DIDACTICS
OF
MATHEMATICS

5-6(9-10)

The Publishing House
of the Wroclaw University of Economics
Wroclaw 2009

Editors
Janusz Łyko
Antoni Smoluk

Referee
Włodzimierz Odyniec
(The Hercen University, St Petersburg)

Proof reading
Joanna Szynal

Setting
Elżbieta Szlachcic

Cover design
Robert Mazurczyk

Front cover painting: W. Tank, *Sower*
(private collection)

PL ISSN 1733-7941

TABLE OF CONTENTS

JAN FLOREK, JACEK JUZWISZYN, ANDRZEJ MISZTAL, JERZY SACAŁA
O ciągu Ulama, równaniu Pella i rotacjach rynku finansowego
[*On Ulam sequence, Pell's equation and rotations of the financial market*] ... 5

MAREK BIERNACKI
Effectiveness of mathematical education
[*Skuteczność edukacji matematycznej*] 19

JAN FLOREK
Równania Cauchy'ego-Riemanna i przekształcenia konforemne
[*Cauchy-Riemann equations and conformal maps*] 33

PIOTR DNIESTRZAŃSKI, ANDRZEJ WILKOWSKI
O paradoksie Halla i rzucaniu monetą
[*On Hall's paradox and coin flipping*] 43

TADEUSZ JANASZAK
O kreśleniu wykresów funkcji wymiernych z użyciem programu Matlab
[*Some remarks about the construction of the rational function with the use of Matlab programme*] 53

ANDRZEJ WILKOWSKI
Notes on normal distribution
[*Uwagi o rozkładzie normalnym*] 71

WIKTOR EJSMONT
Production function as a measure of school education quality
[*Funkcja produkcji jako miernik jakości kształcenia szkoły*] 79

RAFAŁ KORZONEK
Uwagi o granicznych rozkładach ekstremalnych statystyk pozycyjnych
[*Selected issues on the limit distributions of extreme order statistics*] 89

TADEUSZ JANASZAK
O konieczności nauczania liczb rzeczywistych i trygonometrii hiperbolicznej w kontekście użycia programu Matlab
[*Some remarks about the necessity of teaching about complex numbers and hiperbolic trigonometry in the context of Matlab programme*] 99

WIKTOR EJSMONT
Efektywność nauczania we wrocławskich liceach
[*Efficiency of teaching at high schools in Wroclaw*] 111

ANTONI SMOLUK
Corrigendum I 129

ANTONI SMOLUK
Corrigendum II 131

D I D A C T I C S O F M A T H E M A T I C S

No. 5-6 (9-10) 2009

Rafał Korzonek
(Wrocław)

UWAGI O GRANICZNYCH ROZKŁADACH EKSTREMALNYCH STATYSTYK POZYCYJNYCH

Abstract. In many practical issues to deal with extreme situations is often more important than with common situations. It is the observation of cases of significant outliers from the central parameters are the most troublesome. In this article, which has a review character, presents selected issues of the theory of limit distributions of extreme order statistics. Addressed the most important results concerning this theory. Considered distributions limit of extreme order statistics and showing that they may only be three types.

Key words: statystyki pozycyjne, rozkłady graniczne statystyk pozycyjnych, rozkłady minimum i maksimum, badania niezawodności systemów.

1. Wstęp

W wielu zagadnieniach praktycznych radzenie sobie z sytuacjami ekstremalnymi jest często ważniejsze niż z sytuacjami powszechnie spotykanymi. To właśnie przypadki obserwacji znacznie odstających od parametrów centralnych są najbardziej dokuczliwe. Jak wiadomo, o prawdopodobieństwie powstania sytuacji ekstremalnych informują rozkłady maksimów i minimów. W artykule tym, mającym charakter przeglądowy, przedstawiono wybrane zagadnienia z teorii rozkładów granicznych ekstremalnych statystyk pozycyjnych. Zajęto się najważniejszymi wynikami dotyczącymi tej teorii. Zbadano rozkłady graniczne ekstremalnych statystyk pozycyjnych i pokazano, że mogą być one tylko trzech typów. Następnie oszacowano szybkość zbieżności rozkładu k-tej statystyki pozycyjnej do rozkładu granicznego.

