

WIESŁAW DZIUBDZIELA (Kostomłoty Pierwsze)

K -te wartości rekordowe i ich podstawowe własności

(Artykuł dedykuję żonie Dorocie)

Streszczenie W pracy przedstawiony jest model k -tych wartości rekordowych. K -te wartości rekordowe są ważnym przykładem uporządkowanych zmiennych losowych. Pojawiają się one w sposób naturalny w realnym życiu, w przypadku, gdy jesteśmy zainteresowani kolejnymi k -tymi maksymalnymi obserwacjami. Formalną definicję k -tych wartości rekordowych podali Dziubdziela i Kopociński w 1976 roku. Artykuł zawiera probabilistyczną teorię dla tego modelu.

2010 *Klasyfikacja tematyczna AMS (2010)*: Primary: 60G70; Secondary: 62G30.

Słowa kluczowe: k -te wartości rekordowe, k -te czasy rekordowe, liczba k -tych wartości rekordowych, proces punktowy k -tych wartości rekordowych, reprezentacje k -tych wartości rekordowych.

1. Wstęp

Niech $X_1, X_2, \dots$ będzie ciągiem zmiennych losowych. Definiujemy ciąg zmiennych losowych

$$\begin{aligned} L(0) &= 0, \quad L(1) = 1 \\ L(n+1) &= \min \{j : X_j > X_{L(n)}\}, \quad n = 1, 2, \dots \end{aligned}$$

Zmienne losowe $L(n)$, $n = 1, 2, \dots$, nazywamy czasami rekordowymi. Interesować nas będą także następujące zmienne losowe zdefiniowane dla $n = 1, 2, \dots$:
czasy międzyrekordowe

$$\Delta(n) = L(n) - L(n-1)$$

i wartości rekordowe lub krótko *rekordy*

$$X(n) = X_{L(n)}.$$

Chandler (1952) badał ciągi $\{L(n)\}, \{\Delta(n)\}$ i $\{X(n)\}$ przy założeniu, że ciąg $X_1, X_2, \dots$ jest ciągiem niezależnych zmiennych losowych o jednokowej ciągłej dystrybucji F . Między innymi udowodnił następujący lemat i twierdzenie (dowody przedstawimy później).

LEMAT 1.1 *Jeżeli F jest ciągłą dystrybuantą, to rozkład prawdopodobieństwa zmiennej losowej $L(n)$, $n = 1, 2, \dots$, nie zależy od postaci dystrybuanty.*

TWIERDZENIE 1.2 *Rozkład zmiennej losowej $L(2)$ jest postaci*

$$P(L(2) = j) = \frac{1}{j(j-1)}, \quad j = 2, 3, \dots$$

WNIOSEK 1.1 *Dla $n = 2, 3, \dots$*

$$E(L(n)) = +\infty.$$

Znalezienie rozkładów prawdopodobieństwa czasów rekordowych dla $n \geq 3$ jest bardziej skomplikowane. A. Rényi (1962) (see also [47, p. 50–65]) udowodnił, że ciąg $L(n)$, $n = 1, 2, \dots$, tworzy łańcuch Markowa z prawdopodobieństwami przejść

$$P\{L(n) = k | L(n-1) = l\} = \frac{l}{k(k-1)},$$

gdy $k > l \geq n-1$, $n = 2, 3, \dots$. Własność tę wykorzystał do znalezienia dwuwymiarowych łącznych rozkładów czasów rekordowych i w konsekwencji do znalezienia rozkładów brzegowych czasów rekordowych. Praca Chandler (1952) spowodowała pojawienie się w latach pięćdziesiątych serii prac: Foster and Stuart (1954), Foster and Teichroew (1955), Stuart (1956, 1957), w których rekordy wykorzystano w zagadnieniach związanych z testowaniem hipotez statystycznych. Wkrótce jednak nastąpił spadek zainteresowania teorią rekordów. Wydaje się to związane z faktem, że $E(L(2)) = +\infty$, tzn. średni czas czekania na rekord ma nieskończoną wartość oczekiwaną.

Impulsem, który pobudził dalszy rozwój teorii rekordów i jej zastosowań była opublikowana w 1962 roku praca A. Rényi'ego [35], a dokładniej lemat, który nazywamy dzisiaj Lematem Rényi'ego. Przedstawimy go w dalszym ciągu pracy.

Praca poświęcona jest naturalnemu rozszerzeniu pojęcia rekordu. Dziubiela and Kopociński (1976) zdefiniowali tzw. k -te wartości rekordowe. Pojęcie to znalazło uznanie w literaturze. Do chwili obecnej istnieje kilkaset publikacji, w których wykorzystywane jest to pojęcie. Na poszczególnych etapach rozwoju teorii powstały także artykuły przeglądowe, podręczniki i monografie poświęcone tej teorii, zobacz np. Glick (1978), Nevzorov (1987), Nagaraja (1988), Arnold et al. (1998), Nevzorov (2001), Ahsanullah and Nevzorov (2015), Shahbaz et al. (2016).

2. Definicje

Niech $X_1, X_2, \dots$ będzie ciągiem zmiennych losowych. Oznaczmy

$$\begin{aligned} X_{1:n} &= \min\{X_1, X_2, \dots, X_n\} \leq X_{2:n} \leq \dots \leq X_{n-1:n} \leq \\ &\leq X_{n:n} = \max\{X_1, X_2, \dots, X_n\} \end{aligned}$$

statystyki pozycyjne w ciągu $X_1, X_2, \dots, X_n$, gdzie $n \geq 1$.

Rozważmy ciąg wektorów losowych

$$\mathbf{X}_{n-k+1:n} = (X_{n-k+1:n}, X_{n-k+2:n}, \dots, X_{n:n}), \quad n \geq k.$$

Podamy teraz definicję k -tej wartości (statystyki) rekordowej, gdzie $k \geq 1$ jest ustaloną liczbą naturalną. Dziubdziela and Kopociński (1976) (zob. także Dziubdziela (1977, 1990)) wprowadzili k -te czasy rekordowe $L(1, k), n \geq k$, jako kolejne momenty, w których ciąg $\{\mathbf{X}_{n-k+1:n}\}$ zmienia wartości. Oczywiście, mamy

$$L(1, k) = k,$$

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

k -te wartości rekordowe (k -te statystyki rekordowe, k -rekordy) są zdefiniowane przez

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

Co więcej, odpowiednie między- k -rekordowe czasy są określone przez

$$\Delta(1, k) = L(1, k) = k,$$

$$\Delta(n, k) = L(n, k) - L(n-1, k), \quad n \geq 2.$$

W przypadku $k = 1$, otrzymujemy zdefiniowane wcześniej czasy rekordowe $L(n)$, czasy międzyrekordowe $\Delta(n)$ i rekordy $X(n)$.

Wydaje się, że w zrozumieniu sensu definicji k -tych wartości rekordowych pomoże następujący przykład zaczerpnięty z dziedziny sportu. Rozważmy, na przykład, dziesięć najlepszych wyników na świecie w skoku w dal mężczyzn. W danym momencie czasu są to: rekord świata (wartość rekordowa), drugi wynik (druga wartość rekordowa), ..., dziesiąty wynik (dziesiąta wartość rekordowa). Jeżeli w przyszłości zawodnik osiągnie wynik mieszczący się w pierwszej dziesiątce, to ulegają zmianie odpowiednie wartości rekordowe.

Możemy spojrzeć na definicję k -tych wartości rekordowych w inny sposób. Rozważmy niemalejący ciąg k -tych statystyk pozycyjnych

$$-\infty < X_{1:k} \leq X_{2:k+1} \leq \dots \leq X_{n-k:n-1} \leq X_{n-k+1:n} \leq \dots$$

Usuając w tym ciągu elementy powtarzające się, otrzymujemy ściśle rosnący ciąg kolejnych k -tych wartości rekordowych

$$-\infty < X_{1:k} < X_{L(2,k)-k+1:L(2,k)} < X_{L(3,k)-k+1:L(3,k)} < \dots$$

W pracy Dziubdziela and Kopociński (1976) rozważano ciąg $X_1, X_2, \dots$ niezależnych zmiennych losowych z jednakową **ciągłą** dystrybucją. Z założenia ciągłości wynika, że definicja k -tych czasów rekordowych i k -tych rekordowych wartości może być zapisana w równoważnej postaci

$$\widehat{L}(1, k) = 1,$$

$$\widehat{L}(n+1, k) = \min \left\{ j > \widehat{L}(n, k) : X_{\widehat{L}(n, k): \widehat{L}(n, k)+k-1} < X_{j:j+k-1} \right\},$$

$n \geq 1$, oraz

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

[Dembińska and López-Blázquez \(2005\)](#) zauważyli, że w przypadku dyskretnych rozkładów powyższe definicje k -tych czasów rekordowych i k -tych rekordowych wartości nie są równoważne.

Podali oni następujący przykład:

PRZYKŁAD 2.1 Niech ciąg $X_1, X_2, \dots$ ma realizację $0, 1, 1, 1, 0, 4, 0, 3, \dots$. Wtedy

$$\{L(n, 2), n \geq 1\} = \{2, 3, 6, 8, \dots\}$$

i

$$\{\widehat{L}(n, 2), n \geq 1\} = \{1, 2, 7, \dots\}.$$

Oraz

$$\{X_{L(n, 2) - 2 + 1: L(n, 2)}, n \geq 1\} = \{0, 1, 1, 3, \dots\}$$

i

$$\{X_{\widehat{L}(n, 2): \widehat{L}(n, 2) + 2 - 1}, n \geq 1\} = \{0, 1, 3, \dots\}.$$

Wynika stąd, że zbiory drugich rekordów nie są równe.

Kończąc paragraf o definicjach, przedstawimy definicję opierającą się na pojęciu rangi.

Sekwencyjna ranga I_n to ranga X_n wśród $X_1, X_2, \dots, X_n$, tzn.

$$X_n = X_{I_n: n}, \quad n \geq 1$$

albo

$$I_n = \sum_{i=1}^n \mathbf{1}_{\{X_i \leq X_n\}}, \quad n \geq 1,$$

gdzie $\mathbf{1}_A$ oznacza indykator zdarzenia losowego A .

Możemy już teraz zdefiniować k -te czasy rekordowe wzorem

$$L(1, k) = k,$$

$$L(n+1, k) = \min \{j > L(n, k) : n - k + 1 \leq I_j \leq n\}, \quad n \geq 1.$$

Zauważmy, że w pracy [Pickands \(1971\)](#) (zob. także [\[3, 4\]](#)) wprowadzono tzw. ścisłe (typu pierwszego) czasy k -rekordowe i ścisłe k -rekordy, przyjmując odpowiednio

$$L_s(1, k) = k, \quad L_s(n+1, k) = \min \{j > L_s(n, k) : I_j = n - k + 1\}, \quad n \geq 1,$$

i

$$X_s(n, k) = X_{L_s(n, k) - k + 1: L_s(n, k)}, \quad n \geq 1.$$

3. Lemat Rényi'ego i indykatory k -tych rekordów

Od tego paragrafu będziemy zakładać, że ciąg $X_1, X_2, \dots$ jest ciągiem niezależnych zmiennych losowych o jednakowej ciągłej dystrybuancie F .

