (Shorrock, (1972)). Пусть $\{X_n\ ,\ n\geq 1\}$ - последовательность независимых Geo(p_n) - распределенных с.в.. Тогда, если ряд $\sum_{n=1}^{\infty} p_n$ сходится, верно следующее равенство:

$$\overline{\lim_{n\to\infty}}\left|\frac{\log X_n p_n}{\log n}\right| \overset{п.н.}{=} 1$$

Тот факт, что согласно (2.1.3) для любого n распределение $\Delta(n)$ при фиксированном значении $X(n\text{-}1)=x_{n-1}$ имеет Geo(p_n) распределение, где $p_n=\ F_n\left(x_{n-1}\right)$, позволил Пфайферу обобщить лемму Шоррока и получить следующую теорему.

Теорема 2.1 (Pfeifer, (1984)). Пусть в модели Пфайфера рассматриваются с.в. $\{X_k^{(n)},\ k \ge 1,\ n \ge 1\}$, функции распределения F_n которых непрерывны и не убывают по n. Тогда верно следующее равенство:

$$\overline{\lim_{n\to\infty}} \left| \frac{\log \Delta(n) - H_n(X(n-1))}{\log n} \right| \overset{п.н.}{=} 1, \qquad (2.1.6)$$

где $H_n(X) = -\log(1-F(X))$.

Приведем также еще один результат, полученный Пфайфером.

Теорема 2.2 (Pfeifer, (1984)). Пусть в модели Пфайфера функции распределения F_n с.в. $\{X_k^{(n)},\ k \ge 1,\ n \ge 1\}$ имеют вид $1 - F_n(x) = (1-F(x))^{\lambda_n}$, где $\lambda_n = n + \mathrm{C}$,
С – фиксированное натуральное число. Тогда верно следующее:

$$\frac{\log \Delta(n)}{n} - \log n \xrightarrow{d} x_G(C-1), \text{ где}$$

$x_G(n)$ обозначает n-ую нижнюю рекордную величину в последовательности с.в. $\{W_n, n \ge 1\}$, имеющих распределение Гумбеля.

Из этой теоремы видно, что асимптотическая нормальность $\log \Delta(n)$, характерная для классической модели, не всегда присуща модели Пфайфера.

2.2 *k*-рекорды в схеме Пфайфера

В данном параграфе приведем ряд новых результатов для модели Пфайфера. Пусть теперь переход к новой серии случайных величин будет осуществляться лишь после того, как k величин превзойдут предыдущий рекорд, а за новую рекордную величину будет полагаться меньшая из них. В этом случае межрекордные времена $\Delta(n,k)$, k- рекордные моменты и $L(n, k)$ и k- рекордные величины $X(n, k)$ определяются следующим образом:

$$\Delta(1,k) = k,\ \Delta(n+1,k) = \min\left\{ j : X_{j-k+1,j}^{(n+1)} > X_{\Delta(n,k)-k+1,\Delta(n,k)}^{(n)} \right\}, \qquad (2.2.1)$$

$$L(0,k) = 0,\ L(n,k) = \Delta(1,k) + \ldots + \Delta(n,k),$$

$$X(0,k)=0,\ X(n,k)=X^{(n)}_{\Delta(n,k)-k+1,\Delta(n,k)},\ n\geq 1.$$

Рассмотрим свойства данной рекордной модели.
Совместное распределение межрекордных времен $\Delta(1,k),\ \ldots,\Delta(n,k)$ и рекордных величин $X(1,k),\ldots,X(n,k)$ будет иметь следующий вид:

$$\mathrm{P}\{X(1,k)<x_1,\ \ldots,X(n,k)<x_n,\ \Delta(1,k)=k,\ \Delta(2,k)=m(2),\ \ldots,\ \Delta(n,k)=m(n)\}=$$

$$\mathrm{P}\{X^{(1)}_{1,k}<x_1,\ X^{(2)}_{m(2)-k,m(2)}\leq X^{(1)}_{1,k}<X^{(2)}_{m(2)-k+1,m(2)}<x_2,\ \ldots,$$

$$X^{(n)}_{m(n)-k,m(n)}\leq X^{(n-1)}_{m(n-1)-k+1,m(n-1)}<X^{(n)}_{m(n)-k+1,m(n)}<x_n\}=$$

$$=\int_{-\infty}^{x_n}\ldots\int_{-\infty}^{x_1}\psi(v_1,\ldots,v_n)f^{(1)}_{1,k}(v_1)dv_1\ldots f^{(n)}_{1,k}(v_n)dv_n,\quad n\geq 2,\qquad (2.2.2)$$

где $f^{(i)}_{1,k}$ обозначает плотность распределения $\min\{X^{(i)}_1,\ldots,X^{(i)}_k\}$ и задается формулой $f^{(i)}_{1,k}=k(1-F_i(x))^{k-1}f_i(x)$,

$\psi(v_1,\ldots,v_n)=\prod_{i=2}^{n}C^{k-1}_{m(i)-1}F_i^{m(i)-k}(v_{i-1})$, если $-\infty<v_1<\ldots<v_n<\infty$ и $\psi(v_1,\ldots,v_n)=0$, иначе.