Teoria ekstremalnych statystyk pozycyjnych jest powszechnie wykorzystywana. Jednym z głównych zastosowań tej teorii są badania niezawodności systemów elementów. Podstawowymi strukturami niezawodnościowymi są struktury szeregowe i równoległe. Jak wiadomo, struktura niezawodnościowa nazwana jest szeregową, jeżeli system jest zdatny wyłącznie wtedy, gdy zdatne są wszystkie jego elementy. Wówczas czas zdatności (przeżycia) takiego systemu T_S jest równy czasowi zdatności „najgorszego" elementu, zatem jest równy minimalnej statystyce pozycyjnej: $T_S = \min(T_1, T_2, ..., T_n)$, gdzie T_i – czas zdatności elementu i-tego. Analogicznie struktura niezawodnościowa nazwana jest równoległą, jeżeli system jest niezdatny wyłącznie wtedy, gdy niezdatne są wszystkie jego elementy. Wówczas czas zdatności takiego systemu jest równy czasowi zdatności „najlepszego" elementu, czyli $T_S = \max(T_1, T_2, ..., T_n)$. W praktyce pojawia się potrzeba wyznaczenia rozkładu granicznego funkcji niezawodności danego systemu. To zagadnienie pociąga za sobą wyznaczenie rozkładu granicznego maksymalnej i minimalnej statystyki pozycyjnej. Poniższe rozważania mają na celu przybliżyć teorię potrzebną do rozwiązania tego problemu.

2. Rozkłady graniczne statystyk pozycyjnych

Badanie rozkładów granicznych ekstremalnych statystyk pozycyjnych rozpoczęto już w latach dwudziestych minionego wieku, m.in. w pracach Frecheta oraz Fishera. Później Gniedenko udowodnił, że rozkłady graniczne maksimum ciągu zmiennych losowych mogą być jednego z trzech typów, a Smirnow uogólnił jego wyniki na przypadek k-tej statystyki ekstremalnej. W artykule przedstawiono podstawowe fakty tej teorii.

Zanim zajmiemy się badaniem rozkładów granicznych podajemy definicję statystyk pozycyjnych. W tym celu rozpatrzmy ciąg zmiennych losowych $(X_1, X_2, \ldots, X_n)$. Wówczas mamy:

Definicja 1. *Statystykę pozycyjną*

$$Z_k^{(n)},\ k = 1, 2, n,$$

nazywamy zmienną losową będącą funkcją wektora losowego $(X_1, X_2, \ldots, X_n)$, *określoną w następujący sposób*:

Dla każdego zdarzenia elementarnego ω *ciąg realizacji*

$$X_1(\omega) = x_1,\ X_2(\omega) = x_2,\ \ldots,\ X_n(\omega) = x_n$$

porządkujemy rosnąco i otrzymujemy $z_1 \le z_2 \le \ldots \le z_n$.

W tym ciągu z_k *jest realizacją zmiennej losowej* $Z_k^{(n)}$*, tzn.*

$$Z_k^{(n)}(\omega) = z_k .$$

Oczywiście

$$Z_n^{(n)} = \max\left(X_1, X_2, \ldots, X_n\right),$$

$$Z_1^{(n)} = \min\left(X_1, X_2, \ldots, X_n\right).$$

Definicja 2. *Zmienną losową*

$$Z_{n-k+1}^{(n)},$$

gdzie k jest ustaloną liczbą naturalną $k = 1, 2, \ldots, n$ *będziemy nazywać k-tą ekstremalną statystyką pozycyjną.*

Mając zdefiniowane statystyki pozycyjne, oznaczmy ich rozkład przez:

$$F_k^{(n)}(x) = P\left(Z_{n-k+1}^{(n)} < x\right), \quad k = 1, 2, \ldots, n.$$

Łatwo zauważyć następujący fakt:

Fakt 1. Jeżeli $(X_1, X_2, \ldots, X_n)$, są niezależnymi zmiennymi losowymi o jednakowym rozkładzie $F(x)$, to

$$F_k^{(n)}(x) = \sum_{r=0}^{k-1} \binom{n}{r} \left(1 - F(x)\right)^r \left(F(x)\right)^{n-r} = (n-k+1)\binom{n}{k-1} \int_0^{F(x)} u^{n-k}(1-u)^{k-1}\, du .$$