Rényi (1962) udowodnił wspomniany we wstępie

LEMAT 3.1 (RÉNYI'EGO) *Dla ciągu $X_1, X_2, \dots$ niezależnych zmiennych losowych o jednakowej ciągłej dystrybuancie F , rangi $I_1, I_2, \dots$ są wzajemnie niezależne oraz*

$$P(I_n = j) = \frac{1}{n}, \quad j = 1, 2, \dots, n, \quad n \geq 1.$$

DOWÓD. Z założenia ciągłości dystrybuanty F wynika, że równości wśród realizacji zmiennych losowych $X_1, X_2, \dots, X_n$ występują z prawdopodobieństwem zero. Niech $i_1, i_2, \dots, i_n$ będzie permutacją liczb $1, 2, \dots, n$. Każde z możliwych uporządkowań $X_{i_1} < X_{i_2} < \dots < X_{i_n}$ jest jednakowo prawdopodobne. Zauważmy, że istnieje wzajemnie jednoznaczna odpowiedniość pomiędzy tymi uporządkowaniami a realizacjami rang $I_1, I_2, \dots, I_n$. Na przykład, uporządkowaniu $X_3 < X_1 < X_4 < X_2$ odpowiada realizacja $I_1 = 2, I_2 = 4, I_3 = 1, I_4 = 3$. Z powyższego wynika, że realizacje rang są jednakowo prawdopodobne, przy czym

$$P(I_1 = r_1, I_2 = r_2, \dots, I_n = r_n) = \frac{1}{n!}$$

dla $r_i = 1, 2, \dots, i, i = 1, 2, \dots, n$. Zauważmy następnie, że

$$\begin{aligned} P(I_n = j) &= \sum_{(r_1, r_2, \dots, r_{n-1})} P(I_1 = r_1, I_2 = r_2, \dots, I_{n-1} = r_{n-1}, I_n = j) \\ &= \sum_{(r_1, r_2, \dots, r_{n-1})} \frac{1}{n!} = \frac{1 \cdot 2 \cdot \dots \cdot (n-1)}{n!} = \frac{1}{n}, \end{aligned}$$

a stąd wynika niezależność rang

$$P(I_1 = r_1, I_2 = r_2, \dots, I_n = r_n) = \prod_{i=1}^n P(I_i = r_i).$$

Dzisiaj wiadomo, że Lemat Rényi'ego był znany wcześniej, zobacz **Wilks (1941)**, **Dwass (1960)**.

Zmienne losowe

$$\xi_n(k) = \mathbf{1}_{\{X_n > X_{n-k;n-1}\}}, \quad n \geq k,$$

gdzie $\xi_1(1) = 1$, nazywamy indykatorami k -tych wartości rekordowych. Oczywiście

$$\xi_n(k) = \mathbf{1}_{\{n-k+1 \leq I_n \leq n\}}.$$

Wnioskiem z Lematu Rényi'ego jest następujące twierdzenie, które pierwszy sformułował V. B. Nevzorov [30] w roku 1986.

Twierdzenie 3.2 *Jeżeli $X_1, X_2, \dots$ są niezależnymi zmiennymi losowymi o jednakowej ciągłej dystrybucji F , to indykatory k -tych wartości rekordowych $\xi_k(k), \xi_{k+1}(k), \dots$ są wzajemnie niezależne i mają rozkład prawdopodobieństwa*

$$P(\xi_n(k) = 1) = 1 - P(\xi_n(k) = 0) = \frac{k}{n}, \quad k \geq 1, \quad n \geq k. \quad (1)$$

Dowód. Mamy

$$P(\xi_n(k) = 1) = P(n - k + 1 \leq I_n \leq n) = \sum_{r=1}^k P(I_n = r),$$

zatem z Lematu Rényi'ego otrzymujemy niezależność indykatorów k -tych wartości rekordowych i ich rozkład postaci (1).

Zanim przejdziemy do badania liczby k -tych wartości rekordowych, udowodnimy lemat, który wykorzystamy w następnym paragrafie pracy (zob. także [32], Lemma 13.2).

Lemat 3.3 *Niech $X_1, X_2, \dots$ będą niezależnymi zmiennymi losowymi o jednakowej ciągłej dystrybucji F . Oznaczmy $M_n = \max\{X_1, X_2, \dots, X_n\}$. Wtedy, dla każdego $n \geq 1$, indykatory $\xi_1(1), \xi_2(1), \dots, \xi_n(1)$ oraz M_n są wzajemnie niezależne.*

Dowód. Wybierzmy dowolne $n \geq 2$. Zauważmy, że z symetrii (niezależność i jednakowy rozkład) mamy

$$P(\xi_n(1) = 1) = P(X_n = M_n) = P(X_r = M_n), \quad r = 1, 2, \dots, n.$$

Natomiast z założenia ciągłości dystrybuanty

$$P(X_i = X_j) = 0$$

dla $i \neq j$. Wtedy

$$1 = P(X_1 = M_n) + P(X_2 = M_n) + \dots + P(X_n = M_n) = nP(X_n = M_n)$$

i

$$P(\xi_n(1) = 1) = P(X_n = M_n) = \frac{1}{n}.$$

Ponieważ zmienne losowe $\xi_1(1), \xi_2(1), \dots, \xi_n(1)$ przyjmują jedynie dwie wartości, więc, aby udowodnić wzajemną niezależność zmiennych losowych $\xi_1(1), \xi_2(1), \dots, \xi_n(1)$ i M_n , wystarczy pokazać, że dla dowolnego $x, r = 2, 3, \dots, s = 1, 2, \dots, r$ oraz

$$1 \leq \alpha(1) < \alpha(2) < \dots < \alpha(s) \leq r$$

mamy

$$\begin{aligned} P(\xi_{\alpha(1)}(1) = 1, \dots, \xi_{\alpha(s)}(1) = 1, M_r \leq x) = \\ = P(\xi_{\alpha(1)}(1) = 1) \cdots P(\xi_{\alpha(s)}(1) = 1) P(M_r \leq x). \end{aligned} \quad (2)$$

Dowód (2) nie jest trudny, ale w przypadku dowolnego s jest złożony. Dla prostoty rozważmy $s = 2$. Wtedy

$$\begin{aligned} P(\xi_{\alpha(1)}(1) = 1, \xi_{\alpha(2)}(1) = 1, M_r \leq x) = \\ = P(\max(X_1, \dots, X_{\alpha(1)-1}) < X_{\alpha(1)} \leq x; \\ \max(X_{\alpha(1)}, \dots, X_{\alpha(2)-1}) < X_{\alpha(2)} \leq x; \max(X_{\alpha(2)+1}, \dots, X_r) \leq x) = \\ = (F(x))^{r-\alpha(2)} \int_{-\infty}^x \int_u^x (F(u))^{\alpha(1)-1} (F(v))^{\alpha(2)-\alpha(1)-1} dF(v) dF(u) = \\ = \frac{F^r(x)}{\alpha(1)\alpha(2)} = P(\xi_{\alpha(1)}(1) = 1) \cdot P(\xi_{\alpha(2)}(1) = 1) \cdot P(M_r \leq x). \end{aligned}$$

Oznaczmy przez $N(n, k)$, $n \geq k$, liczbę k -tych wartości rekordowych wśród zmiennych losowych $X_1, X_2, \dots, X_n$. Zauważmy, że

$$N(n, k) = \xi_k(k) + \xi_{k+1}(k) + \dots + \xi_n(k),$$

zatem zmienna losowa $N(n, k)$ ma rozkład dwumianowy ze zmiennymi prawdopodobieństwami sukcesów p_n , $n = k, k+1, \dots, n$, danymi przez (1).

Ciągi niezależnych zmiennych Bernoulliego ze zmiennymi prawdopodobieństwami sukcesu w i -tej próbie

$$P_i(\{s\}) = p_i, \quad i = 1, 2, \dots, \quad (3)$$

rozważał już w 1837 roku S. D. Poisson [34]. Obecnie badanie takich ciągów zmiennych Bernoulliego w dalszym ciągu budzi duże zainteresowanie i wciąż otrzymywane są nowe wyniki. My wykorzystamy twierdzenie z pracy Ch. A. Charalambidesa ([8], Theorem 3.1, zobacz także [7]).

Zanim go jednak sformułujemy, musimy wprowadzić tzw. uogólnione liczby Stirlinga pierwszego rodzaju [the generalized signless Stirling numbers of the first kind] $|s(n, k; \mathbf{a})|$ (zob. np. [8]).

Uogólnione liczby Stirlinga pierwszego rodzaju można zdefiniować przez

$$\prod_{i=1}^n (t + a_i) = \sum_{k=0}^n |s(n, k; \mathbf{a})| t^k,$$

gdzie $\mathbf{a} = (a_1, a_2, \dots, a_n)$. Zauważmy, że

$$|s(n, n-k; \mathbf{a})| = \sum a_{i_1} a_{i_2} \dots a_{i_k},$$

gdzie sumowanie przebiega po wszystkich k -elementowych kombinacjach zbioru indeksów $1, 2, \dots, n$, jest symetryczną funkcją n zmiennych $(a_1, a_2, \dots, a_n)$. W szczególnym przypadku, gdy $k = 1, 2$ i $n = k, k + 1, \dots$, mamy

$$|s(n, 1; \mathbf{a})| = \left(\prod_{i=1}^n a_i \right) \left(\sum_{i=1}^n \frac{1}{a_i} \right), \quad |s(n, 2; \mathbf{a})| = \left(\prod_{i=1}^n a_i \right) \left(\sum_{j=2}^n \sum_{i=1}^{j-1} \frac{1}{a_i a_j} \right).$$

TWIERDZENIE 3.4 *Rozkład prawdopodobieństwa liczby sukcesów ζ_n w n próbach Bernoulliego ze zmiennymi prawdopodobieństwami sukcesu (3) jest dany przez*

$$P(\zeta = x) = \frac{|s(n, x; \mathbf{a})|}{\prod_{i=1}^n (1 + a_i)}, \quad x = 0, 1, \dots, n, \quad (4)$$

gdzie $a_i = (1 - p_i)/p_i$, $i = 1, 2, \dots, n$, a j -ty moment silniowy

$$E[(\zeta_n)_j] = E[\zeta_n(\zeta_n - 1) \cdots (\zeta_n - j + 1)]$$

jest dany przez

$$E[(\zeta_n)_j] = \frac{j! |s(n, j; \mathbf{a} + \mathbf{1})|}{\prod_{i=1}^n (1 + a_i)}, \quad j = 1, 2, \dots, n, \quad (5)$$

i

$$E[(\zeta_n)_j] = 0, \quad j = n + 1, n + 2, \dots \quad (6)$$

W szczególności, wartość oczekiwana i wariancja wynoszą

$$E(\zeta_n) = \sum_{i=1}^n \frac{1}{1 + a_i} = \sum_{i=1}^n p_i, \quad \text{Var}(\zeta_n) = \sum_{i=1}^n \frac{a_i}{(1 + a_i)^2} = \sum_{i=1}^n p_i(1 - p_i). \quad (7)$$

Dowód Twierdzenia 3.4 można znaleźć w pracy ([8]).