Дифференцированием (2.2.2) по переменным $x_1,\ldots,x_n$ получаем функцию $f_{(\Delta(1,k),X(1,k),\ldots,\Delta(n,k),X(n,k))}$, которая представляет плотность по отношению к рекордным величинам $X(1,k),\ldots,X(n,k)$ и распределение вероятностей рекордных моментов $\Delta(1,k),\ \ldots,\Delta(n,k)$:

$$f_{(\Delta(1,k),X(1,k),\ldots,\Delta(n,k),X(n,k))}(s(1),\ldots,s(n),x_1,\ldots,x_n)=\psi(x_1,\ldots,x_n)f^{(1)}_{1,k}(x_1)\cdot\ldots\cdot f^{(n)}_{1,k}(x_n).\qquad (2.2.3)$$

Найдем распределение $X(n,k)$ при условии, что фиксированы значения $X(1,k)=x_1,\ldots,X(n\text{-}1,k)=x_{n-1}$. Воспользуемся тем фактом, что $\{Y_i=X_{L(n-1)+i},\ i\geq 1\}$ не зависят от $X(1,k),\ldots,X(n\text{-}1,k)$ и независимы между собой. Обозначим за $Y_{\tau(u)}$ - первую величину в последовательности $\{Y_i,\ i\geq 1\}$, которая превышает u. Т.е. рекордную величину можно представить в новых обозначениях как $X(n,k)=Y_{\tau(X(n-1))}$. Верно следующее равенство:

$$\mathrm{P}\{X(n)>x_n\mid X(1)=x_1,\ldots,X(n-1)=x_{n-1}\}=\qquad (2.2.4)$$

$$\mathrm{P}\{X_{L(n-1)+\tau(X(n-1))}>x_n\mid X(1)=x_1,\ldots,X(n-1)=x_{n-1}\}=\mathrm{P}\{X_{L(n-1)+\tau(x_{n-1})}>x_n\}=\mathrm{P}\{Y_{\tau(x_{n-1})}>x_n\}=$$

$$=\mathrm{P}\{Y^{(n)}_{1,k}>x\}+\mathrm{P}\{Y^{(n)}_{1,k+1}\leq x_{n-1},Y^{(n)}_{2,k+1}>x\}+\mathrm{P}\{Y^{(n)}_{2,k+2}\leq x_{n-1},Y^{(n)}_{3,k+2}>x\}+\ldots=$$

$$= \left(1 - F_n(x)\right)^k + C_k^1 F_n(x_{n-1})\left(1 - F_n(x)\right)^k + C_{k+1}^2 \left(F_n(x_{n-1})\right)^2 \left(1 - F_n(x)\right)^k + \ldots =$$

$$= \left(1 - F_n(x)\right)^k \sum_{i=0}^{\infty} C_{k+i-1}^i \left(F_n(x_{n-1})\right)^i = \frac{\left(1 - F_n(x)\right)^k}{\left(1 - F_n(x_{n-1})\right)^k}, \quad x > x_{n-1} .$$

Здесь мы воспользовались следующей формулой:

$$\sum_{i=0}^{\infty} C_{k+i-1}^i x^i = (1-x)^{-k} .$$

Из (2.2.4) также следует, что рекордные величины $\{X(n, k), n \geq 1\}$ образуют цепь Маркова и условная плотность $X(n, k)$ при фиксированном значении $X(n\text{-}1, k) = x_{n-1}$ имеет вид:

$$f_{(X(n,k)|\ x_1,\ldots,x_{n-1})}(x) = \frac{k f_n(x)\left(1 - F_n(x)\right)^{k-1}}{\left(1 - F_n(x_{n-1})\right)^k} , \quad x > x_{n-1} . \tag{2.2.5}$$

Напишем также выражение для плотности совместного распределения рекордных величин $X(1, k), \ldots , X(n, k)$:

$$f_{(X(1,k),\ldots,\ X(n,k))}(x_1,\ldots,x_n) = f_{(X(n,k)|\ x_1,\ldots,x_{n-1})}(x_n) f_{(X(1,k),\ldots,\ X(n-1,k))}(x_1,\ldots,x_{n-1}) = \ldots = \tag{2.2.6}$$

$$= k^n f_1(x_1)\left(1 - F_1(x_1)\right)^{k-1} \prod_{i=2}^{n} W_i(x_i),$$

где $W_i(x_i) = \dfrac{f_i(x_i)\left(1 - F_i(x_i)\right)^{k-1}}{\left(1 - F_i(x_{i-1})\right)^k}$.

Из (2.2.6) получаем, что функция распределения $X(n, k)$ представляется в виде:

$$F_{X(n,k)}(x_n) = \int \ldots \int f_{(X(1,k),\ldots,\ X(n,k))}(v_1,\ldots,v_n) dv_1 \cdot \ldots \cdot dv_n , \tag{2.2.7}$$

где область интегрирования задается неравенствами: $-\infty < v_1 < \ldots < v_n < x_n < \infty$.