Na podstawie powyższego faktu możemy w prosty sposób wyznaczyć rozkłady maksimum i minimum ciągu niezależnych zmiennych losowych. Mając dane rozkłady ekstremalnych statystyk pozycyjnych, zajęto się badaniem ich rozkładów granicznych. Okazuje się, że jeżeli ciąg $\max(X_1, X_2, \ldots, X_n)$ ma niezdegenerowaną dystrybuantę graniczną, to jest ona jednej z trzech wymienionych w twierdzeniu 2.1 postaci. Poniższe twierdzenie zostało zaczerpnięte z pracy Czekały (M. Czekała (2001)), natomiast dowody tego twierdzenia można znaleźć w pracach Resnicka i Gniedenki (S.I. Resnick (1987); B.W. Gniedenko (1943)).

Twierdzenie 1. *Niech*

$$\omega(F) = \sup\{x : F(x) < 1\}$$

oraz

$$\alpha(F) = \inf\{x : F(x) > 0\}.$$

Niech zmienne losowe X_i (i = 1, 2, …) będą niezależne o dystrybuancie F.

a) Załóżmy, że $\omega(F) = +\infty$ oraz że istnieje taka stała $\gamma > 0$, że dla każdego $x > 0$ spełniony jest warunek

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\gamma}.$$

Wówczas istnieje ciąg stałych b_n $(b_n > 0)$ taki, że

$$\lim_{n\to\infty} P\{\max(X_1, X_2, \ldots, X_n) < b_n x\} = H_{1,\gamma}(x),$$

gdzie ciąg b_n może być postaci

$$b_n = \inf\left\{x : 1 - F(x) \le \frac{1}{n}\right\},$$

zaś

$$H_{1,\gamma}(x) = \begin{cases} \exp(-x^{-\gamma}) & \text{dla } x > 0, \\ 0 & \text{dla } x \le 0. \end{cases}$$

b) Załóżmy, że $f(x) > 0$ w pewnym przedziale (x_0, x_F) oraz $f(x) = 0$ dla $x > x_F$ (f jest gęstością dystrybuanty F).

Jeżeli

$$\lim_{t\uparrow x_F} \frac{(x_F - t)f(t)}{1-F(t)} = \gamma > 0,$$

to

$$\lim_{n\to\infty} P\{\max(X_1, X_2, \ldots, X_n) < a_n + b_n x\} = H_{2,\gamma}(x),$$

gdzie

$$a_n = \omega(F)$$

i

$$b_n = \omega(F) - \inf\left\{x : 1 - F(x) \le \frac{1}{n}\right\}$$

oraz

$$H_{2,\gamma}(x) = \begin{cases} 1 & \text{dla } x \geq 0, \\ \exp(-(-x)^{\gamma}) & \text{dla } x < 0. \end{cases}$$

c) Załóżmy, że f ma ujemną pochodną f' dla wszystkich x w pewnym przedziale

$$(x_0, x_F), \quad (x_F \leq \infty) \quad \text{oraz} \quad f(x) = 0$$

dla $x > x_F$ (zauważmy, że dopuszczamy tutaj przedział $(x_0, +\infty)$). Jeżeli

$$\lim_{t \uparrow x_F} \frac{f'(t)(1-F(t))}{f^2(t)} = -1,$$

to

$$\lim_{n\to\infty} P\{\max(X_1, X_2, \ldots, X_n) < a_n + b_n x\} = H_{3,0}(x),$$

gdzie ciągi liczbowe a_n i b_n mogą być postaci

$$a_n = \inf\left\{x : 1 - F(x) \leq \frac{1}{n}\right\}$$

oraz

$$b_n = R(a_n),$$

gdzie

$$R(t) = (1 - F(t))^{-1} \int_t^{\omega(F)} (1 - F(t))dy$$

i

$$H_{3,0}(x) = \exp(-\exp(-x)), \; x \in R.$$

Zauważmy, że powyższe twierdzenie podaje warunki wystarczające, aby dystrybuanta graniczna miała jedną z podanych trzech postaci. W literaturze te trzy klasy dystrybuant nazywamy odpowiednio: dystrybuantami typu I, typu II i typu III.