Jeżeli rozważymy zdefiniowaną wcześniej liczbę k -rekordów $N(n, k)$ w ciągu zmiennych losowych $X_1, X_2, \dots, X_n$, to mamy

$$N(n, k) = \zeta_{n-k+1},$$

gdzie

$$p_i = \frac{k}{i + k - 1}, \quad i = 1, 2, \dots, n - k + 1. \quad (8)$$

Z Twierdzenia 3.4 otrzymujemy następujący wniosek

WNIOSEK 3.1 *Niech $N(n, k)$ będzie liczbą k -rekordów w ciągu niezależnych zmiennych losowych $X_1, X_2, \dots, X_n$ o ciągłej dystrybucji, to*

$$P(N(n, k) = x) = \frac{k^{n-k} |s(n - k + 1, x; \mathbf{a})|}{(k + 1)(k + 2) \cdots n}, \quad x = 0, 1, \dots, n - k + 1, \quad (9)$$

gdzie $a_i = \frac{i-1}{k}$, $i = 1, 2, \dots, n-k+1$, natomiast j -ty moment silniowy jest równy

$$E \left[(N(n, k))_j \right] = \frac{j! k^{n-k} |s(n-k+1, j; \mathbf{a} + \mathbf{1})|}{(k+1)(k+2) \cdots n}, \quad (10)$$

$j = 1, 2, \dots, n-k+1$, oraz

$$E \left[(N(n, k))_j \right] = 0, \quad j = n-k, n-k+1, \dots. \quad (11)$$

DOWÓD. Korzystając ze wzorów (8), otrzymujemy

$$a_i = \frac{1-p_i}{p_i} = \frac{1 - \frac{k}{i+k-1}}{\frac{k}{i+k-1}} = \frac{i-1}{k}$$

oraz

$$\prod_{i=1}^{n-k+1} (1+a_i) = \prod_{i=1}^{n-k+1} \frac{k+i-1}{k} = k^{-(n-k+1)} k(k+1) \cdots n.$$

Zatem z (4) otrzymujemy (9) i z (5), (6) otrzymujemy (10), (11).

Wartość oczekiwaną i wariancję liczby k -tych rekordów $N(n, k)$ otrzymuje się z (7):

$$E(N(n, k)) = \sum_{i=1}^{n-k+1} p_i = \sum_{i=1}^{n-k+1} \frac{k}{i+k-1} = k \sum_{m=k}^n \frac{1}{m}$$

i

$$\text{Var}(N(n, k)) = \sum_{i=1}^{n-k+1} p_i - \sum_{i=1}^{n-k+1} p_i^2 = k \sum_{m=k}^n \frac{1}{m} - k^2 \sum_{m=k}^n \frac{1}{m^2}$$

(porównaj [31], wzór (19.14) i (19.16)).

4. k -te czasy rekordowe

Z Twierdzenia 3.2 wiemy, że indykatory k -tych wartości rekordowych $\xi_k(k)$, $\xi_{k+1}(k)$, $\dots$ są wzajemnie niezależne. Fakt ten posłuży nam do obliczenia rozkładów k -tych czasów rekordowych

$$P_{n,k} = P\{L(1, k) = k, L(2, k) = m(2), \dots, L(n, k) = m(n)\}$$

(zobacz np. [32], str. 84).

Mamy

$$P_{n,k} = P(\xi_k(k) = 1, \xi_{k+1}(k) = 0, \dots, \xi_{m(2)-1}(k) = 0, \xi_{m(2)}(k) = 1, \dots, \\ \xi_{m(n)-1}(k) = 0, \xi_{m(n)}(k) = 1),$$

stąd

$$P_{n,k} = P(\xi_k(k) = 1)P(\xi_{k+1}(k) = 0) \cdots P(\xi_{m(2)-1}(k) = 0) \times \\ \times P(\xi_{m(2)}(k) = 1) \cdots P(\xi_{m(n)-1}(k) = 0)P(\xi_{m(n)}(k) = 1)$$

i ostatecznie, ponownie korzystając z Twierdzenia 3.2,

$$P_{n,k} = \frac{k! (m(n) - k)!}{m(n)!} \cdot \frac{k^{n-1}}{(m(2) - k)(m(3) - k) \cdots (m(n) - k)}, \quad (12)$$

gdzie $k < m(2) < \cdots < m(n)$.

Twierdzenie 4.1 Ciąg $L(1, k), L(2, k), \dots$ jest łańcuchem Markowa z prawdopodobieństwami przejścia

$$P(L(n+1, k) > r \mid L(n, k) = m) = \\ = \frac{(m+1-k) \cdots (r-k-1)(r-k)}{(m+1) \cdots (r-1)r}, \quad r \geq m \geq n+k-1 \quad (13)$$

i

$$P(L(n+1, k) = r \mid L(n, k) = m) = \\ = \frac{(m+1-k) \cdots (r-k-1)r}{(m+1) \cdots (r-1)r}, \quad r > m \geq n+k-1. \quad (14)$$

Dowód. Zauważmy, że

$$P(L(n+1, k) > r \mid L(n, k) = m) = \\ = P(L(n+1, k) > r \mid L(n, k) = m, L(n-1, k) = m_{n-1}, \dots, L(1, k) = k) = \\ = P(\xi_{m+1}(k) = 0, \dots, \xi_r(k) = 0), \quad r \geq m \geq n+k-1,$$

i

$$P(L(n+1, k) = r \mid L(n, k) = m) = \\ = P(L(n+1, k) = r \mid L(n, k) = m, L(n-1, k) = m_{n-1}, \dots, L(1, k) = k) = \\ = P(\xi_{m+1}(k) = 0, \dots, \xi_{r-1}(k) = 0, \xi_r(k) = 1), \quad r > m \geq n+k-1.$$

Z powyższych równości i Twierdzenia 3.2 otrzymujemy wzory (13) i (14).

5. Rozkład wartości rekordowej

Wykorzystując definicję czasu rekordowego $L(n, 1)$, wartość rekordową $X(n, 1)$ zapisujemy

$$X(n, 1) = X_{L(n, 1)} = X_{L(n, 1); L(n, 1)} = M_{L(n, 1)}, \quad n = 1, 2, \dots$$

Zauważmy, że ze wzoru na prawdopodobieństwo całkowite, mamy

$$\begin{aligned} P(X(n, 1) \leq x) &= P(M_{L(n)} \leq x) = \\ &= \sum_{m=1}^{\infty} P(M_{L(n)} \leq x | L(n) = m) \cdot P(L(n) = m) = \\ &= \sum_{m=1}^{\infty} P(M_m \leq x | L(n) = m) \cdot P(L(n) = m) \end{aligned}$$

Z Lematu 3.3 wynika, że dla $n, m = 1, 2, \dots$ zdarzenia

$$\{L(n) = m\} = \{\xi_1(1) + \dots + \xi_{m-1}(1) = n - 1, \xi_m(1) = 1\}$$

i zmienna losowa M_n są niezależne. Wiadomo także, że

$$P(M_n \leq x) = F^n(x).$$

Stąd

$$\begin{aligned} P(X(n, 1) \leq x) &= \sum_{m=1}^{\infty} P(M_m \leq x) \cdot P(L(n) = m) = \\ &= \sum_{m=1}^{\infty} F^m(x) \cdot P(L(n) = m) = E\left(F^{L(n)}(x)\right) = Q_n(F(x)), \end{aligned}$$

gdzie

$$Q_n(s) = E\left(s^{L(n)}\right)$$

jest funkcją tworzącą czasu rekordowego $L(n)$. W monografii Nevzorova [31] (Exercise 14.1) podana jest postać tej funkcji tworzącej

$$\begin{aligned} Q_n(s) &= 1 - (1-s) \sum_{k=0}^{n-1} \frac{(-\log(1-s))^k}{k!} = \\ &= \int_0^{-\log(1-s)} \frac{v^{n-1} \exp(-v)}{(n-1)!} dv. \end{aligned}$$

Zatem rozkład wartości rekordowej jest postaci

$$\begin{aligned} P(X(n, 1) \leq x) &= 1 - (1 - F(x)) \sum_{k=0}^{n-1} \frac{(-\log(1 - F(x)))^k}{k!} = \\ &= \int_0^{-\log(1-F(x))} \frac{v^{n-1} \exp(-v)}{(n-1)!} dv, \quad n = 1, 2, \dots \end{aligned}$$

W przypadku standardowego rozkładu wykładniczego

$$F(x) = 1 - \exp(-x), \quad x > 0,$$

mamy

$$P(X(n, 1) \leq x) = \int_0^x \frac{v^{n-1} \exp(-v)}{(n-1)!} dv.$$

Zatem w tym przypadku wartość rekordowa ma $X(n, 1)$ na rozkład gamma z parametrem n .

6. Procesy punktowe

W najbliższych dwóch paragrafach przedstawimy aparat matematyczny, który będziemy wykorzystywać w dalszej części pracy. W pierwszym omówimy procesy punktowe, a w następnym procesy Markowa.

Procesy punktowe omówimy w oparciu o podręcznik Serfozo [40] (zob. także [26], [27], [36], [37]).

„Proces punktowy” to proces liczący punkty w losowym podzbiorze pewnej przestrzeni. Zwykle tą przestrzenią jest prosta rzeczywista $\mathbb{R}$, płaszczyzna $\mathbb{R}^2$, wielowymiarowa przestrzeń euklidesowa $\mathbb{R}^d$, czy bardziej ogólnie zupełna, ośrodkowa przestrzeń metryczna (przestrzeń polska) $\mathbf{S}$. Oznaczmy przez $\mathcal{S}$ rodzinę zbiorów borelowskich z $\mathbf{S}$, oznaczmy także rodzinę ograniczonych zbiorów borelowskich przez $\hat{\mathcal{S}}$.

W naszych zastosowaniach teorii procesów punktowych zazwyczaj będziemy przyjmować, że $\mathbf{S} \subset \mathbb{R}$.

Losowy zbiór punktów w $\mathbf{S}$ jest przeliczalnym zbiorem elementów losowych $\{\tau_n\}$ o wartościach z $\mathbf{S}$ takim, że w dowolnym ograniczonym podzbiorze może znajdować się tylko skończona liczba punktów. Wtedy

$$N(B) = \sum_n \delta_{\tau_n}(B), \quad B \in \mathcal{S},$$

oznacza liczbę punktów w B , gdzie $\delta_x(B) = \mathbf{1}_{(x \in B)}$ jest miarą Diraca z jednostkową masą w x . Tą liczącą miarę N nazywamy procesem punktowym na $\mathbf{S}$ z lokalizacjami punktów $\{\tau_n\}$.