Замечание 1. Получаем, что в схеме Пфайфера, так же как и в классической рекордной модели, имеет место равенство:

$$X(n, k) \overset{d}{=} Y(n),$$

где $X(n, k)$ – k-рекорды в модели Пфайфера $\left\{X_m^{(n)},\ m \geq 1,\ n \geq 1\right\}$и F_n - функция распределения с.в.$\left\{X_m^{(n)},\ m \geq 1\right\}$,

а $Y(n)$ – рекордные величины в схеме Пфайфера $\left\{Y_m^{(n)},\ m \geq 1,\ n \geq 1\right\}$,

определяемые следующим образом:

$$Y_m^{(n)} = \min\left\{X_{k(m-1)+1}^{(n)},\ldots,X_{km}^{(n)}\right\},$$

при этом функцией распределения $\left\{Y_m^{(n)},\ m\ge 1\right\}$ является $G_n = 1-\left(1-F_n\right)^k$.

Теперь с помощью (2.2.3) и (2.2.5) найдем распределение межрекордных времен при условии, что фиксированы значения рекордных величин:

$$\mathrm{P}\{\Delta(2,k)=s(2),\ldots,\Delta(n,k)=s(n)\mid X(1,k)=x_1,\ldots,X(n,k)=x_n\}= \tag{2.2.8}$$

$$=\prod_{i=2}^{n} C_{s(i)-1}^{k-1} F_i\left(x_{i-1}\right)^{s(i)-k}\left(1-F_i\left(x_{i-1}\right)\right)^k$$

Из (2.2.8) видно, что межрекордные времена условно независимы и верно равенство:

$$\mathrm{P}\{\Delta(n,k)=s(n)\mid X(n-1,k)=x_{n-1}\}=C_{s(n)-1}^{k-1}F_n\left(x_{n-1}\right)^{s(n)-k}\left(1-F_n\left(x_{n-1}\right)\right)^k\ . \tag{2.2.9}$$

Заметим, что распределение $\Delta(n,k)$ при фиксированном значении $X(n-1,k)=x_{n-1}$ представляет собой отрицательное биномиальное распределение с параметром p=$F_n\left(x_{n-1}\right)$.

Теперь рассмотрим важный частный случай, когда $\left\{Z_m^{(n)},\ m\ge 1,\ n\ge 1\right\}$ являются $\mathrm{E}(\lambda_m)$-распределенными независимыми случайными величинами.

Будем интересоваться представлением распределения Z (n, k). Воспользуемся представлением для порядковых статистик экспоненциально распределенных случайных величин (1.1.20). Из указанного представления ясно, что первая рекордная величина распределена как:

$$Z(1,k)=Z_{1,n}=\frac{\xi_1\lambda_1}{k}\ .$$

Напомним, что $\xi_i - \mathrm{E}(1)$распределенные с.в. , $i\ge 1$.

Пусть Z (n, k) принимает значение z_n . Найдем представление для Z (n+1, k).

Выделим из последовательности $\left\{Z_m^{(n+1)},\ m\ge 1\right\}$ подпоследовательность $\left\{Z_{m_j}^{(n+1)} : Z_{m_j}^{(n+1)} > z_n\right\}$.

Согласно (2.2.4) имеем:

$$\mathrm{P}\{Z(n+1,k)>x\}=e^{-(x-z_n)k/\lambda_{n+1}}\ ,\ x>z_n\ ,\quad \text{то есть } Z(n+1,k)\overset{d}{=}Z_n+\frac{\xi_{n+1}\lambda_{n+1}}{k}.$$

Откуда следует, что

$$\mathrm{Z}\,(n{+}1,\,k) \overset{d}{=} \mathrm{Z}\,(n,\,k) + \frac{\xi_n\lambda_n}{k}$$

Таким образом, получаем следующее представление для экспоненциальных величин:

$$\mathrm{Z}\,(n,\,k) \overset{d}{=} \frac{1}{k}\sum_{i=1}^{n}\xi_i\lambda_i\,. \qquad (2.2.10)$$

Также можно было рассматривать и более общую ситуацию, когда значение k не фиксировано, а представляет собой функцию k=k(n) от номера серии n. В этой ситуации представление(2.2.10) для экспоненциально распределенных величин перепишется в виде:

$$\mathrm{Z}\,(n,\,k) \overset{d}{=} \sum_{i=1}^{n}\frac{\xi_i\lambda_i}{k(i)}\,.$$

Из замечания 1 следует, что Z (n, k) $\overset{d}{=} \widetilde{Z}(n)$, где $\left\{Z_m^{(i)},\ m\geq 1\right\}$ - $\mathrm{E}(\lambda_i)$-распределенные с.в., а $\left\{\widetilde{Z}_m^{(i)},\ m\geq 1\right\}$-$\mathrm{E}\left(\widetilde{\lambda}_i\right)$-распределенные с.в., $\widetilde{\lambda}_i=\lambda_i/k(i)$.