Analogicznie, korzystając z równości:

$$\max(X_1, X_2, \ldots, X_n) = -\min(-X_1, -X_2, \ldots, -X_n),$$

można wykazać, że dystrybuantami granicznymi zmiennych

$$\min(X_1, X_2, \cdots, X_n)$$

mogą być:

Typ I: $$L_{1,\gamma}(x) = \begin{cases} 1 - \exp(-(-x)^{-\gamma}), & x < 0, \\ 1, & x > 0. \end{cases}$$

Typ II: $$L_{2,\gamma}(x) = \begin{cases} 1 - \exp(-x^{-\gamma}), & x > 0, \\ 0, & x < 0. \end{cases}$$

Typ III: $$L_{3,0}(x) = 1 - \exp(-\exp(x)), \quad x \in R.$$

Aby zilustrować przydatność twierdzenia, rozważmy następujący przykład.

Przykład 1. Można wykazać, że do obszaru przyciągania dystrybuanty typu I należy dystrybuanta postaci

$$F(x) = \frac{1}{2} + \frac{1}{\pi} arctgx, \quad x \in R,$$

czyli dystrybuanta rozkładu Cauchy''ego.

Elementarnie wykazuje się, że

$$\lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = \frac{1}{x},$$

a zatem

$$\lim_{n \to \infty} P\{\ \max(X_1, X_2,\ \ldots,\ X_n) < b_n\} = \exp\left(-\frac{1}{x}\right),\ x > 0.$$

Stałe b_n można wyznaczyć z relacji

$$b_n = tg\left(\frac{\pi}{2} - \frac{\pi}{n}\right).$$

Przykładem innego rozkładu należącego do obszary przyciągania typu I jest np.: rozkład Pareto. Do obszaru przyciągania dystrybuanty typu II należą m.in.: rozkład jednostajny oraz standardowy rozkład Plancka, natomiast do obszaru przyciągania dystrybuanty typu III należą m.in.:

rozkład normalny $N(0, 1)$ oraz rozkład wykładniczy. Okazuje się także, że nie dla wszystkich rozkładów istnieje zdegenerowana dystrybuanta graniczna jednego z tych trzech typów. Przykładami mogą być rozkład geometryczny oraz rozkład Poissona, dla których przy dowolnym wyborze stałych a_n i b_n prawdopodobieństwo zdarzenia

$$\max(X_1, X_2, \dots, X_n) < a_n + b_n x$$

wynosi 0 lub 1. Jak wspomniano na początku, Smirnow uogólnił teorię granicznych rozkładów maksimum na przypadek k-tej statystyki ekstremalnej. Mówi o tym poniższe twierdzenie, zaczerpnięte z pracy Czekały (M. Czekała (2001)).

Twierdzenie 2 (rozkład *k*-tych statystyk pozycyjnych). *Jeżeli dla pewnych ciągów a_n, $b_n > 0$ oraz $k > 1$ zachodzi relacja*

$$\lim_{n\to\infty} P\{X_{n-k+1:n} < a_n + b_n x\} = H^{(k)}(x),$$

wówczas dla

$$Z_n = \max(X_1, X_2, \dots, X_n)$$

zachodzi

$$\lim_{n\to\infty} P\{Z_n < a_n + b_n x\} = H(x),$$

gdzie $H(x)$ jest jedną z dystrybuant maksymalnie stabilnych (tzn. $H_{1,\gamma}(x)$, $H_{2,\gamma}(x)$ oraz $H_{3,0}(x)$). Ponadto funkcja $H^{(k)}(x)$ ma postać:

$$H^{(k)}(x) = H(x) \cdot \sum_{t=0}^{k-1} \frac{1}{t!} \left(-\ln H(x)\right)^t .$$

Twierdzenie odwrotne jest również prawdziwe. Zatem można stwierdzić, że przynależność dystrybuanty granicznej do danego obszaru przyciągania determinuje postać rozkładu k-tej statystyki pozycyjnej. Analogicznie twierdzenie zachodzi dla minimów (M. Czekała (2001)).

Zajmiemy się teraz szacowaniem szybkości zbieżności rozkładów statystyk pozycyjnych do rozkładów granicznych. Okazuje się, że zagadnienie to sprowadza się do szacowania szybkości zbieżności rozkładu dwumianowego do rozkładu Poissona. Poniższe dwa twierdzenia znajdujemy w pracy Dziubdzieli (W. Dziubdziela (1977)).