Oznaczmy przez $\mathbb{M}$ zbiór wszystkich liczących punkty miar ν na $(\mathbf{S}, \mathcal{S})$, które są localnie skończone, tzn. $\nu(B) < \infty$ dla ograniczonych zbiorów B . Miary z $\mathbb{M}$ są postaci

$$\nu(B) = \sum_n \delta_{x_n}(B), \quad B \in \mathcal{S},$$

gdzie $\{x_1, \dots, x_k\}$, $0 \leq k \leq \infty$, jest związanym z nią zbiorem punktów. Ze zbiorem $\mathbb{M}$ wiążemy generowaną przez zbiory $\{\nu \in \mathbb{M} : \nu(B) = n\}$, dla $B \in \mathcal{S}$ i $n \geq 0$, σ -algebrę $\mathcal{M}$.

Możemy już teraz podać formalną definicję procesu punkowego.

DEFINICJA 6.1 Proces punktowy na przestrzeni $\mathbf{S}$ jest mierzalnym przekształceniem N z przestrzeni probabilistycznej $(\Omega, \mathcal{F}, P)$ w przestrzeń $(\mathbb{M}, \mathcal{M})$. Zmienna losowa $N(B)$ jest liczbą punktów w zbiorze $B \in \mathcal{S}$.

Oczywiście

$$N(B) = \sum_n \delta_{\tau_n}(B), \quad B \in \mathcal{S},$$

gdzie τ_n oznaczają lokalizacje punktów przekształcenia N .

W przypadku, gdy $\mathbf{S} = \mathbb{R}^d$, to N jest lokalnie skończony, tzn. $N(B) < \infty$ dla wszystkich ograniczonych zbiorów B .

Rozkład prawdopodobieństwa procesu punktowego N , tzn. $P(N \in \cdot)$, jest określony przez jego skończenie wymiarowe rozkłady

$$P(N(B_1) = n_1, \dots, N(B_k) = n_k), \quad B_1, \dots, B_k \in \hat{\mathcal{S}}, \quad (15)$$

gdzie $n_j \in \{0, 1, 2, \dots\}$, $j = 1, 2, \dots, k$ i $k = 1, 2, \dots$.

Dwa procesy punktowe N i N' na $\mathbf{S}$ są równe według rozkładu, co piszemy $N \stackrel{d}{=} N'$, jeżeli ich skończenie wymiarowe rozkłady są równe

$$(N(B_1), \dots, N(B_k)) \stackrel{d}{=} (N'(B_1), \dots, N'(B_k)), \quad B_1, \dots, B_k \in \hat{\mathcal{S}}.$$

W przypadku, gdy $\mathbf{S} = \mathbb{R}^d$, wystarczy zdefiniować prawdopodobieństwa (15) dla prostokątów $(a, b] = \{x \in \mathbb{R}^d : a_i < x_i \leq b_i, 1 \leq i \leq d\}$.

Miarą intensywności albo miarą średnią procesu punktowego N jest

$$\mu(B) = E[N(B)], \quad B \in \mathcal{S}.$$

Miara $\mu(B)$ może być nieskończona, nawet jeśli B jest ograniczonym zbiorem. Jeżeli $\mathbf{S} = \mathbb{R}^d$, to czasami używamy zapisu postaci

$$\mu(B) = \int_B \lambda(x) dx,$$

gdzie $\lambda(x)$ jest intensywnością procesu N w punkcie x , a dx oznacza miarę Lebesgue'a. Funkcję $\lambda(x)$ nazywamy zależną od lokalizacji funkcją intensywności procesu N .

W naszych rozważaniach kluczową rolę będą odgrywać procesy punktowe Poissona.

DEFINICJA 6.2 Proces punktowy N na przestrzeni $\mathbf{S}$ jest procesem Poissona z miarą intensywności μ , która jest lokalnie skończona, jeżeli następujące warunki są spełnione:

- N ma niezależne przyrosty: zmienne losowe $N(B_1), \dots, N(B_k)$ są niezależne dla rozłącznych zbiorów $B_1, \dots, B_k \in \hat{\mathcal{S}}$.
- dla każdego $B \in \hat{\mathcal{S}}$, zmienna losowa $N(B)$ ma rozkład Poissona z wartością oczekiwaną $\mu(B)$.

W następnych paragrafach istotną rolę odegrają następujące dwa twierdzenia.

Twierdzenie 6.1 (Suma niezależnych procesów Poissona) *Niech $N_1, \dots, N_n$ będą niezależnymi procesami Poissona na $\mathbf{S}$ z miarami intensywności $\mu_1, \dots, \mu_n$, odpowiednio. Wtedy ich suma (superpozycja) $N = \sum_{i=1}^n N_i$ jest procesem Poissona z miarą intensywności $\mu = \sum_{i=1}^n \mu_i$. Twierdzenie jest także prawdziwe dla $n = \infty$, przy założeniu, że μ jest lokalnie skończona.*

Twierdzenie 6.2 (Przekształcenie procesu Poissona) *Niech $\mathbf{S}_i$, $i = 1, 2$, będą przestrzeniami polskimi. Niech $\mathbf{S}_i$, $i = 1, 2$, będą związanymi z nimi σ -algebrami. Niech $T: (\mathbf{S}_1, \mathcal{S}_1) \rightarrow (\mathbf{S}_2, \mathcal{S}_2)$ będzie mierzalnym przekształceniem. Jeżeli N jest procesem Poissona na $\mathbf{S}_1$ z miarą intensywności μ , to*

$$\check{N} = N \circ T^{-1}$$

jest procesem Poissona na $\mathbf{S}_2$ z miarą intensywności $\check{\mu} = \mu \circ T^{-1}$. Symbol „ $\circ$ ” oznacza superpozycję (złożenie).

Jeżeli proces punktowy N ma reprezentację

$$N = \sum_n \delta_{\tau_n},$$

to $\check{N}$ ma reprezentację

$$\check{N} = N \circ T^{-1} = \sum_n \delta_{T\tau_n}.$$

Na zakończenie paragrafu rozważmy klasyczny przypadek procesu Poissona na $\mathbb{R}_+$. Proces Poissona N na $\mathbb{R}_+$ (lub na $\mathbb{R}$) z miarą intensywności μ nazywamy niejednorodnym procesem Poissona. Oznaczmy przez $0 < \tau_1 \leq \tau_2 \leq \dots$ punkty lokalizacji procesu N . Wtedy $N(B) = \sum_n \delta_{\tau_n}(B)$ ma rozkład Poissona ze średnią $\mu(B)$.

Będziemy wykorzystywać zapis $N(t) = N((0, t])$ dla $t > 0$. Jeżeli

$$\mu(B) = \int_B \lambda(t) dt,$$

to mówimy, że N jest procesem Poissona z funkcją intensywności $\lambda(t)$. W przypadku, gdy $E[N(t)] = \lambda t$ dla ustalonego $\lambda > 0$, to N jest klasycznym procesem Poissona. Nazywamy go także jednorodnym lub stacjonarnym procesem Poissona z intensywnością λ .

Przypomnijmy, proces punktowy $N = \{N(t) : t \geq 0\}$ na $\mathbb{R}_+$ definiuje się jako proces liczący $N(t) = \sum_{n=1}^{\infty} \mathbf{1}_{\{\tau_n \leq t\}}$, gdzie $0 = \tau_0 \leq \tau_1 \leq \tau_2 \leq \dots$ są losowymi punktami (czasami) takimi, że $\tau_n \rightarrow \infty$ prawie na pewno, gdy

$n \rightarrow \infty$. Proces punktowy N nazywamy prostym, gdy jego punkty są różne, tzn. $0 = \tau_0 < \tau_1 < \tau_2 < \dots$ prawie na pewno.

Ważna jest następująca charakteryzacja jednorodnego procesu Poissona na $\mathbb{R}_+$. Przypomnijmy, że zmienna losowa X ma rozkład wykładniczy z intensywnością $\lambda > 0$, jeżeli $P(X > t) = e^{-\lambda t}$ dla wszystkich $t \geq 0$. Oczywiście $EX = \lambda^{-1}$.

TWIERDZENIE 6.3 *Niech $N(t)$ będzie prostym procesem punktowym na $\mathbb{R}_+$ z punktami $\tau_1 < \tau_2 < \dots$, przyjmujemy, że $\tau_0 = 0$. Proces $N(t)$ jest jednorodnym procesem Poissona na $\mathbb{R}_+$ z intensywnością λ , wtedy i tylko wtedy, gdy różnice $\tau_n - \tau_{n-1}$, $n = 1, 2, \dots$, są niezależnymi zmiennymi losowymi o rozkładzie wykładniczym z wartością oczekiwaną λ^{-1} .*

7. Procesy Markowa

Przyjmujemy, że czytelnik zna aparat procesów Markowa (zob. np. [27], [28], [13]). Tutaj wprowadzimy oznaczenia i przypomnimy jedynie wybrane pojęcia i wyniki.

Rozważmy zbiory $T, S \subset \mathbb{R}$, gdzie $\mathbb{R}$ jest zbiorem liczb rzeczywistych. Proces stochastyczny $X = \{X_t, t \in T\}$, gdzie $X_t \in S$, nazywamy procesem Markowa, jeżeli dla każdego $t \in T$ ciąg X_s , $s \leq t$ i zmienna losowa X_u dla $u \geq t$ są warunkowo niezależne względem zmiennej losowej X_t . To znaczy dla każdego zbioru borelowskiego $B \subset S$ prawie na pewno mamy

$$P(X_u \in B | X_s, s \leq t) = P(X_u \in B | X_t)$$

dla $u \geq t$, gdzie $P(\cdot | \cdot)$ jest prawdopodobieństwem warunkowym.

Zbiór T nazywamy czasem (skalą czasu), a zbiór S — zbiorem wartości procesu.

Dla każdych $s, t \in T$, przy czym $s \leq t$, istnieją regularne warunkowe rozkłady (zob. np. [27]) takie, że prawie na pewno

$$K_{s,t}(X_s, B) = P(X_t \in B | X_s) = P(X_t \in B | X_u, u \leq s),$$

gdzie $B \subset \mathbb{R}$ są zbiorami borelowskimi.

Rozkłady $K_{s,t}(X_s, B)$ nazywamy jądrami markowskimi (przejsć) na S .

W przypadku, gdy $T = \mathbb{R}_+$ albo $T = \mathbb{Z}_+$, mówimy, że mamy proces Markowa jednorodny w czasie, jeżeli jądra przejść są postaci $K_{s,t} = K_{t-s}$, to znaczy

$$P(X_t \in B | X_s) = K_{t-s}(X_s, B)$$

prawie na pewno, gdzie $B \subset S$ jest zbiorem borelowskim, $s, t \in T$, $s \leq t$.

Wtedy, wprowadzając rozkład początkowy $\nu_{t_0} = P \circ X_0^{-1}$, rozkład skończenie wymiarowy procesu X jest postaci

$$P(X_{t_0} \in B_0, X_{t_1} \in B_1, \dots, X_{t_n} \in B_n) = \int_{B_0} d\nu_{t_0}(x_0) \int_{B_1} K_{t_1-t_0}(x_0, dx_1) \cdots \int_{B_n} K_{t_n-t_{n-1}}(x_{n-1}, dx_n). \quad (16)$$

Relacje Chapmana–Kołmogorowa tożsamościowo postać

$$K_{s+u} = K_s K_u \quad s \leq t \leq u.$$

Rodzinę (K_t) nazywamy półgrupą jąder przejścia.