В схеме Пфайфера параметры λ_i можно интерпретировать, как параметры изменения системы, а $k(i)$ - параметры управления этой системой.

Далее будем интересоваться предельным поведением Z (n)= $\sum\limits_{i=1}^{n}\xi_i\lambda_i$.

Из представления распределения Z (n) в виде суммы независимых с.в. (в (2.2.10) полагаем k=1) сразу получаем выражения для математического ожидания и дисперсии:

$$\mathrm{E}Z(n)=\sum_{i=1}^{n}\lambda_i\,, \qquad \mathrm{D}Z(n)=\sum_{i=1}^{n}\lambda_i^{\,2}\,. \qquad (2.2.11)$$

Из ЦПТ для независимых и ограниченных с.в. имеем следующее:

$$\frac{Z(n)-\mathrm{E}Z(n)}{\sqrt{\mathrm{D}Z(n)}}\xrightarrow{\ d\ }N(0,1)\ \text{при}\ \ n\to\infty, \qquad (2.2.12)$$

если ряд $\sum\limits_{i=1}^{n}\lambda_i^{\,2}$ расходится.

Если ряд $\sum\limits_{i=1}^{n}\lambda_i^{\,2}$ сходится, то существует такая с.в. V, что

$$Z(n)-\mathrm{E}Z(n)\xrightarrow{\ п.н.\ }V\,. \qquad (2.2.13)$$

В случае, когда $\{Z_m^{(i)},\ m \ge 1\}$ - $\mathrm{E}(\lambda_i)$- распределенные с.в. , функция H_n в условии теоремы 2.1 имеет вид: $H_n(x) = \lambda_n x$. Теорема для экспоненциально распределенных с.в. переписывается следующим образом:

$$\overline{\lim_{n\to\infty}} \left| \frac{\log \Delta(n) - \lambda_n Z(n-1)}{\log n} \right| \overset{п.н.}{=} 1 \quad \text{или} \quad \overline{\lim_{n\to\infty}} \frac{\lambda_n}{\log n} \left| \frac{\log \Delta(n)}{\lambda_n} - Z(n-1) \right| \overset{п.н.}{=} 1. \qquad (2.2.14)$$

Если ряд $\sum_{i=1}^{n} \lambda_i^{\,2}$ расходится и при этом также $\dfrac{\lambda_n}{\log n} \to \infty$, то тогда из соотношений (2.2.12) и (2.2.14) следует асимптотическая нормальность центрированных и нормированных с.в. $W_n = \dfrac{\log \Delta(n)}{\lambda_n}$.

Мы рассмотрели рекордную модель, в которой в каждой серии за рекордную величину полагалась наименьшая из k с.в., превосходящих предыдущий рекорд. Если бы мы принимали за рекорд максимальную из этих случайных величин, то тогда получили бы обобщение рассмотренной В.Б. Невзоровым схемы так называемых рекордов с подтверждениями (2008). Рекордные моменты $\hat{L}(n,k)$ и рекордные величины $\hat{X}(n,k)$для данной серии определяются соответственно:

$$\Delta(1,k)=k,\ \Delta(n+1,k) = \min\left\{ j : X_{j-k+1,j}^{(n+1)} > X_{\Delta(n,k),\Delta(n,k)}^{(n)} \right\}, \qquad (2.2.15)$$

$$\hat{L}(0,k) = 0,\ \hat{L}(n,k) = \Delta(1,k) + \ldots + \Delta(n,k),$$

$$\hat{X}(0,k) = 0,\ \hat{X}(n,k) = X_{\Delta(n,k),\Delta(n,k)}^{(n)}\ .$$

В случае экспоненциально распределенных случайных величин, используя представление (1.1.20), аналогично (2.2.10) получаем следующее представление для рекордных величин $\hat{Z}(n,k)$:

$$\hat{Z}(n,k) \overset{d}{=} \sum_{i=1}^{n} \lambda_i \left(\frac{\xi_1^{(i)}}{k} + \frac{\xi_2^{(i)}}{k-1} + \ldots + \xi_k^{(i)} \right) = \sum_{i=1}^{n} \sum_{j=1}^{k} \lambda_i \frac{\xi_j^{(i)}}{k-j+1}\ . \qquad (2.2.16)$$

Из представления (2.2.14) сразу следуют формулы для математического ожидания и дисперсии $\hat{Z}(n,k)$:

$$\mathrm{E}\hat{Z}(n,k)=\sum_{i=1}^{n}\sum_{j=0}^{k-1}\lambda_i\frac{1}{k-j}\ ,\quad \mathrm{D}\hat{Z}(n,k)=\sum_{i=1}^{n}\sum_{j=0}^{k-1}\lambda_i^{\,2}\frac{1}{(k-j)^2}\quad . \tag{2.2.17}$$

Заметим, что если в этой схеме рассматривать k=k(n), то мы получим обобщение рекордной схемы Йенга.