Twierdzenie 3. *Jeżeli*

$$B(x;n,p)=\sum_{0\le r<x}\binom{n}{r}p^r(1-p)^{n-r}$$

jest dystrybuantą rozkładu dwumianowego oraz

$$0<p<\frac{1}{2},$$

wówczas

$$\sup_{-\infty<x<\infty}\left|B(x;n,p)-P(x,np)+\frac{1}{2}np^2P^*(x,np)\right|\le\frac{1}{n^2}C(n,p),$$

gdzie

$$C(n,p)=\frac{1}{2}\pi\exp(2np)(np)^3\left[\frac{4}{3(1-2p)}+\right.$$

$$\left.+\left(\frac{16}{9}np^3\frac{1}{(1-2p)^2}+\frac{8}{3}np^2\frac{1}{1-2p}+np\right)\exp\left(2np^2\left(1+\frac{4}{3}p\frac{1}{1-2p}\right)\right)\right].$$

Następne twierdzenie daje oszacowanie szybkości zbieżności w twierdzeniu granicznym dla ekstremalnych statystyk pozycyjnych.

Twierdzenie 4. *Jeżeli*

$$\frac{1}{2}<F(y_n)<1,$$

to dla każdego $k=1,2,\ldots$ *mamy*

$$\left|F_k^{(n)}(y_n)-\Phi_k(x)+\frac{1}{2}np^2(y_n)P^*(k,np(y_n))\right|\le$$

$$\le\frac{1}{n^2}C(n,p(y_n))+\left|\frac{1}{(k-1)!}\int_{np(y_n)}^{L(x)}u^{k-1}e^{-u}du\right|,$$

gdzie

$$p(y_n)=1-F(y_n),\ L(x)=\lim_{n\to\infty}np(y_n).$$

Dowody powyższych twierdzeń także można znaleźć w pracy Dziubdzieli (W. Dziubdziela (1977)).

3. Podsumowanie

Powyższe rozważania miały na celu przybliżenie podstawowych twierdzeń i faktów dotyczących teorii rozkładów granicznych ekstremalnych statystyk pozycyjnych. Teoria ta jest często wykorzystywana nie tylko w wspomnianych badaniach niezawodności, ale również w wielu innych zagadnieniach praktycznych.

Przykładami takich zagadnień opisanych w literaturze mogą być:

- badanie wytrzymałości materiałów (E.J. Gumbel (1962)),
- badanie maksymalnego czasu czekania na obsługę w teorii obsługi masowej (W. Whitt (1974)),
- badanie maksymalnej długości kolejek w teorii obsługi masowej (C.W. Anderson (1970)),
- szacowanie maksymalnych przyborów wody w rzece (E.J. Gumbel (1962)),
- obliczanie nośności konstrukcji budowlanych (B. Kopociński, Z. Kowal (1972)),
- weryfikacja złożeń modelu Blacka-Scholesa w ekonometrii (M. Czekała (2001)).

Literatura

C.W. Anderson (1970). *Extreme value theory for a class of discrete distributions with applications to some stochastic processes*. Journal of Applied Probability 7.

M. Czekała (2001). *Statystyki pozycyjne w modelowaniu ekonometrycznym*. Wrocław.

W. Dziubdziela (1977). *Rozkłady graniczne ekstremalnych statystyk pozycyjnych.* Matematyka stosowana.

B.W. Gniedenko (1943). *Sur la distribution limite du terme maximum d'une serie aleatorie*. Annals of Mathematics 44.

E.J. Gumbel (1962). *Statistics of extremes*. New York.

R. Harris (1970). *An application of extreme value theory*. Annals of Mathematical Statistics 41.

B. Kopociński (1973). *Zarys teorii odnowy i niezawodności*. Warszawa.

B. Kopociński, Z. Kowal (1972). *Losowa nośność graniczna konstrukcji o dwóch minimalnych krytycznych zbiorach elementów mających elementy wspólne*. Archiwum Inżynierii Lądowej 18.

S.I. Resnick (1987). *Extreme values, regular variation*, *and point processes*. New York.

W. Whitt (1974). *Heavy traffic limit theorems for queues: A survey*. Lecture Notes In Economics and Mathematical Systems 98.

M.R. Leadbetter, G. Lindgren, H. Rootzen (1986). *Extremes and related properties of random sequences and processes.* Berlin – New York.