Ustalmy półgrupę jąder (K_t) na przestrzeni $S \subset T$. Wtedy z Twierdzenia Kołmogorowa o istnieniu (zob. np. [27]) istnieje proces Markowa X_ν związany z tą półgrupą i rozkładem początkowym ν . Wiadomo wtedy, że istnieje jednoznacznie wyznaczona przez ν miara P_ν na przestrzeni trajektorii $(S^T, \mathcal{S}^T)$. Dla miary skupionej δ_x w punkcie $x \in S$ będziemy pisać P_x zamiast P_{δ_x} . Wartości oczekiwane (całkowanie) względem P_ν albo P_x będziemy oznaczać przez E_ν albo E_x , odpowiednio.

8. Procesy Markowa rekordów. Rozkład k -wartości rekordowej

Niech $X_1, X_2, \dots$ będą niezależnymi zmiennymi losowymi o jednakowej (absolutnie) ciągłej dystrybucji F z gęstością f .

Nasze rozważania rozpoczynamy od znalezienia łącznej gęstości rozkładu wektorów statystyk pozycyjnych (zob. np. [9])

$$(X_{1:n}, X_{2:n}, \dots, X_{n:n}), \quad n \geq 1.$$

Gęstość tego wektora dla ustalonego n jest postaci

$$n! f(x_1) \cdots f(x_n), \quad x_1 < \cdots < x_n. \quad (17)$$

Ustalmy $k \in \mathbb{Z}_+$. Rozważmy ciąg wektorów statystyk rekordowych

$$\mathbf{X}(n, k) = (X(n, k), X(n, k-1), \dots, X(n, 1)), \quad n \geq 1, \quad (18)$$

gdzie $X(n, k) = X_{L(n, k) - k + 1 : L(n, k)}$.

Zauważmy, że ciąg (18) jest jednorodnym procesem Markowa z czasem $\mathbf{T} = \mathbb{Z}_+$. Zauważmy, że znając stan procesu $\mathbf{X}(t, k)$ w chwili $t = n$, przyszła jego ewolucja w czasie zależy jedynie od ciągu $X_{n+1}, X_{n+2}, \dots$, zatem jest ona niezależna od przeszłości opisanej przez $X_1, \dots, X_{n-1}$.

Rozważmy najpierw ciąg rekordów $X(n, 1)$, $n \geq 1$. Niech $f_n(x)$ będzie gęstością rekordu $X(n, 1)$. Mamy

$$f_1(x) = f(x), \quad x \in \mathbb{R},$$

oraz

$$f_{n+1}(x) = \int_0^x f_n(y) f_{n+1}(x|y) dy,$$

gdzie $f_{n+1}(x|y)$ jest warunkową gęstością $X(n+1, 1)$ przy danym $X(n, 1) = y$.
Następnie

$$f_{n+1}(x|y) = \frac{f(x)}{1 - F(y)}, \quad y < x,$$

(zob. także [31], wzór (15.8)).

Ostatecznie otrzymujemy

$$f_{n+1}(x) = \begin{cases} \int_0^x f_n(y) \frac{f(x)}{1 - F(y)} dy, & \text{gdy } y < x, \\ 0, & \text{w przeciwnym razie.} \end{cases}$$

Rozważmy teraz ciąg dwuwymiarowych wektorów

$$(X(n, 2), X(n, 1)), \quad n \geq 1.$$

Wtedy z (17)

$$f_1(x_1, x_2) = \begin{cases} 2! f(x_1) f(x_2), & \text{gdy } x_1 < x_2, \\ 0, & \text{w przeciwnym razie.} \end{cases}$$

Obliczymy $f_{n+1}(x_1, x_2)$. Niech $X(n, 2) = y$. Rozważymy dwa przypadki:

1) kolejna w wyjściowym ciągu zmienna losowa X , która powoduje zmianę drugiej wartości rekordowej jest drugą maksymalną statystyką pozycyjną, tzn. drugą co do wielkości wartością w zaobserwowanym ciągu; $X = X_{-1, \cdot}$,

2) kolejna w wyjściowym ciągu zmienna losowa X , która powoduje zmianę drugiej wartości rekordowej jest wartością maksymalną.

Mamy wtedy, rozumując jak poprzednio,

$$f_{n+1}(x_1, x_2) = \begin{cases} \sum_{i=1}^2 \int_{-\infty}^{x_1} f_n(y, x_i) \frac{f(x_i)}{1 - F(y)} dy, & \text{gdy } x_1 < x_2, \\ 0, & \text{w przeciwnym razie.} \end{cases}$$

Ostatecznie obliczamy $f_{n+1}(x_1, x_2, \dots, x_k)$. Niech $X(n, k) = y$. Rozważymy teraz k przypadków: kolejna w wyjściowym ciągu zmienna losowa X , która powoduje zmianę k -wartości rekordowej jest: 1) k maksymalną statystyką pozycyjną, 2) $k-1$ maksymalną statystyką pozycyjną, ..., k) pierw-

szą maksymalną statystyką pozycyjną, czyli maksimum. Wtedy, dla $n \geq 1$,

$$f_{n+1}(x_1, x_2, \dots, x_k) = \begin{cases} \sum_{i=1}^k \int_{-\infty}^{x_i} f_n(y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k) \frac{f(x_i)}{1-F(y)} dy, \\ 0, \end{cases} \begin{matrix} \text{gdy } x_1 < x_2 < \dots < x_k, \\ \text{w przeciwnym razie,} \end{matrix}$$

gdzie

$$f_n(y, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_k) \frac{f(x_i)}{1-F(y)}$$

dla $k = 1$, wynosi

$$f_n(y, x_2, \dots, x_k) \frac{f(x_1)}{1-F(y)}.$$

Wykorzystując powyższe równania rekurencyjne, w następnym twierdzeniu podamy gęstość k -tej wartości rekordowej.

Twierdzenie 8.1 *Gęstość $f_n^{(k)}$ k -tej wartości rekordowej jest postaci*

$$f_n^{(k)}(y) = \frac{k}{(n-1)!} [-k \log(1-F(y))]^{n-1} (1-F(y))^{k-1} f(y), \quad n \geq 1.$$

Dowód. Zauważmy, że gęstości $f_n(x_1, x_2, \dots, x_k)$, $n \geq 1$ spełniają równania

$$f_n(x_1, x_2, \dots, x_k) = \begin{cases} k! g_n(x_1) f(x_1) f(x_2) \dots f(x_k), & \text{gdy } x_1 < x_2 < \dots < x_k, \\ 0, & \text{w przeciwnym razie,} \end{cases} \quad (19)$$

gdzie

$$g_1(z) = 1, \\ g_{n+1}(z) = k \int_{-\infty}^z g_n(y) \frac{f(y)}{1-F(y)} dy, \quad n \geq 1.$$

Stąd

$$g_n(z) = \frac{1}{(n-1)!} [-k \log(1-F(z))]^{n-1}, \quad n \geq 1.$$

Zatem mamy już postać łącznych gęstości $f_n(x_1, x_2, \dots, x_k)$, $n \geq 1$, wektorów

$$(X(n, k), X(n, k-1), \dots, X(n, 1)), \quad n \geq 1.$$

Możemy zatem znaleźć gęstości $f_n^{(k)}(y)$, $k \geq 1$, $n \geq 1$, k -tych wartości rekordowych $\{X(n, k), k \geq 1, n \geq 1\}$.

Z (19) mamy

$$\begin{aligned} f_n^{(k)}(y) &= \int_y^\infty \int_{x_2}^\infty \cdots \int_{x_{k-1}}^\infty \frac{k!}{(n-1)!} [-k \log(1-F(y))]^{n-1} \times \\ &\quad \cdot f(y) f(x_2) \cdots f(x_k) dx_k \dots dx_2 = \\ &= \frac{k}{(n-1)!} [-k \log(1-F(y))]^{n-1} (1-F(y))^{k-1} f(y), \end{aligned}$$

to kończy dowód Twierdzenia 8.1.

Korzystając z Twierdzenia 8.1, łatwo jest znaleźć dystrybuantę k -tej wartości rekordowej $F_n^{(k)}(x) = P(X(n, k) \leq x)$.

WNIOSEK 8.1 Dla $k \geq 1$, $n \geq 1$ mamy

$$F_n^{(k)}(x) = \int_0^{-k \log(1-F(x))} \frac{y^{n-1}}{(n-1)!} e^y dy.$$

W przypadku wartości rekordowej $X(n, 1)$, tzn. dla $k = 1$, mamy

$$f_n^{(1)}(y) = \frac{1}{(n-1)!} [-k \log(1-F(y))]^{n-1} f(y), \quad n \geq 1.$$

oraz

$$F_n^{(1)}(x) = \int_0^{-\log(1-F(x))} \frac{y^{n-1}}{(n-1)!} e^y dy, \quad n \geq 1.$$

Rozważmy dwa ciągi niezależnych o jednakowym rozkładzie zmiennych losowych: $X_1, X_2, \dots$ z ciągłą dystrybuantą F i

$$Y_1 = \min\{X_1, \dots, X_k\}, \quad Y_2 = \min\{X_{k+1}, \dots, X_{2k}\}, \dots$$

z dystrybuantą

$$G(x) = 1 - (1 - F(x))^k.$$

Niech $Y(n, 1)$ oznaczają wartości rekordowe w ciągu $Y_1, Y_2, \dots$

Twierdzenie 8.2 Dla ustalonego $k \geq 1$, mamy

$$X(n, k) \stackrel{d}{=} Y(n, 1), \quad n \geq 1.$$

Dowód. Zauważmy, że

$$P(X(n, k) \leq x) = \int_0^{-k \log(1-F(x))} \frac{y^{n-1}}{(n-1)!} e^y dy$$

oraz

$$P(Y(n, 1) \leq x) = \int_0^{-k \log(1-F(x))} \frac{y^{n-1}}{(n-1)!} e^y dy.$$

To kończy dowód twierdzenia 8.2.

9. Proces wartości rekordowych

Niech $E_1, E_2, \dots$ będzie ciągiem niezależnych zmiennych losowych z jednokową standardową dystrybuantą wykładniczą

$$F(x) = \begin{cases} 1 - \exp(-x) & x > 0, \\ 0, & x \leq 0. \end{cases}$$

Pokazaliśmy poprzednio, że dystrybuanta $F_{n,E}^{(1)}(x)$ wartości rekordowej $E(n, 1)$, $n \geq 1$, w ciągu $E_1, E_2, \dots$ jest postaci

$$F_{n,E}^{(1)}(x) = \int_0^x \frac{y^{n-1}}{(n-1)!} e^y dy, \quad n \geq 1.$$

Zatem w tym przypadku wartość rekordowa ma $E(n, 1)$ na rozkład gamma z parametrem n , czyli

$$E(n, 1) \stackrel{d}{=} E_1 + E_2 + \dots + E_n, \quad n \geq 1.$$

Rozważmy teraz proces wartości rekordowych

$$E(1, 1), E(2, 1), \dots, E(n, 1), \dots \quad (20)$$

Z rozważań z Rozdziału 8 wynika, że ciąg (20) jest jednorodnym procesem Markowa z czasem $\mathbf{T} = \mathbb{Z}_+$ dla którego

$$P(E(n+1, 1) > x | E(n, 1) = y) = e^{-(x-y)}$$

dla $n \geq 1$ i $x > y$.