В зависимости от рассматриваемой задачи за рекордную величину мы можем принимать любую порядковую статистику вариационного ряда, построенного по k превышающим предыдущий рекорд случайным величинам.

Например, можно было бы рассматривать схему, где за новый рекорд принимается не обязательно крайнее, а, например, среднее значение в вариационном ряду. В этом случае для экспоненциально распределенных случайных величин получили бы следующее представление:

$$\mathrm{Z}\ (n,\,k)\overset{d}{=}\sum_{i=1}^{[n/2]}\sum_{j=1}^{k}\lambda_i\frac{\xi_j^{(i)}}{k-j+1}\ . \tag{2.2.18}$$

2.3 Схема Пфайфера в случае дискретных распределений

Рассмотрим схему Пфайфера в случае дискретно распределенных с.в., принимающих с ненулевыми вероятностями целые неотрицательные значения.

Введем те же обозначения, что и при доказательстве марковского свойства рекордных величин(2.2.4). Вновь воспользовавшись тем, что $\{Y_i=X_{L(n-1)+i},\ i\ge 1\}$ не зависят от X(1, k),… , X(n-1, k) и независимы между собой, докажем аналогичный факт в случае дискретной модели:

$$\mathrm{P}\{X(n)=x_n\mid X(1)=x_1,\dots,X(n-1)=x_{n-1}\}=\mathrm{P}\{X_{L(n)+\tau(X(n-1))}=x_n\mid X(1)=x_1,\dots,X(n-1)=x_{n-1}\}=$$

$$\mathrm{P}\{X_{L(n)+\tau(x_{n-1})}=x_n\}=\mathrm{P}\{Y_{\tau(x_{n-1})}=x_n\}=\mathrm{P}\{Y_1=x_n\}+\mathrm{P}\{Y_1\le x_{n-1},Y_2=x_n\}+\ldots=$$

$$=\mathrm{P}\{X=x_n\}\left(1+F(x_{n-1})+F(x_{n-1})^2+\ldots\right)=\frac{\mathrm{P}\{X=x_n\}}{1-F(x_{n-1})},\ \ n\ge 2. \tag{2.3.1}$$

Из (2.3.1) сразу получаем выражение для совместного распределения рекордных величин в дискретном случае:

$$P\{X(1)=x_1,...,X(n)=x_n\}=P\{X=x_1\}\prod_{i=2}^{n}\frac{P\{X=x_i\}}{1-F(x_{i-1})}\ , n\geq 2. \qquad (2.3.2)$$

Откуда следует:

$$P\{X(n)=x_n\} \ = \sum_{x_1<x_2<...<x_{n-1}} P\{X=x_1\}\prod_{i=2}^{n}\frac{P\{X=x_i\}}{1-F(x_{i-1})}, \text{ где } x_i<x_n,\ i=1, ..., n-1., n\geq 2 \qquad (2.3.3)$$

Рассмотрим важный частный случай. Пусть $\{X_k^{(n)},\ k\geq 1,\ n\geq 1\}$ - Geo(p_n)-распределенные с параметрами $\{p_n,\ n\geq 1\}$ с.в..
Тогда формулы (2.3.1) и (2.3.2) перепишутся соответственно в виде:

$$P\{X(n)=x_n \mid X(1)=x_1,...,X(n-1)=x_{n-1}\}=\frac{p_n^{\,x_n-1}(1-p_n)}{p_n^{\,x_{n-1}}}=p_n^{\,(x_n-x_{n-1})-1}(1-p_n), \qquad (2.3.4)$$

$$P\{X(1)=x_1,...,X(n)=x_n\}=p_1^{\,x_1-1}(1-p_1)\prod_{i=2}^{n}p_i^{\,(x_i-x_{i-1})-1}(1-p_i). \qquad (2.3.5)$$

Заметим, что если от рассмотрения рекордных величин $\{X(n),\ n\geq 1\}$ перейти к величинам $\{V(1)=X(1),\ V(n)=X(n)-X(n-1),\ n\geq 2\}$, то соотношение (2.3.4) представится в виде произведения независимых геометрически распределенных с параметрами $\{p_n,\ n\geq 1\}$ случайных величин:

$$P\{V(1)=y_1,V(2)=y_2,...,X(n)=y_n\}=p_1^{\,y_1-1}(1-p_1)\prod_{i=2}^{n}p_i^{\,y_i-1}(1-p_i). \qquad (2.3.6)$$

Из (2.3.5) получаем представление для рекордной величины $X(n)$:

$$X(n) \overset{d}{=} \sum_{i}^{n}V(i), \qquad (2.3.7)$$

где $\{V(n),\ n\geq 1\}$ -геометрически распределенные с.в. с параметрами $\{p_n,\ n\geq 1\}$.
Заметим, что в частном случае, когда $p_1=...=p_n$, в (2.3.7) получаем представление Шоррока для классической рекордной модели.

Из ЦПТ для независимых и ограниченных с.в. следует, что

$$\frac{X(n)-EX(n)}{\sqrt{DX(n)}}\xrightarrow{d}N(0,1) \text{ при } n\to\infty, \qquad (2.3.8)$$

где $\quad \mathrm{E}X(n)=\sum_{i=1}^{n}\frac{1}{1-p_i}, \quad \mathrm{D}Z(n)=\sum_{i=1}^{n}\frac{p_i}{(1-p_i)^2}$.

3 Рекорды с ограничениями

3.1 Определения и свойства рекордных величин

Рассмотрим последовательность независимых одинаково распределенных с.в. X_1, X_2 ... Также фиксируем $C>0$. Определим рекордные моменты $L(n)$ и рекордные величины $X(n)$ следующим образом:

$$L(0)=0,\ L(0)=1,\ L(n+1)=\min\{j>L(n) : X(n) < X_j \le X(n)+C\}, \qquad (3.1.1)$$

$$X(0)=0,\ X(n)= X_{L(n)},\ n \ge 1.$$

Таким образом, отличие данной рекордной модели от классической состоит в том, что новая рекордная величина не должна превышать предыдущий рекорд более, чем на заранее фиксированное $C>0$. Случайные величины, которые значительно превысили предыдущее рекордное значение, в рамках данной модели считаются сомнительными и в рассмотрении не учитываются.

Для данной схемы в случае непрерывных распределений было показано (Невзоров, (2013)), что рекордные величины образуют цепь Маркова. При этом условная плотность распределения $X(n)$ при фиксированном значении $X(n-1)=x_{n-1}$ имеет следующий вид:

$$f_{(X(n)|\, x_{n-1})}(x_n)=\frac{f(x_n)}{F(x_{n-1}+C)-F(x_{n-1})},\ n \ge 2, \qquad (3.1.2)$$

если $x_{n-1} < x_n \le x_{n-1}+C$, и $f_{(X(n)|\, x_{n-1})}(x_n)=0$, иначе,

а плотность совместного распределения $X(1)$, $X(2)$, ..., $X(n)$ представляется в виде:

$$f_{(X(1),\dots,X(n))}(x_1,\dots,x_n)== f(x_1)\prod_{i=2}^{n}\frac{f(x_i)}{F(x_{i-1}+C)-F(x_{i-1})},\ n \ge 2, \qquad (3.1.3)$$

если $x_{i-1} < x_i \le x_{i-1} + C$, $i \ge 2$, и $f_{(X(1),\ldots,X(n))}(x_1,\ldots,x_n) = 0$, иначе.

Также было показано, что в случае $E(\lambda)$- распределенных исходных случайных величин имеет место следующее представление для $X(n)$:

$$X(n) \stackrel{d}{=} \sum_{i=1}^{n} V_i , \tag{3.1.4}$$

где V_n, $n \ge 1$ независимые с.в. , V_1-представляет собой $E(\lambda)$-распределенную случайную величину, а плотность распределения V_n , $n \ge 2$ определяется соотношением:

$$g(v) = \frac{1}{\lambda} \frac{\lambda e^{-v/\lambda}}{1 - e^{-C/\lambda}},$$

если $0 \le v \le C$, и $g(v)=0$, иначе.

Заметим, что схема рекордов с ограничениями очень похожа на рекордную модель Пфайфера: каждый раз после достижения очередного рекорда $X(n) = x_n$ осуществляется переход к новой серии величин с плотностью распределения, определяемой соотношением (3.1.2).

3.2 Обобщение модели рекордов с ограничениями. Случай непрерывных распределений

Рассмотрим более общий случай схемы рекордов с ограничениями.
Пусть даны две последовательности $\{C_n , n \ge 1\} \ge 0$, $\{c_n , n \ge 1\} > 0$.
Положим $L(0)=0$, $L(1)=1$, $X_0 = 0$. Последующие рекордные моменты $L(n)$ и рекордные величины $X(n)$ определим следующим образом:

$$L(n+1) = min\{j > L(n) : X(n) + c_n < X_j \le X(n) + C_n\},\ X(n) = X_{L(n)} , \tag{3.2.1}$$

$$X(n) = X_{L(n)} ,\ n \ge 1.$$

Рассуждая схожим образом, как и при доказательстве соотношений (2.2.5) и (2.2.6), приходим к формулам, аналогичным (3.1.2) и (3.1.3), которые в рамках рассматриваемой модели перепишутся следующим образом:

$$f_{(X(n,)|\, x_{n-1})}(x_n) = \frac{f(x_n)}{F(x_{n-1}+C_{n-1}) - F(x_{n-1}+c_{n-1})}, \quad n \geq 2, \qquad (3.2.2)$$

если $x_{n-1}+c_{n-1} < x_n \leq x_{n-1}+C_{n-1}$, и $f_{(X(n)|\, x_{n-1})}(x_n) = 0$, иначе,

$$f_{(X(1),\dots,\, X(n))}(x_1,\dots,x_n) = f(x_1)\prod_{i=2}^{n}\frac{f(x_i)}{F(x_{i-1}+C_{i-1}) - F(x_{i-1}+\mathrm{c}_{i-1})}, \quad n \geq 2, \qquad (3.2.3)$$

если $x_{i-1}+c_{i-1} < x_i \leq x_{i-1}+C_{i-1}$, $i \geq 2$, и $f_{(X(1),\dots,\, X(n))}(x_1,\dots,x_n) = 0$, иначе.