Rzeczywiście, wiemy, że

$$f_{n+1,E}(x|y) = \frac{e^{-x}}{e^{-y}} = e^{-(x-y)}, \quad y < x,$$

$f_{n+1,E}(x|y)$ jest warunkową gęstością $E(n+1, 1)$ przy danym $E(n, 1) = y$ oraz $f_{n,E}(x)$ jest gęstością rekordu $E(n, 1)$.

Zatem

$$\begin{aligned} P(E(n+1, 1) > x | E(n, 1) = y) &= \\ &= \int_x^\infty f_{n+1,E}(u|y) du = \int_x^\infty e^{-(u-y)} du = e^{-(x-y)}, \quad y < x. \end{aligned}$$

Następne twierdzenie odgrywa kluczową rolę w teorii procesów wartości rekordowych.

Twierdzenie 9.1 *Zmienne losowe $E(n, 1)$, $n \geq 1$, są punktami jednorodnego procesu punkowego Poissona na $(0, \infty)$. Oznacza to, że*

$$\{E(n, 1), n \geq 1\} \stackrel{d}{=} \{\Gamma_n, n \geq 1\} \quad (21)$$

w $\mathbb{R}^\infty$, gdzie $\Gamma_n = E_1 + E_2 + \dots + E_n$ i $\{E_j, j \geq 1\}$ jest ciągiem niezależnych zmiennych losowych z jednakową standardową dystrybucją wykładniczą.

Dowód. Zauważmy, że

$$P(E(n+1, 1) > x \mid E(n, 1) = y) = e^{-(x-y)} = P(\Gamma_{n+1} > x \mid \Gamma_n = y)$$

dla $y < x$.

Pamiętamy także, że $X(1, 1) = X_1 \stackrel{d}{=} \Gamma_1$.

Wiadomo, że dwa ciągi Markowa ze stacjonarnymi prawdopodobieństwami przejść są równe według rozkładu, jeżeli ich wyjściowe rozkłady i prawdopodobieństwa przejść są sobie równe, odpowiednio. Zatem

$$\{E(n, 1), n \geq 1\} \stackrel{d}{=} \{\Gamma_n, n \geq 1\}.$$

Otrzymaliśmy więc tezę (21) Twierdzenia 9.1.

10. Twierdzenie Ignatova–Deheuvelsa Przypomnijmy, sekwencyjna ranga I_n to ranga X_n wśród $X_1, X_2, \dots, X_n$, tzn.

$$X_n = X_{I_n:n}, n \geq 1$$

albo

$$I_n = \sum_{i=1}^n \mathbf{1}_{\{X_i \leq X_n\}}, n \geq 1,$$

gdzie $\mathbf{1}_A$ oznacza indykator zdarzenia losowego A .

Czasy rekordowe można zdefiniować wzorem

$$\begin{aligned} L(1, 1) &= 1, \\ L(n+1, 1) &= \min \{j > L(n, 1) : I_j = n\}, n \geq 1. \end{aligned}$$

Wprowadziliśmy poprzednio tzw. ścisłe czasy k -rekordowe przyjmując odpowiednio, dla $k = 1, 2, \dots$,

$$\begin{aligned} L_s(1, k) &= k \\ L_s(n+1, k) &= \min \{j > L_s(n, k) : I_j = n - k + 1\}, n \geq 1. \end{aligned}$$

i ścisłe k -rekordy,

$$X_s(n, k) = X_{L_s(n, k) - k + 1 : L_s(n, k)}, n \geq 1.$$

Wtedy

$$X(n, 1) = X_s(n, 1), \quad n \geq 1.$$

Niech $k \geq 1$ będzie ustaloną liczbą naturalną. Definiujemy procesy liczące punkty związane z procesami k -tych rekordów.

Niech

$$\{X(1, k), X(2, k), \dots, X(n, k), \dots\}$$

będzie procesem punktowym k -tych rekordów, a

$$N^k(x) = \sum_{n=1}^{\infty} \mathbf{1}_{\{X(n, k) \leq x\}}$$

związujemy z nim procesem liczącym punkty. Podobnie, niech

$$\{X_s(1, k), X_s(2, k), \dots, X_s(n, k), \dots\}$$

będzie procesem punktowym ścisłych k -tych rekordów, a

$$N_s^k(x) = \sum_{n=1}^{\infty} \mathbf{1}_{\{X_s(n, k) \leq x\}}$$

związujemy z nim procesem liczącym punkty.

Możemy już sformułować Twierdzenie Ignatova–Deheuvelsa.

Twierdzenie 10.1 *Procesy ścisłych k -rekordów N_s^k , $k = 1, 2, \dots$ są niezależne i mają taki sam rozkład jak proces rekordów N^1 .*

Istnieją kontrowersje, do kogo należy wynik przedstawiony w Twierdzeniu 10.1 (zob. [11]). Twierdzenie 10.1, w przypadku ciągłej wyjściowej dystrybucyjnej F , zostało udowodnione przez P. Deheuvelsa w pracy [10] przesłanej do druku 10 stycznia 1982 roku. Niezależnie, Z. Ignatov [24], [25] trochę wcześniej przedstawił bez dowodu omawiane twierdzenie.

W następnych latach pojawiły się prace podające dowody Twierdzenia Ignatova–Deheuvelsa nie tylko dla ciągłych dystrybucyj, ale i dla dowolnych dystrybucyj.

Wyliczmy niektóre z nich w kolejności czasu publikacji: C. M. Goldie and L. C. G. Rogers (1984) [23], A. J. Stam (1985) [43], S. I. Resnick (1987) [36], Rozdział 4.6, R. Engelen, P. Tommassen and W. Vervaat (1988) [18], L. C. G. Rogers (1989) [38], S. M. Samuels (1992) [39], Y-C Yao (1997) [49].

W tym opracowaniu nie podajemy dowodu Twierdzenia 10.1 (Ignatova–Deheuvelsa). Wymagałoby to osobnego opracowania. Zwróćmy jedynie uwagę, że pojawiają się interesujące zastosowania tego twierdzenia. Zobaczmy, na przykład: G. Simons, Y-C. Yao L. Yang (2002) [42], A. Gnedin (2008) [22], B. Buchmann, R. Maller and S. I. Resnick (2018) [5].

W następnym paragrafie potrzebna będzie nam wersja Twierdzenia Ignatova–Deheuvelsa dla standardowego rozkładu wykładniczego.

Niech $E_1, E_2, \dots$ będzie ciągiem niezależnych zmiennych losowych z jednakową standardową dystrybucją wykładniczą. Niech

$$\{E_s(1, k), E_s(2, k), \dots, E_s(n, k), \dots\}$$

będzie procesem punktowym ścisłych k -tych rekordów w ciągu $E_1, E_2, \dots$, a

$$N_{s,E}^k(x) = \sum_{n=1}^{\infty} \mathbf{1}_{\{E_s(n,k) \leq x\}}$$

związany z nim procesem liczącym punkty. Podobnie dla procesu punkowego k -tych rekordów

$$\{E(1, k), E(2, k), \dots, E(n, k), \dots\}$$

przez

$$N_E^k(x) = \sum_{n=1}^{\infty} \mathbf{1}_{\{E(n,k) \leq x\}}$$

oznaczamy związany z nim proces liczący punkty.

Z Twierdzenia 9.1 wynika, że zmienne losowe $\{E(n, 1), n \geq 1\}$ są punktami jednorodnego standardowego procesu punkowego Poissona na $(0, \infty)$, tzn. z intensywnością $\lambda = 1$. Zatem wnioskiem z Twierdzenia 10.1 jest następujące twierdzenie.

Twierdzenie 10.2 *Procesy ścisłych k -rekordów $N_{s,E}^k, k = 1, 2, \dots$, są niezależnymi jednorodnymi standardowymi procesami punktowymi Poissona na $(0, \infty)$.*

11. Proces k -tych wartości rekordowych

Do konstrukcji procesu k -tych wartości rekordowych wykorzystamy rozumowanie P. Deheuvelsa z pracy [11].

Rozważmy, dla $k = 1, 2, \dots$, ciąg k -tych czasów rekordowych

$$\{L(1, k), L(2, k), \dots\}$$

oraz ciąg ścisłych k -tych czasów rekordowych

$$\{L_s(1, k), L_s(2, k), \dots\}.$$

Zauważmy, że

$$\{L(j, k), j \geq 1\} = \bigcup_{m=1}^k \{L_s(j, m), j \geq 1\}. \quad (22)$$

Z równania (22) wynika, że proces punktowy k -tych rekordów

$$N_E^k(x) = \sum_{n=1}^{\infty} \mathbf{1}_{\{E(n,k) \leq x\}}$$

jest superpozycją procesów punktowych ścisłych m -tych rekordów

$$N_{s,E}^m(x) = \sum_{n=1}^{\infty} \mathbf{1}_{\{E_s(n,m) \leq x\}},$$

gdzie $m = 1, 2, \dots, k$.

Możemy już teraz udowodnić twierdzenie o procesie k -tych rekordów dla próby z rozkładu wykładniczego.

Twierdzenie 11.1 *Procesy k -rekordów N_E^k , $k = 1, 2, \dots$, są jednorodnymi procesami punktowymi Poissona na $(0, \infty)$ z intensywnościami $\lambda(x) = k$, $x \in (0, \infty)$, odpowiednio.*

Dowód. Korzystając z Twierdzeń 10.2 i 6.1 mamy, że $N_{s,E}^k$ jest procesem Poissona z intensywnością

$$\lambda(x) = \sum_{m=1}^k 1 = k, \quad x \in (0, \infty).$$

To kończy dowód Twierdzenia 11.1.

W celu znalezienia postaci procesu punktowego k -tych rekordów wprowadzamy pojęcie odwrotności monotonicznej funkcji i przypomnimy kilka jej własności (zob. np. [36], Rozdział 0.2. i Rozdział 4.1.).

Założmy, że H jest niemalejącą funkcją na $\mathbb{R}$. Przyjmujemy konwencję, że infimum zbioru pustego jest równe $+\infty$. Definiujemy (lewostronnie ciągłą) odwrotność H jako

$$H^{\leftarrow}(y) = \inf \{s : H(s) \geq y\}.$$

Niech X będzie zmienną losową o dystrybucanie F . Określmy przekształcenie

$$R(x) = -\log(1 - F(x)), \quad x \in \mathbb{R}.$$

Niech E będzie zmienną losową o standartowym rozkładzie wykładniczym. Przypomnijmy, że wtedy $P(E > x) = e^{-x}$, $x > 0$. Zmienne losowe $R^{\leftarrow}(E)$ i X mają ten sam rozkład prawdopodobieństwa, co zapisujemy

$$R^{\leftarrow}(E) \stackrel{d}{=} X.$$

Niech $X_1, X_2, \dots$ będzie ciągiem niezależnych zmiennych losowych z jednakową dystrybuantą F oraz $E_1, E_2, \dots$ będzie ciągiem niezależnych zmiennych losowych z jednakową standardową dystrybuantą wykładniczą. Oznaczmy $M_n = \max(X_1, X_2, \dots, X_n)$, $M_{n,E} = \max(E_1, E_2, \dots, E_n)$. Wtedy

$$\{\max(R^{\leftarrow}(E_1), R^{\leftarrow}(E_2), \dots, R^{\leftarrow}(E_n)), n \geq 1\} \stackrel{d}{=} \{M_n, n \geq 1\}.$$

Ponieważ $R^{\leftarrow}$ jest funkcją niemalejącą, więc

$$\{\max(R^{\leftarrow}(E_1), R^{\leftarrow}(E_2), \dots, R^{\leftarrow}(E_n)), n \geq 1\} \stackrel{d}{=} \{R^{\leftarrow}(M_{n,E}), n \geq 1\}.$$

W rozważanym w tej pracy przypadku, gdy F jest dystrybuantą ciągłą, funkcja R jest ciągła i stąd $R^{\leftarrow}$ jest ściśle rosnącą funkcją. Wtedy, dla $k = 1, 2, \dots$, mamy

$$\{R^{\leftarrow}(E(n, k)), n \geq 1\} \stackrel{d}{=} \{X(n, k), n \geq 1\}. \quad (23)$$

Pamiętamy, że $\{E(n, k), n \geq 1\}$ jest jednorodnym procesem Poissona na $(0, \infty)$ z intensywnością $\lambda(x) = k$, $x \in (0, \infty)$ (zobacz Twierdzenie 11.1). Następny wynik o własnościach procesu k -tych rekordów sformułujemy jako twierdzenie.

Twierdzenie 11.2 *Niech F będzie ciągłą dystrybuantą. Określmy*

$$R(x) = -\log(1 - F(x))$$

i połóżmy $x_l = \inf\{y : F(y) > 0\}$, $x_0 = \sup\{y : F(y) < 1\}$. Ponieważ

$$R: (x_l, x_0) \rightarrow (0, \infty),$$

więc $R^{\leftarrow}: (0, \infty) \rightarrow (x_l, x_0)$.

Wtedy procesy k -rekordów N^k , $k = 1, 2, \dots$, są procesami punktowymi Poissona na (x_l, x_0) ze średnią miarą

$$E[N^k((a, b))] = kR(b) - kR(a) \quad (24)$$

dla $x_l < a < b < x_0$.

Dowód. Wykorzystamy Twierdzenie 6.2 o przekształceniu procesu Poissona. Rozważmy proces Poissona

$$N_E^k = \sum_{n=1}^{\infty} \delta_{E(n,k)}$$

na $(0, \infty)$ z intensywnością $\lambda(x) = k$, $x \in (0, \infty)$ i przekształcenie $R^{\leftarrow}$. Wtedy

$$\check{N}_E^k = N_E^k \circ (R^{\leftarrow})^{-1} = \sum_{n=1}^{\infty} \delta_{R^{-1}(E(n,k))}.$$

jest procesem Poissona na (x_l, x_0) ze średnią miarą

$$\dot{\mu} = (km) \circ (R^{\leftarrow})^{-1} \text{ na } (x_l, x_0), \quad (25)$$

gdzie m jest miarą Lebesguea.

Ze wzoru (23) mamy

$$N^k = \sum_{n=1}^{\infty} \delta_{X(n,k)} = \sum_{n=1}^{\infty} \delta_{R^{-1}(E(n,k))}$$

oraz z (25) otrzymujemy (24).

To kończy dowód Twierdzenia 11.2.

12. Uogólnienie reprezentacji Tata na przypadek k -tych wartości rekordowych

Zauważmy, że dla $n = 1, 2, \dots$ i $k = 1, 2, \dots$ mamy

$$(X(1, k), \dots, X(n, k)) \stackrel{d}{=} (R^{\leftarrow}(E(1, k)), \dots, R^{\leftarrow}(E(n, k))),$$

gdzie $R^{\leftarrow}(x) = F^{-1}(1 - e^{-x})$. Przypominamy, po raz kolejny, że w tej części pracy zakładamy ciągłość dystrybucyjności F .

Rozważymy najpierw przypadek $k = 1$. Z Twierdzenia 9.1 wynika, że

$$\{E(n, 1), n \geq 1\} \stackrel{d}{=} \{\Gamma_n, n \geq 1\}$$

w $\mathbb{R}^{\infty}$, gdzie

$$\Gamma_n = E_1 + E_2 + \dots + E_n$$

i $\{E_j, j \geq 1\}$ jest ciągiem niezależnych zmiennych losowych z jednakową standardową dystrybucją wykładniczą.

Zatem

$$(X(1, 1), \dots, X(n, 1)) \stackrel{d}{=} (R^{\leftarrow}(\Gamma_1), \dots, R^{\leftarrow}(\Gamma_n)), \quad (26)$$

Równanie (26) nazywamy reprezentacją Tata (zob. [46])

Dla ustalonego $k = 1, 2, \dots$, zmienne losowe

$$\{X(1, k), X(2, k), \dots, X(n, k), \dots\}$$

są punktami procesu punktowego k -tych rekordów N^k . Proces k -rekordów N^k jest procesem punktowym Poissona na (x_l, x_0) ze średnią miarą

$$E[N^k((a, b])] = kR(b) - kR(a),$$

gdzie $R(x) = -\log(1 - F(x))$, $x \in \mathbb{R}$. K -te rekordy tworzą proces Markowa z czasem $\mathbf{T} = \mathbb{Z}_+$ dla którego (zobacz np. [1])

$$P(X(n+1, k) > x | X(n, k) = y) = \left(\frac{1 - F(x)}{1 - F(y)} \right)^k$$

dla $n \geq 1$ i $x > y$. W przypadku standardowego rozkładu wykładniczego mamy

$$P(E(n+1, k) > x | E(n, k) = y) = e^{-k(x-y)}$$

dla $n \geq 1$ i $x > y$. Zatem

$$\{E(1, k), E(2, k), \dots, E(n, k), \dots\}$$

jest jednorodnym procesem Poissona z intensywnością λ .

Ostatecznie

$$(X(1, k), \dots, X(n, k)) \stackrel{d}{=} \left(\frac{1}{k} (R^{\leftarrow}(\Gamma_1), \dots, R^{\leftarrow}(\Gamma_n)) \right), \quad (27)$$

Równanie (27) nazywamy uogólnioną reprezentacją Tata (zob. np. [31], a także [17]).

13. Zakończenie

W poprzednich paragrafach przedstawiliśmy jedynie część wyników dotyczących teorii k -tych wartości rekordowych. Pominęliśmy prawie zupełnie wyniki o k -tych wartościach rekordowych w próbach z dyskretnym rozkładem. Nie wspomnieliśmy nic o dobrze rozwiniętej asymptotycznej teorii k -tych wartości rekordowych. Rozwija się także teoria tzw. k -tych procesów ekstremalnych i procesów punktowych związanych z nimi, którą można wykorzystać do badania k -tych rekordów.

Zwróćmy uwagę, że pojawiają się w znacznej ilości artykuły dotyczące wnioskowań startystycznych w oparciu o obserwowane k -te wartości rekordowe. Powstają także interesujące zastosowania w teorii niezawodności, w badaniach dotyczących ocieplenia klimatu, czy też w teorii systemów dynamicznych.

Nawet pobieżne przedstawienia któregoś z powyższych zagadnień znacznie przekracza ramy prezentowanego artykułu.

14. Podziękowanie

Dziękuję Profesorowi Władysławowi Szczotce z Instytutu Matematycznego Uniwersytetu Wrocławskiego za zachętę do napisania pracy i uwagi, które pozwoliły ją ulepszyć.

Literatura

- [1] M. Ahsanullah and V. B. Nevzorov. Records statistics. In M. Lovric, editor, *International Encyclopedia of Statistical Science*, pages 1195–1202. Springer, 2011.
- [2] M. Ahsanullah and V. B. Nevzorov. *Records via probability theory*. Amsterdam: Atlantis Press, 2015. ISBN 978-94-6239-135-2/hbk; 978-94-6239-136-9/ebook. doi: [10.2991/978-94-6239-136-9](https://doi.org/10.2991/978-94-6239-136-9).

- [3] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *Records*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, 1998. ISBN 0-471-08108-6. doi: [10.1002/9781118150412](https://doi.org/10.1002/9781118150412). A Wiley-Interscience Publication.
- [4] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *A first course in order statistics*, volume 54 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008. ISBN 978-0-89871-648-1. doi: [10.1137/1.9780898719062](https://doi.org/10.1137/1.9780898719062). Unabridged republication of the 1992 original.
- [5] B. Buchmann, R. Maller, and S. I. Resnick. Processes of r -th largest. *Extremes*, Feb 2018. ISSN 1572-915X. doi: [10.1007/s10687-018-0308-x](https://doi.org/10.1007/s10687-018-0308-x).
- [6] K. N. Chandler. The distribution and frequency of record values. *J. Roy. Statist. Soc. Ser. B.*, 14:220–228, 1952. ISSN 0035-9246. URL [http://links.jstor.org/sici?sici=0035-9246\(1952\)14:2<220:TDAFOR>2.0.CO;2-D&origin=MSN](http://links.jstor.org/sici?sici=0035-9246(1952)14:2<220:TDAFOR>2.0.CO;2-D&origin=MSN).
- [7] C. A. Charalambides. Exact distributions of the number of r -records and the r -record and inter- r -record times. *Comm. Statist. Theory Methods*, 36(5-8): 1305–1317, 2007. ISSN 0361-0926. doi: [10.1080/03610920601076974](https://doi.org/10.1080/03610920601076974).
- [8] C. A. Charalambides. Discrete q -distributions on Bernoulli trials with a geometrically varying success probability. *J. Statist. Plann. Inference*, 140(9):2355–2383, 2010. ISSN 0378-3758. doi: [10.1016/j.jspi.2010.03.024](https://doi.org/10.1016/j.jspi.2010.03.024).
- [9] H. A. David and H. N. Nagaraja. *Order statistics*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, third edition, 2003. ISBN 0-471-38926-9. doi: [10.1002/0471722162](https://doi.org/10.1002/0471722162).
- [10] P. Deheuvels. The strong approximation of extremal processes. II. *Z. Wahrsch. Verw. Gebiete*, 62(1):7–15, 1983. ISSN 0044-3719. doi: [10.1007/BF00532159](https://doi.org/10.1007/BF00532159).
- [11] P. Deheuvels. Strong approximations of k th records and k th record times by Wiener processes. *Probab. Theory Related Fields*, 77(2):195–209, 1988. ISSN 0178-8051. doi: [10.1007/BF00334037](https://doi.org/10.1007/BF00334037).
- [12] A. Dembińska and F. López-Blázquez. k th records from discrete distributions. *Statist. Probab. Lett.*, 71(3):203–214, 2005. ISSN 0167-7152. doi: [10.1016/j.spl.2004.11.001](https://doi.org/10.1016/j.spl.2004.11.001).
- [13] Dembo, A. . Stochastic processes. Technical report, Department of Statistics, Stanford University, 2013. URL <http://statweb.stanford.edu/~adembo/math-136/mnotes.pdf>. Revised by K. Ross. Online: accessed 11 May of 2018.
- [14] M. Dwass. Some k -sample rank-order tests. In *Contributions to probability and statistics*, pages 198–202. Stanford Univ. Press, Stanford, Calif., 1960. Essays in Honour of H. Hotelling.
- [15] W. Dziubdziela. Limit distributions of extreme order statistics. *Mat. Stos. (3)*, 9:45–71, 1977. ISSN 0137-2890.