При этом функция распределения $X(n)$ будет иметь следующий вид:

$$F_{X(n)}(x_n) = \int\limits_{x_{n-1}-C_n}^{x_{n-1}-c_n} \dots \int\limits_{x_2-C_1}^{x_2-c_1} f_{(X(1,k),\dots,\, X(n,k))}(v_1,\dots,v_n)\,dv_1\cdot\ldots\cdot dv_n, \qquad (3.2.4)$$

где область интегрирования задается неравенствами:

$$v_n < x_n, \quad v_i - C_{i-1} < v_{i-1} < v_i - c_{i-1}, \quad 2 \leq i \leq n.$$

Пусть теперь $\{X_n,\ n \geq 1\}$ имеют $\mathrm{E}(\lambda)$-распределение, $C_0 = C_1 = \ldots = C \geq 0$, $c_0 = c_1 = \ldots = c > 0$. Также предположим, что $L(1)$ определяется соотношением (3.2.1).

Из (3.2.2) получаем следующее:

$$\mathrm{P}\{X(n) > x_n \mid X(n-1) = x_{n-1}\} = \frac{\exp(-(x_n - x_{n-1})/\lambda)}{\exp(-c/\lambda) - \exp(-C/\lambda)}, \quad x_{n-1}+c < x_n \leq x_{n-1}+C, \quad n \geq 1$$

Откуда следует, что

$$X(n) \overset{d}{=} x_{n-1} + V_n,$$

где плотность распределения V_n определяется следующим выражением:

$$g(v) = \frac{e^{-\mathrm{v}/\lambda}}{\lambda\left(e^{-\mathrm{c}/\lambda} - e^{-\mathrm{C}/\lambda}\right)}, \qquad (3.2.5)$$

если $c < v \leq C$, и $g(v) = 0$, иначе.

Значит, для всех $n \geq 1$ верно:

$$X(n) \overset{d}{=} X(n-1) + V_n ,$$

и следовательно

$$X(n) \overset{d}{=} \sum_{i=1}^{n} V_i , \quad n \geq 1, \tag{3.2.6}$$

где V_i, $1 \leq i \leq n$ независимые с.в., плотность которых определяется соотношением (3.2.5).

Если в данной модели положить все $C_i = 0$, а $c_1 = \ldots = c_n = \ldots = \delta$, то получим в этом частном случае так называемую δ-схему (Balakrishnan, Balasubramanian, Panchapakesan, (1997)).

3.3 Обобщение модели рекордов с ограничениями. Случай дискретных распределений

Рассмотрим прежнюю схему в случае дискретно распределенных случайных величин, принимающих с ненулевыми вероятностями целые неотрицательные значения. Будем предполагать в данном случае, что $\{C_n , n \geq 1\}$ и $\{c_n , n \geq 1\}$ принимают натуральные значения.
Действуя, как и при доказательстве соотношения (2.3.1), получаем следующий вид для условного распределения $X(n)$:

$$\mathrm{P}\{X(n) = x_n \mid X(1) = x_1, \ldots, X(n-1) = x_{n-1}\} = \tag{3.3.1}$$
$$= \mathrm{P}\{X(n) = x_n \mid X(n-1) = x_{n-1}\} = \frac{\mathrm{P}\{X = x\}}{F(x_{n-1} + C_{n-1}) - F(x_{n-1} + c_{n-1})} , \; n \geq 2,$$

если $x_{n-1} + c_{n-1} < x_n \leq x_{n-1} + C_{n-1}$, и $\mathrm{P}\{X(n) = x_n \mid X(1) = x_1, \ldots, X(n-1) = x_{n-1}\} = 0$, иначе.

Дискретный аналог формулы (3.2.3) представляется в следующем виде:

$$\mathrm{P}\{X(1) = x_1, \ldots, X(n) = x_n\} = \mathrm{P}\{X = x_1\} \prod_{i=2}^{n} \frac{\mathrm{P}\{X = x_i\}}{F(x_{i-1} + C_{i-1}) - F(x_{i-1} + \mathrm{c}_{i-1})} , \; n \geq 2 , \tag{3.3.2}$$

если $x_{i-1}+c_{i-1}<x_i\le x_{i-1}+C_{i-1}$, $i\ge 2$, и $\mathrm{P}\{X(1)=x_1,\ldots,X(n)=x_n\}=0$, иначе,

Рассмотрим частный случай, когда $\{X_n,\ n\ge 1\}$ имеют Geo(p) распределение, $C_0= C_1=\ldots = C$, $c_0= c_1=\ldots = c$. Также предположим, что $L(1)$ определяется соотношением (3.2.1).
В этом случае (3.3.2) будет иметь следующий вид:

$$\mathrm{P}\{X(1)=x_1,\ldots,X(n)=x_n\}=\prod_{i=1}^{n}\frac{p_i^{(x_i-x_{i-1})-1}(1-p_i)}{\left(p^c-p^C\right)}\ ,\ n\ge 1\ , \qquad (3.3.3)$$