-
- [16] W. Dziubdziela. On record times and record values in a sequence of random variables. *Wiadom. Mat.*, 29(1):57–70, 1990. ISSN 0373-8302.
- [17] W. Dziubdziela and B. Kopociński. Limiting properties of the k -th record values. *Zastos. Mat.*, 15(2):187–190, 1976. ISSN 0044-1899.
- [18] R. Engelen, P. Tommassen, and W. Vervaat. Ignatov’s theorem: a new and short proof. *J. Appl. Probab.*, Special Vol. 25A:229–236, 1988. ISSN 0021-9002. A celebration of applied probability.
- [19] F. G. Foster and A. Stuart. Distribution-free tests in time-series based on the breaking of records. *J. Roy. Statist. Soc. Ser. B.*, 16:1–13; discussion 13–22, 1954. ISSN 0035-9246.
- [20] F. G. Foster and D. Teichroew. A sampling experiment on the powers of the records tests for trend in a time series. *J. Roy. Statist. Soc. Ser. B.*, 17:115–121, 1955. ISSN 0035-9246.
- [21] N. Glick. Breaking records and breaking boards. *Amer. Math. Monthly*, 85(1): 2–26, 1978. ISSN 0002-9890. doi: [10.2307/2978044](https://doi.org/10.2307/2978044).
- [22] A. V. Gnedin. Corners and records of the Poisson process in quadrant. *Electron. Commun. Probab.*, 13:187–193, 2008. ISSN 1083-589X. doi: [10.1214/ECP.v13-1351](https://doi.org/10.1214/ECP.v13-1351).
- [23] C. M. Goldie and L. C. G. Rogers. The k -record processes are i.i.d. *Z. Wahrsch. Verw. Gebiete*, 67(2):197–211, 1984. ISSN 0044-3719. doi: [10.1007/BF00535268](https://doi.org/10.1007/BF00535268).
- [24] Z. Ignatov. Ein von der variationsreihe erzeugter poissonscher punctprozess. *Annuaire Univ. Sofia Fac. Math. Méch.*, 77:79–94, 1971. Opublikowane w 1986 roku.
- [25] Z. Ignatov. Point processes generated by order statistics and their applications. In P. Bártfai, editor, *Point Processes and Queing problems*, volume 24 of *Coll. Math. Soc. János Bolyai*, pages 109–116. North Holland, Amsterdam, 1981.
- [26] O. Kallenberg. *Random measures*. Akademie-Verlag, Berlin; Academic Press, Inc., London, fourth edition, 1986. ISBN 0-12-394960-2.
- [27] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, 1997. ISBN 0-387-94957-7.
- [28] S. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009. ISBN 978-0-521-73182-9. doi: [10.1017/CBO9780511626630](https://doi.org/10.1017/CBO9780511626630). With a prologue by Peter W. Glynn.
- [29] H. N. Nagaraja. Record values and related statistics. A review. *Comm. Statist. Theory Methods*, 17(7):2223–2238, 1988. ISSN 0361-0926. doi: [10.1080/03610928808829743](https://doi.org/10.1080/03610928808829743).
- [30] V. B. Nevzorov. k th record times and their generalizations. *Zapiski Nauchn. Semin LOMI*, 153:115–121, 1986. English transl., *J. Soviet Math* **44** (1986) 510–515.

- [31] V. B. Nevzorov. Records. *Teor. Veroyatnost. i Primenen.*, 32(2):219–251, 1987. ISSN 0040-361X.
- [32] V. B. Nevzorov. *Records: mathematical theory*, volume 194 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 2001. ISBN 0-8218-1945-3. Translated from the Russian manuscript by D. M. Chibisov.
- [33] I. Pickands. The two-dimensional Poisson process and extremal processes. *J. Appl. Probability*, 8:745–756, 1971. ISSN 0021-9002.
- [34] S. D. Poisson. *Recherchés sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile*. Précédées des Regles Générales du Calcul des Probabilités, Bechelier, Imprimeur-Libraire pour des Mathematiques, la Physique, etc., Paris, 1837. With a bibliography of Rényi's works.
- [35] A. Rényi. Théorie des éléments saillants d'une suite d'observations. *Ann. Fac. Sci. Univ. Clermont-Ferrand No.*, 8:7–13, 1962. See also *Théorie des éléments saillants d'une suite d'observations*, Colloquim on Combinatorial Methods in Probability Theory (Math. Inst., Aarhus, Denmark August 1-10, 1962), pp. 104-117; and *On the extreme elements of observations*, Selected Papers of Alfred Rényi, vol. 3, Akademiai Kiado, Budapest 1976, pp. 50-65.
- [36] S. I. Resnick. *Extreme values, regular variation, and point processes*, volume 4 of *Applied Probability. A Series of the Applied Probability Trust*. Springer-Verlag, New York, 1987. ISBN 0-387-96481-9. doi: [10.1007/978-0-387-75953-1](https://doi.org/10.1007/978-0-387-75953-1).
- [37] S. I. Resnick. *Heavy-tail phenomena*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2007. ISBN 978-0-387-24272-9; 0-387-24272-4. Probabilistic and statistical modeling.
- [38] L. C. G. Rogers. Ignatov's theorem: an abbreviation of the proof of Engelen, Tommassen and Vervaat. Comment on: "Ignatov's theorem: a new and short proof" [J. Appl. Probab. **1988**, Special Vol. 25A, 229–236; MR0974584 (90a:60094)] by R. Engelen, P. Tommassen and W. Vervaat. *Adv. in Appl. Probab.*, 21(4):933–934, 1989. ISSN 0001-8678. doi: [10.2307/1427776](https://doi.org/10.2307/1427776).
- [39] S. M. Samuels. An all-at-once proof of Ignatov's theorem. In *Strategies for sequential search and selection in real time (Amherst, MA, 1990)*, volume 125 of *Contemp. Math.*, pages 231–237. Amer. Math. Soc., Providence, RI, 1992. doi: [10.1090/conm/125/1160623](https://doi.org/10.1090/conm/125/1160623).
- [40] R. Serfozo. *Basics of applied stochastic processes*. Probability and its Applications (New York). Springer-Verlag, Berlin, 2009. ISBN 978-3-540-89331-8. doi: [10.1007/978-3-540-89332-5](https://doi.org/10.1007/978-3-540-89332-5).
- [41] M. Q. Shahbaz, M. Ahsanullah, S. Hanif Shahbaz, and B. M. Al-Zahrani. *Ordered random variables: theory and applications*, volume 9 of *Atlantis Studies in Probability and Statistics*. Atlantis Press, Paris, 2016. ISBN 978-94-6239-224-3; 978-94-6239-225-0. doi: [10.2991/978-94-6239-225-0](https://doi.org/10.2991/978-94-6239-225-0).
- [42] G. Simons, Y.-C. Yao, and L. Yang. Doob, Ignatov and optional skipping. *Ann. Probab.*, 30(4):1933–1958, 2002. ISSN 0091-1798. doi: [10.1214/aop/1039548377](https://doi.org/10.1214/aop/1039548377).

-
- [43] A. J. Stam. Independent Poisson processes generated by record values and inter-record times. *Stochastic Process. Appl.*, 19(2):315–325, 1985. ISSN 0304-4149. doi: [10.1016/0304-4149\(85\)90033-X](https://doi.org/10.1016/0304-4149(85)90033-X).
- [44] A. Stuart. The efficiencies of tests of randomness against normal regression. *J. Amer. Statist. Assoc.*, 51:285–287, 1956. ISSN 0162-1459. URL [http://links.jstor.org/sici?sici=0162-1459\(195606\)51:274<285:TEOTOR>2.0.CO;2-Y&origin=MSN](http://links.jstor.org/sici?sici=0162-1459(195606)51:274<285:TEOTOR>2.0.CO;2-Y&origin=MSN).
- [45] A. Stuart. The efficiency of the records test for trend in normal regression. *J. Roy. Statist. Soc. Ser. B.*, 19:149–153, 1957. ISSN 0035-9246. URL [http://links.jstor.org/sici?sici=0035-9246\(1957\)19:1<149:TEOTRT>2.0.CO;2-N&origin=MSN](http://links.jstor.org/sici?sici=0035-9246(1957)19:1<149:TEOTRT>2.0.CO;2-N&origin=MSN).
- [46] M. N. Tata. On outstanding values in a sequence of random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 12:9–20, 1969. doi: [10.1007/BF00538520](https://doi.org/10.1007/BF00538520). MR 0247655.
- [47] P. Turán, editor. *Selected papers of Alfréd Rényi, Vol. III: 1962–1970*. Akadémiai Kiadó, Budapest, 1976. With a bibliography of Rényi’s works.
- [48] S. S. Wilks. Determination of sample sizes for setting tolerance limits. *Ann. Math. Statistics*, 12:91–96, 1941. ISSN 0003-4851. MR 0004451.
- [49] Y.-C. Yao. On independence of k -record processes: Ignatov’s theorem revisited. *Ann. Appl. Probab.*, 7(3):815–821, 1997. ISSN 1050-5164. doi: [10.1214/aoap/1034801255](https://doi.org/10.1214/aoap/1034801255). MR 1459272.

***K*-th Record Values and Their Basic Properties**

Wiesław Dziubdziela

Abstract In this paper, the k -th record value model is presented. The k -th record values has emerged as an important model of ordered random variables. They appears naturally in real life where one interested in successive K -th maximum observations. The k -th record values are formally defined by Dziubdziela i Kopociński (1976). The paper contains distributional theory for this model.

2010 *Klasyfikacja tematyczna AMS (2010)*: 60G70,60G30.

Słowa kluczowe: k -th record values, k -th record times , number of k -th record values, point process of k -th record values, representations of k -th record values.



Wiesław Dziubdziela po pięćdziesięciu latach zajmowania się teorią wartości ekstremalnych i rekordowych obecnie poświęca się pracy w ogródku.

WIESŁAW DZIUBDZIELA
POLSKIE TOWARZYSTWO MATEMATYCZNE
ODDZIAŁ KIELECKI
INSTYTUT MATEMATYKI UNIWERSYTETU JANA KOCHANOWSKIEGO W KIELCACH
UL. ŚWIĘTOKRZYSKA 15, KIELCE 25-406
E-mail: wiedziu@poczta.fm

Communicated by: Krzysztof Szajowski

(Received: 24th of March 2018; revised: 11th of May 2018)