После замены $v_n=x_n-x_{n-1}$, что соответствует переходу к с.в. $V(n)=X(n)-X(n-1)$, $n\ge 1$, (3.3.3) представится в виде:

$$\mathrm{P}\{V(1)=m_1,V(2)=m_2,\ldots,V(n)=m_n\}=\prod_{i=1}^{n}\frac{p^{m_i-1}(1-p)}{\left(p^c-p^C\right)}\quad . \qquad (3.3.4)$$

Из (3.3.4) получаем следующее представление для $X(n)$:

$$X(n)\ \overset{d}{=}\ \sum_{i=1}^{n}V_i\ ,\quad n\ge 1, \qquad (3.3.5)$$

где V_i, $1\le i\le n$ независимые с.в., распределение которых определяется выражением:

$$\mathrm{P}\{V_n=m_n\}=\frac{p^{m_n-1}(1-p)}{\left(p^c-p^C\right)},\ \text{где}\quad c<m_n\le C\ .$$

4 Заключение

В первой части данной работы приведены некоторые необходимые сведения из теории рекордов. Дан обзор классической рекордной схемы и некоторых неклассических схем - F^{α}- схемы и схемы с линейным сносом.

Во второй и третьей частях были рассмотрены две рекордные схемы – модель Пфайфера и схема рекордов с ограничениями.

В модели Пфайфера были рассмотрены *k*-рекордные моменты и *k*-рекордные величины и получены некоторые новые результаты для данной обобщенной модели. В частности, было установлено, что последовательность *k*-рекордных величин образует цепь Маркова, найдены соответствующие переходные вероятности (2.2.4) и выражение (2.2.6) для плотности совместного распределения первых *n* рекордных величин. Также была отмечена связь *k*-ых и обычных рекордов. При дополнительном предположении об экспоненциальном $E(\lambda_m)$- распределении, исходных случайных величин $\{Z_m^{(n)},\ m \ge 1,\ n \ge 1\}$, найдено представление для рекордной величины Z (*n*, *k*) и установлена её асимптотическая нормальность в случае, когда ряд $\sum_{i=1}^{n} \lambda_i^{\ 2}$ расходится.

В пункте 2.3 была рассмотрена схема Пфайера в дискретном случае. В предположении геометрического Geo(p_n)- распределения, исходных случайных величин $\{X_k^{(n)},\ k \ge 1,\ n \ge 1\}$, найдено представление для рекордной величины $X(n)$.

В модели рекордов с ограничениями был рассмотрен более общий случай – схема с верхней и нижней границами. Для непрерывного и дискретного случая были найдены совместные распределения рекордных величин((3.2.3) и (3.3.3)). В непрерывном случае при предположении об экспоненциальном распределении, а в дискретном – при предположении геометрического распределения исходных случайных величин были найдены соответствующие представления для *n*-ой рекордной величины((3.2.6) и (3.3.5)).
Рассмотренные рекордные схемы являются некоторым обобщением классической модели.

Список литературы

[1] Невзоров В.Б., Рекорды. Математическая теория, М, Фазис, 2000.

[2] Arnold B.C., Balakrishnan N., Nagaraja. H.N., A Wiley-Interscience Publication JOHN WILEY & SONS, INC., 1998.

[3] Ahsanullah M., Nevzorov V.B. Ordered random variables. Huntington, New York, 2001.

[4] Renyi A., Theorie des elements saillants d'une suite d'observations, Collquium on Combinatorial Methods in Probability Theory, Nathematisk Institut, Aarhus University, Aarhus, Denmark. English translation in Selected Papers of Alfred Renyi, Volume 2, Academic Press, New York (1962).

[5] Ballerini R., Resnick S. I.,Records from improving populations, Journal of Applied Probability, 22 (1985).

[6] Ballerini R., Resnick S. I., Records in the presence of a linear trend, Advances in Applied Probability, 19, 801-828 (1987).

[7] Shorrock, R. W. A limit theorem for inter-record times, Journal of Applied Probability, 9, 219-223. (Correction 9, p. 877.) (1972).

[8] Pfeifer, D., Limit laws for inter-record times from nonhomogeneous record values, Journal of Organizational Behavior and Statistics, 1 69-74 (1984).

[9] Yang, M. C. K. On the distribution of the inter-record times in an increasing population, Journal of Applied Probability, 12, 148- 154 (1975).

[10] Невзоров В.Б., Рекордные величины с ограничением. Вестник СПбГУ, Сер.1(2013), выпуск 3, С. 70-74.

[11] Balakrishnan N., Balasubramanian K., Panchapakesan S., δ -exceedance records Невзоров В.Б., Journal of Applied Statistical Science, 4, 123-132. (1997).